STUDIES ON THE ASSEMBLY OF MEMBRANE PROTEINS: TOPOLOGICAL ORGANIZATION AND INTERACTIONS OF THE MEMBRANE PROTEINS OF VESICULAR STOMATITIS VIRUS

By

JOHN P. CAPONE, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

September, 1983

TOPOLOGY AND INTERACTIONS OF THE MEMBRANE PROTEINS OF VESICULAR STOMATITIS VIRUS.

DOCTOR OF PHILOSOPHY
(Biochemistry)

McMaster University
Hamilton, Ontario

TITLE: Studies on the Assembly of Membrane Proteins: Topological Organization and Interactions of the Membrane Proteins of Vesicular Stomatitis Virus

AUTHOR: John P. Capone, B.Sc. (University of Western Ontario)

SUPERVISOR: Dr. Hara P. Ghosh

NUMBER OF PAGES: xix, 426

ABSTRACT

The purpose of this study was to investigate the assembly of membrane proteins by examining the topological organization and interactions of the membrane proteins G and M of vesicular stomatitis virus (VSV).

Exhaustive proteolytic digestion of VSV resulted in the complete degradation of G protein, however, a small peptide fragment having an apparent molecular weight of approximately 9,000 D remained associated with the viral membrane. Characterization of this membrane embedded fragment by tryptic peptide analysis and amino acid sequence determination demonstrated that it was derived from the COOH-terminal end of G protein. Furthermore, comparison of the partial amino acid sequence of this peptide fragment with the predicted amino acid sequence of G protein indicated that this peptide contains an uninterrupted hydrophobic domain of sufficient length to span the viral envelope. Thus, the G protein of VSV is anchored in the viral membrane by a hydrophobic domain located near the COOH-terminus.

In addition to the COOH-terminus, the NH_2-terminus of G protein was also shown to be protected from proteolytic attack by the integrity of the viral envelope. This may be related to the tertiary structure of G protein that is imposed by the viral membrane or the fact that the NH_2-terminus may be in close proximity to the surface of the membrane and thus protected from proteolytic attack.

It has recently been reported that the VSV G protein contains

tightly bound fatty acid residues. Proteolytic digestion of VSV labeled with [^{3}H]-palmitic acid demonstrated that all the fatty acid residues present in G protein are localized exclusively in the membrane interacting domain. Thus, the lipophilic fatty acids in conjunction with the hydrophobic domain may play an important role in the interaction of G protein with the viral envelope.

VSV grown in the presence of ω-[9-^{3}H] diazirinophenoxy nonanoate resulted in the biosynthetic incorporation of this photoreactive fatty acid into the viral phospholipids as well as into the membrane anchoring domain of G protein. Photolysis of the virus resulted in extensive phospholipid crosslinking to the G protein but not to the M protein. This confirms that the COOH-terminal region of G protein is in intimate contact with the hydrophobic core of the lipid bilayer and directly demonstrates that M protein does not penetrate the viral membrane to a significant extent. In addition, a new product was obtained following photolysis and identified as a G-G dimer on the basis of its molecular weight and immunoreactivity. This product arose presumably from protein crosslinking mediated by the photoreactive fatty acid attached to G protein. Thus, the biosynthetic incorporation of this photoreactive fatty acid makes it possible to not only identify integral membrane proteins but also to make photoaffinity probes of membrane proteins which are normally fatty acid acylated.

The nature of the association of M protein with membranes was examined by reconstituting the purified protein into artif- -al phospholipid vesicles. The M protein was shown to have a strong

affinity for such vesicles and the association, once made, could not be disrupted by salt treatment. This suggests that the interaction was at least partly hydrophobic in nature. The nucleocapsid protein N was shown to have no affinity for artificial lipid vesicles, however, it could associate with vesicles in the presence of M protein. The same results were obtained when in vitro synthesized proteins were used for reconstitution. These results support the concept that the M protein is involved in virus maturation through its ability to interact with both the plasma membrane and the ribonucleoprotein core containing N protein.

The maturation of membrane proteins is thought to be directed by discrete polypeptide domains present in the protein which are recognized and decoded by specialized mechanisims of the cell. As an approach to study the function of these various polypeptide domains, a hybrid gene was constructed from plasmids containing cDNA copies of the complete coding sequences of the G and M mRNAs. A chimeric gene, which contained the signal sequence coding region of G protein fused to the bulk of the coding sequence specific for M, was constructed and placed into an expression vector. Introduction of this hybrid gene into mammalian cells by DNA mediated gene transfer resulted in the synthesis of a 36,000 D polypeptide immunoreactive with anti-M antibody. This plasmid should thus be useful for the functional examination of the signal sequence of G protein. As well, it should be useful for examining polypeptide domains which are involved in glycosylation since M protein contains a cryptic glycosylation target

site.

The studies reported in this thesis have been presented, in part, in the following publications.

1. Ghosh, H.P., Capone, J., Irving, R., Kotwal, G., Hofmann, T., Levine, G., Rachubinski, R., Shore, G., and Bergeron, J. (1981) Viral Membrane Proteins: Assembly and Structure. In Replication of Negative Strand Viruses. D.H. Bishop and R.W. Compans ed. p. 655 Elsivier North Holland.
2. Capone, J., Toneguzzo, F., and Ghosh, H.P. (1982) Synthesis and Assembly of Membrane Glycoproteins: Membrane Anchoring COOH-Terminal Domain of Vesicular Stomatitis Virus Envelope Glycoprotein G contains Fatty Acids. J. Biol. Chem. 257, 16
3. Leblanc, P., Capone, J., and Gerber, G.E. (1982) Synthesis and Biosynthetic Utilization of Radioactive Photoreactive Fatty Acids. J. Biol. Chem. 257, 14586
4. Capone, J., Leblanc, P., Gerber, G.E., and Ghosh, H.P. (1983) Localization of Membrane Proteins by the use of a Photoreactive Fatty Acid Incorporated in vivo into Vesicular Stomatitis Virus. J. Biol. Chem. 258, 1395.
5. Kotwal, G., Capone, J., Irving, R., Toneguzzo, F., Bilan, P., Rhee, S., Hofmann, T., and Ghosh, H.P. (1983) Viral Membrane Glycoproteins: Comparison of the Amino Terminal Amino Acid Sequences of the Precursor and Mature Glycoproteins from different Serotypes of Vesicular Stomatitis Virus. Virology (in press).

ACKNOWLEDGEMENTS

I wish to express my appreciation and gratitude to my supervisor, Dr. Hara P. Ghosh, for his patience, encouragement, and inspiration throughout the course of this work.

I thank the members of my supervisory committee, Dr. L. Prevec and Dr. G.E. Gerber for stimulating discussions and helpful criticisms. In particular I am most grateful to Dr. G.E. Gerber and P. Leblanc for providing material which made some of the studies reported here possible.

I wish to thank Dr. J.K. Rose and Dr. R. Kaufman for providing recombinant plasmids.

I would also like to express my thanks to the members of the Ghosh laboratory, both past and present, for many discussions and for making my stay a pleasant one.

I especially wish to thank my wife Julie, without whose love, encouragement, and willingness to sacrifice this work would not have been possible, and my parents, whose efforts, perhaps more than my own, have made the attainment of this goal a reality.

The financial support of the Ontario Ministry of Colleges and Universities, and McMaster University is gratefully acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A	adenine or adenosine
A_{260nm} unit	An amount of material that has an absorbance of 1.0 at 260 nm when dissolved in 1 ml water and measured with a 1 cm light path.

Amino Acids:

Ala, A,	alanine
Arg, R,	arginine
Asn, N,	asparginine
Asp, D,	aspartic acid
Asx, B,	asparginine/aspartic acid
Cys, C,	cysteine
Gln, Q,	glutamine
Glu, E,	glutamic acid
Glx, Z,	glutamine/glutamic acid
Gly, G,	glycine
His, H,	histidine
Ile, I,	isoleucine
Leu, L,	leucine
Lys, K,	lysine
Met, M,	methionine
Phe, F,	phenylalanine
Pro, P,	proline
Ser, S,	serine
Thr, T,	threonine
Trp, W,	tryptophan
Tyr, Y,	tyrosine
Val, V,	valine

ATP	adenosine 5' triphosphate
ATZ	anilinothiozolinone

AUFS	absorbance units, full scale
BHK cells	Baby Hamster Kidney Cells
bp	base pair
BSA	Bovine Serum Albumin
C	cytosine
CEF cells	Chick Embryo Fibroblast Cells
CHO cells	Chinese Hamster Ovary Cells
cpm	counts per minute
D	dalton
DAP	diazirinophenoxy
DFP	diisopropylphosphofluridate
DMBA	N-N-dimethyl amino benzylamine
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
cDNA	complementary deoxyribonucleic acid
DOC	deoxycholate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol bis(β-aminoethylether) N,N,N',N' tetraacetic acid
FBS	fetal bovine serum
FCS	fetal calf serum
dFCS	dialyzed fetal calf serum
Fuc	fucose
G	guanine or guanosine
Gal	galactose
Gluc	glucose

GluNAc	N-acetylglucosamine
GTP	guanosine 5' triphosphate
Hepes	N-2-hydroxyethyl piperazine-N'-2-ethane sulfonic acid
HFBA	heptafluorobutyric acid
HPLC	high pressure liquid chromatography
Man	mannose
MEM	minimal essential medium
moi	multiplicity of infection
NBCS	new born calf serum
NeuNAc	N-acetylneuraminic acid
PBS	phosphate buffered saline
PC	phosphatidylcholine
PE	phosphatidylethanolamine
PFU	plaque forming unit
Pipes	piperazine-N-N' bis(2-ethane sulfonic acid)
PITC	phenylisothiocyanate
PMSF	phenylmethylsulfonyl fluoride
PPO	2,5-diphenyloxazole
PS	phosphatidylserine
PTH	phenylthiohydantoin
Quadrol	N,N,N',N'-tetrakis(2-hydroxypropyl) ethylenediamine trifluoroacetic acid
RER	rough endoplasmic reticulum
RNA	ribonucleic acid
mRNA	messenger ribonucleic acid
tRNA	transfer ribonucleic acid

RNP	ribonucleoprotein
SDS	sodium dodecyl sulfate
SFV	Semliki Forest Virus
T	thymine or thymidine
TCA	trichloroacetic acid
TEMED	N,N,N',N'-tetramethylethylene diamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
TPCK	tolylsulfonyl phenylalanyl chloromethyl ketone
Tris	Tris(hydroxymethyl)aminomethane
U	Uracil or uridine
VSV	Vesicular Stomatitis Virus

Introduction

The rapid progress in recent years of our knowledge concerning the high degree of organizational and functional complexity of eukaryotic organisms has been possible in part from the wide use of animal viruses as simple, well defined model systems.

Viruses are obligate parasites since the limited amount of information encoded in their genomes makes them almost entirely dependant on the host cell machinery for their replication and metabolic requirements. As such, the processes involved in the replication, synthesis, and assembly of virion components is a reflection of normal functions carried out by the cell.

The approximately 600 known animal viruses exhibit a wide spectrum of complexity relating to differences in strategies of replication and maturation. Viruses can be broadly classified on the basis of the chemical nature of their nucleic acid (DNA or RNA; single or double stranded) and whether or not they contain a lipid envelope. Because of the increasing awareness of the importance of membranes in the biology of the cell, a great deal of attention has been focused on lipid containing viruses. Since their membranes are acquired by budding from cellular membranes, structure, composition, and assembly of the viral envelope reflects quantitatively and qualitatively that of the host cell membrane from which the virus matures.

Enveloped animal viruses can be classified into two groups, i) those that acquire their membranes by budding from cytoplasmic membranes such as the plasma membrane or the endoplasmic reticulum, and ii) those that assemble their membranes de novo or as in the case of the herpesviruses, acquire their lipid by budding from the nuclear membrane (Patzer et al., 1979). The most extensively studied enveloped viruses belong to the former group and include the rhabdoviruses, togaviruses, paramyxoviruses, orthomyxoviruses and the retroviruses. Members of these virus groups are surrounded by a cytoplasmically derived unit membrane containing 1-3 types of virus specified glycoprotein species (Lenard, 1978). In addition, the rhabdo, paramyxo, and orthomyxo groups contain a nonglycosylated matrix protein in close association with the viral membrane.

The work described in this thesis concerns the study of the membrane proteins of vesicular stomatitis virus, a negative stranded RNA containing enveloped virus belonging to the rhabdovirus group. VSV serves as a model virus to study membrane assembly and structure since the viral envelope contains only two species of protein; the glycoprotein G, and the nonglycosylated matrix protein M (Wagner, 1975). Infection of cultured cells with VSV results in a rapid shut off of host protein synthesis (Petric and Prevec, 1970; Wagner et al., 1970) thus membrane protein biogenesis becomes exclusively virus directed and virally specialized. In addition, the broad host range specificity and high virus yield affords large quantities of purified

material suitable for the study of many aspects of membrane assembly and viral morphogenesis.

1. Vesicular Stomatitis Virus

1.1 Morphology and composition

Rhabdoviruses are a group of highly infectious agents of both plants and animals whose classification was originally based on their rod shaped morphology (rhabdo = rod). The molecular biology of this virus group is based almost entirely on the prototype, vesicular stomatitis virus (VSV). VSV is pathogenic to cattle and horses, causing diseases similar to foot and mouth disease (Wagner, 1975).

When analyzed by negative contrast electron microscopy, VSV appears as an elongated, bullet shaped particle (B-particle) approximately 170 nm in length and 70 nm in diameter (Knudson, 1973; Nakai and Howatson, 1968; Simpson and Hauser, 1966). In addition to the bullet shaped morphology for the virus, a bacilliform morphology has been proposed (Orenstein et al., 1976).

Electron microscopic visualization of the virus demonstrates that it is composed of two distinct structural components; the outer surface and the internal core. The outer surface consists of a unit membrane from which radiate surface projections or peplomers of approximately 10 nm in length. These surface spikes appear to consist of hollow knobs at the end of short stalks (Nakai and Howatson, 1968; Brown et al., 1974). The lipoprotein envelope surrounds an internal core containing RNA and protein and having a

buoyant density of 1.31 gm/ml (Kang and Prevec, 1969; Wagner et al., 1969). The ribonucleoprotein core (RNP) appears to be helical, giving the appearance in negative stain of approximately 34 characteristic cross striations (Nakai and Howatson, 1968).

Chemical analysis of purified virions has shown that the virion particle contains 3% RNA, 64% protein, 13% carbohydrate, and 20% lipid by weight (McSharry and Wagner, 1971). Purified VSV contains five distinct structural polypeptides identifiable by SDS polyacrylamide gel electrophoresis and which are all specified by the viral genome (Wagner et al., 1972). These proteins have been designated L, G, NS, N, and M.

1.2 Structural Components and their Function

1.2.1 RNA

The genome of VSV is composed of a single RNA strand of negative polarity (complementary to the viral mRNA) having a molecular weight of approximately 4×10^6 D and consisting of approximately 11,000 nucleotides (Repik and Bishop, 1973; McGeoch, 1980). The RNA is present in the RNP in association with the three viral proteins L, NS, and N. On the average, each nucleocapsid contains a single RNA molecule complexed with 60 molecules of L protein, 230 molecules of NS, and 2,000 molecules of N protein (Wagner, 1975). L and NS can be readily dissociated from the RNP by treatment with high salt concentrations, however, the N protein remains tightly bound to the RNA under these conditions (Emerson and Wagner, 1973; Emerson and Yu, 1975). Electron microscopy of extended

unwound nucleocapsids has demonstrated that the N protein is regularily spaced at 3.5 nm intervals along the entire length of the RNA (Nakai and Howatson, 1968).

The negative stranded RNA molecule serves as the template for the production of the five viral specific mRNA species as well as the template for the production of full length progeny RNA molecules (Wagner, 1975).

1.2.2 Viral Proteins

(a) L Protein

L protein is the largest protein present in VSV and has an apparent molecular weight of approximately 190,000 D as determined by polyacrylamide gel electrophoresis (Wagner et al., 1972). As noted above, the L protein is found in association with nucleocapsid where together with the NS protein it serves as the endogenous RNA dependant RNA polymerase (Emerson and Wagner, 1972, 1973; Emerson and Yu, 1975; Inblum and Wagner, 1975; Hunt et al., 1976).

(b) NS Protein

The NS protein is found in large amounts in the infected cell but is only a minor component in the completed virus (Wagner et al., 1975). The NS protein is composed of a single polypeptide of 222 amino acids (25,110 D) as predicted from the mRNA nucleotide sequence determined from cDNA clones (Gallione et al., 1981). NS protein, however, migrates on SDS polyacrylamide gels with an apparent molecular weight in the range of 40,000 to 52,000 D. NS is the major phosphoprotein present in VSV and this led early investigators to

postulate that the phosphorylation was responsible for the aberrant electrophoretic mobility (Sokol and Clark, 1973; Inblum and Wagner, 1974). This does not appear to be the case since dephosphorylation of NS increases rather than decreases its mobility on SDS containing gels (Hsu and Kingsbury, 1980). The predicted amino acid sequence of NS as deduced from cDNA clones of the mRNA reveals a large cluster of negatively charged residues. Variable binding of SDS to this region may account for the anamolous electrophoretic behaviour (Gallione et al., 1981).

NS protein along with L protein is responsible for the virus specific mRNA transcription (Emerson and Yu, 1975). Reconstitution experiments have demonstrated that the binding of L protein to the ribonucleoprotein template is dependant on the presence of NS (Mellon and Emerson, 1978).

NS exists in at least two forms, NS I and NS II, which differ in their degree of phosphorylation (Clinton et al., 1978, 1979; Kingsford and Emerson, 1980). While the least phosphorylated form (NS I) was found associated with VSV core particles (Clinton et al., 1978), the more highly phosphorylated species was found to be more active in a reconstituted transcriptase assay (Kingsford and Emerson, 1980). Thus, phosphorylation-dephosphorylation may affect the ability of NS to regulate VSV RNA synthesis.

It has been postulated that NS has a dual role in that it may also function as a replicase. This has been suggested from studies with ts mutants that have demonstrated an interdependance between

transcription and replication suggesting that there is a subunit common to both pathways (Lesnaw and Reichman, 1975). More recently, using a ts mutant of VSV that is replication defective at the restrictive temperature, it was shown that the ts lesion resides in the NS protein. This supports the suggestion that NS protein plays a role in the replicative pathway (Lesnaw et al., 1979).

(c) N Protein

N protein is the major structural protein of VSV and has an apparent molecular weight of approximately 50,000 D as judged by SDS polyacrylamide gel electrophoresis (Kang and Prevec, 1969). Sequence analysis of the mRNA from a cDNA clone predicts a polypeptide of 422 amino acids (Gallione et al., 1981).

The N protein remains tightly associated with the viral RNA under conditions that release the NS and L proteins. The N-RNA complex serves as the template for transcription. Although N protein itself has no transcriptase activity, it is essential for this function to occur since deproteinized viral RNA cannot serve as a transcription template (Emerson and Wagner, 1973; Wagner, 1975).

Recent studies suggest that N protein may control the balance between transcription and replicative RNA synthesis in the infected cell (Blumberg et al., 1981; Lazzarini et al., 1981).

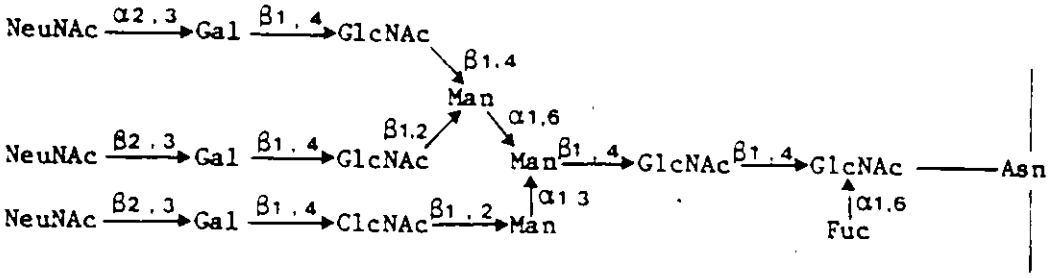
(d) G Protein

The G protein has an apparent molecular weight of 69,000 D and is found exclusively in association with the viral envelope (Wagner, 1975). The nucleic acid sequence of the G mRNA predicts a

polypeptide of 511 amino acids (Rose and Gallione, 1981).

G protein is the only glycosylated protein species present in the virion (Burge and Huang, 1970; Wagner et al., 1970; Mudd and Summers, 1970a) with the carbohydrate moiety accounting for approximately 10% of its weight (Etchison and Holland, 1974a, 1974b).

G protein contains two apparently identical Asn linked complex oligosacchardies of the type



(Etchison et al., 1977; Reading et al., 1978; Li et al., 1978).

These oligosaccharide chains are located at fractional distances of 0.35 and 0.66 from the amino terminus (Rose and Gallione, 1981). In addition to the major oligosaccharide chains, G protein contains minor amounts of sialic acid, fucose, and mannose attached to the polypeptide backbone (Kingsford et al., 1980).

It is apparent from various studies that while host cell glycosyl transferases are responsible for the synthesis of the oligosaccharide moieties, it is specified by information present in the G protein molecule (Sefton, 1976; Chatis and Morrison, 1981; Robertson et al., 1982) although there are minor differences in

glcosylation patterns which are host cell dependent (Etchison and Holland, 1974b; Moyer and Summers, 1974a, Etchison et al., 1981). These differences have been generally attributed to differences in sialic acid content.

In addition to the carboydrate moieties, it has been demonstrated that G protein contains 1-2 molecules of fatty acid that appear to be directly linked to the polypeptide backbone (Schmidt and Schlesinger, 1979).

G protein is present on the external surface of the viral envelope and forms the projections or spikes observed by electron microscopy of negatively stained preparations. This has been demonstrated by a variety of techniques. Surface labelling of VSV under proper conditions by lactoperoxidase catalyzed radioiodination labels only the G protein (McSharry, 1977), while digestion of VSV with a variety of proteases to remove the external spikes show that only G protein is susceptable to proteolytic degradation (Mudd, 1974; Schloemer and Wagner, 1975a). The oligosaccharide moieties are also external to the bilayer as demonstrated by their sensitivity to neuraminidase or a mixture of glycosidases (Schloemer and Wagner, 1974; Cartwright and Brown, 1977; Capone et al., 1982). Evidence that G protein actually penetrates the viral membrane was provided from the observation that a small peptide remains associated with the envelope of the virion after proteolytic removal of the protruding spikes (Schloemer and Wagner, 1975a; Mudd, 1974). This fragment is of sufficient length to span the membrane. The conclusion that

G protein actually spans the membrane has come from studies utilizing bifunctional crosslinking reagents that have shown that G protein can be crosslinked to internal viral components such as N and M protein (Mudd and Swanson, 1978; Duvobi and Wagner, 1977; Zakowski and Wagner, 1980). These same studies have indicated that G protein may exist in the membrane in oligomeric forms. The general picture then is that G protein exists in the viral envelope as a transmembranal protein with the bulk of its mass including the carbohydrate residues on the external surface of the bilayer and attached to the membrane by a small polypeptide fragment which can interact with internal viral components.

G protein is the major antigenic determinant of VSV (Kang and Prevec, 1970; Kelly et al., 1972; Wiktor et al., 1973) and plays several key roles in the life cycle of the virus. G protein is responsible for both the binding of the virus to host cells and inducing viral uptake. Virus particles from which the G protein is removed by proteolytic digestion are not infectious, however, the addition of G protein to these particles restores infectivity (Bishop et al., 1975a). In addition, G protein appears to play a role in the effecient assembly and maturation of the virus although its presence is not essential for budding to occur (Schnitzer and Lodish, 1980).

The function of the carbohydrate residues is not clearly understood. There seems to be a loss of infectivity following removal of sialic acid (Schloemer and Wagner, 1974) with recovery of infectivity following resialyation, however, other studies have

demonstrated that sialic acid residues are not essential for infectivity (Robertson et al., 1978). It has been suggested that the oligosaccharide chains are required to maintain G protein in a proper unaggregated form during its intracellular transit (Gibson et al., 1978, 1979, 1981).

(e) M Protein

The other envelope associated protein in VSV is the matrix protein M. M is the most abundant protein present in the virus and has an apparent molecular weight of 29,000 D (Wagner et al., 1972). The nucleotide sequence of the M mRNA predicts a polypeptide of 229 amino acids (Rose and Gallione, 1981).

Evidence that M protein occupies an internal position in the virus comes from studies demonstrating its resistance to externally added proteases (Mudd, 1974; Schloemer and Wagner, 1975a) and its inaccessibility to surface labelling reagents (McSharry, 1977; Egar et al., 1975).

The abundance of M protein (2-4000 molecules/virion) suggests that it has a major structural role in the virus. Crosslinking studies have indicated that M protein is located on the inner surface of the membrane and in close proximity to the nucleocapsid as well as the G protein (Mudd and Swanson, 1977; Dubovi and Wagner, 1977; Zakowski and Wagner, 1980; Mancarella and Lenard, 1981) thus forming a bridge between the RNP and the envelope. The use of hydrophobic probes to examine the membrane interaction of M protein has demonstrated that M protein is in close association with the

viral envelope, however, it does not seem to penetrate the membrane to a significant extent (Stoffel et al., 1978; Pepinsky and Vogt, 1979; Zakowski and Wagner, 1980).

Direct evidence that M protein is able to interact with G protein present in the membrane of infected cells has been provided by fluorescent photobleaching studies that have demonstrated that M protein can modulate the lateral mobility of G protein (Reidler et al., 1981; Johnson et al., 1981).

Another recently observed structural property of the M protein appears to be its ability to maintain nucleocapsids in a condensed form. In the absence of M protein the nucleocapsid assumes an extended form, however, in the presence of M protein the nucleocapsid more closely resembles the compact structure seen in the native virus. This effect is dependent on the ionic environment, however, the presence of a viral envelope is not essential (Newcomb and Brown, 1981; Newcomb et al., 1982).

Functionally, M protein plays an important role in virus maturation. It is believed that M protein provides a nucleation site on the plasma membrane to initiate the budding process. Studies with temperature sensitive mutants have indicated that while there is no absolute requirement of G protein for budding to take place, M protein is essential for this function to occur (Schnitzer and Lodish, 1979; Lodish and Porter, 1980a). This conclusion has been supported by experiments involving pseudotypic virus formation which have shown that the M protein is the key protein for budding to

occur (McSharry et al., 1971; Schnitzer et al., 1977; Weiss and Bennet, 1980).

Biochemical and genetic studies have indicated that M protein plays a role in the replication of the virus by modulating transcriptase activity. M protein has been shown to drastically reduce the level of transcription in vitro, an effect that is dependent on the ionic environment (Clinton et al., 1978, Martinet et al., 1979; Carroll and Wagner, 1979; Combard and Printz Ane, 1979; De et al., 1982). These functions may be modulated by the recent observation that M protein is a phosphoprotein which contains phosphothreonine, phosphoserine, as well as phosphotyrosine residues (Clinton and Huang, 1981; Clinton et al., 1982). Protein phosphorylation is a post translational modification which regulates the activity of many enzymes (Krebs and Beavo, 1979). The importance of phosphorylation in the case of M protein is not clear since experiments designed to test the effects of endogenous phosphorylation on M protein function did not show any significant changes (Clinton et al., 1982).

1.2.3 Lipids

VSV is surrounded by a typical unit membrane structure which accounts for approximately 20% of its weight (Patzer et al., 1979). Electron microscopic studies have revealed that VSV obtains its membrane by budding through the plasma membrane from a preformed pool of cellular phospholipids. In the case of VSV, virus directed synthesis of specific lipids does not occur (Lenard, 1978).

In general, the composition of the viral envelope reflects that of the plasma membrane of the host cell. The differences usually involve the fatty acid constituents of both the phospholipids and the neutral lipids which generally reflect a higher degree of saturation in comparison to the host cell (Lenard, 1978; Patzer et al., 1979). The viral membrane also shows a higher degree of cholesterol and sphingomyelin. However, this does not necessarily imply that the virions have an intrinsic ability to select for certain lipids. It is thought that the higher cholesterol content is responsible for the lower membrane fluidity as compared to the plasma membrane (Rothman and Lenard, 1977; Moore et al., 1978).

Studies with intact virions using surface labeling reagents such as trinitrobenzene sulfonic acid or reaction with phospholipases have demonstrated the asymmetric nature of the viral envelope (Fong et al., 1976; Shaw et al., 1979; Moore et al., 1977; Patzer et al., 1979). When VSV was grown in BHK 21 cells, choline containing phospholipids were found predominantly on the outer layer of the membrane while the inner layer was composed mostly of aminophospholipids. Phosphatidylethanolamine, the most abundant phospholipid species was more equally distributed on both sides. Gangliosides appear to reside exclusively on the outer surface of the bilayer as evidenced by their sensitivity to neuraminidase (Stoffel et al., 1975).

The fatty acids also have an asymmetric distribution with the external layer enriched in saturated fatty acids while almost all of

the polyunsaturated fatty acids are present in the internal monolayer (Patzer et al., 1978).

Studies utilizing ESR spectroscopy and fluorescent depolarization have revealed that the viral membrane is considerably less fluid than that of the host cell (Sefton and Keegstra, 1974; Barenholz et al., 1976). This increase in rigidity is a reflection of both the lipid composition and the presence of the membrane proteins G and M (Patzer et al., 1979; Altstiel and Landsberger, 1981). It has been speculated that this increase in membrane rigidity is a prerequisite for viral budding to occur (Lenard 1978).

1.3 Life Cycle of VSV

The infectious cycle of VSV involves a number of interrelated steps; i) penetration of the virus into the cell and introduction of the genome into cytoplasm, ii) transcription and translation of viral specific mRNA, iii) replication of progeny viral genomes, iv) assembly of viral components, and v) egression of completed virus particles into the extracellular space.

1.3.1 Absorption and Penetration

Viral infection requires an initial interaction at the cell surface which results in the penetration of the virus through the plasma membrane and into the cytoplasm where it can be replicated.

The G protein is essential for the infectivity of VSV since removal of G by proteolysis or incubation of cells with purified G inhibit viral infection (Bishop et al., 1975a; Bishop and Smith,

1977). It has been assumed that this is a cell surface phenomena whereby G protein modulates the binding of the virus to the plasma membrane. Recent evidence, however, using purified G protein incorporated into phospholipid vesicles or in rosette formation indicate that the requirement of G protein for infectivity is an intracellular rather than a cell surface event (Miller and Lenard, 1980; Miller et al., 1980; Thimmig et al., 1980).

Several lines of evidence have suggested that specific membrane receptors for VSV do not exist. These include the fact that VSV has an extremely broad host range extending from insects to mammals. In addition, trypsin or neuraminidase treatment of cells does not inhibit viral infection (Scholemer and Wagner, 1974) and saturable binding of VSV to BHK cells at 37°C could not be demonstrated (Miller and Lenard, 1980). However, in a recent study in which binding of VSV to Vero cells was examined at 4°C rather than at 37°C in order to prevent endocytosis, a saturable surface component was detected for both VSV binding and uptake (Schlegel et al., 1982a). This component was insensitive to proteases, neuraminidase, and heating but it was sensitive to phospholipase C and soluble in organic solvents. The saturable component was identified as phosphatidylserine (Schlegel et al., 1983). It was further demonstrated that phosphatidylserine could specifically bind to VSV, probably through G protein, and inhibit both VSV attachment and infectivity. Thus, phosphatidylserine appears to be the binding

site for VSV or an important component of it. The use of specific phospholipid as a binding site for VSV would explain both its broad host range and the observation that proteolytic treatment of the host cell surface does not inhibit virus infection.

The mechanism by which the virus actually penetrates the cell and is uncoated is controversial. Evidence exists both for fusion at the cell surface whereby penetration and uncoating occur simultaneously (Heine and Schnaitman, 1969, 1971) and endocytosis whereby the virus penetrates target cells by phagocytosis (Simpson et al., 1969; Dahlberg, 1974; Fan and Sefton, 1978). It is becoming clear that the principle pathway by which VSV enters cells is endocytotic. Electron microscopic visualization has demonstrated that VSV attaches to membrane surface sites in regions containing coated pits which are then engulfed by lysosomes (Dickson et al., 1981). Coated pits are known to mediate cell entry of various extracellular components by receptor mediated endocytosis (Pearse and Bretscher, 1981). Compounds which are known to inhibit receptor mediated endocytosis, such as dansylcadaverine and amantadine, are able to inhibit viral infection by preventing sequestration of VSV into coated pits (Schlegal et al., 1982b).

There is now compelling evidence that viral uncoating is induced by the low pH present in lysosomes (Lenard and Miller, 1982). VSV has a low pH induced fusion activity, thus the acidic environment of the lysosomes causes the viral and lysosomal membranes to fuse, possibly by a mechanism requiring G protein (Matlin et al., 1982).

Since the virus membrane becomes contiguous with the cytoplasmic face of the vacuolar membrane, the RNP becomes exposed to the cytoplasm where replication and viral protein synthesis can begin.

1.3.2 Transcription

The single stranded genome of VSV is transcribed by the virion associated RNA dependent RNA polymerase (Baltimore et al., 1970) into two distinct classes of RNA species. One class consists of the five monocistronic, capped, methylated, and polyadenylated mRNAs coding for the viral proteins L, G, M, NS, and N (Morrison et al., 1974; Moyer and Banerjee, 1975; Rose and Knipe, 1975; Knipe et al., 1975; Both et al., 1975a; Banerjee et al., 1977). A second class consists of a small untranslated 47 nucleotide nonpolyadenylated RNA species, known as the leader RNA, that is complementary to the 3' terminal end of the VSV genome (Colonno and Banerjee, 1976, 1977, 1978).

Both L protein and NS protein are required for VSV transcription (Wagner, 1975). L protein is probably the catalytic activity of RNA synthesis while NS may function as an initiator for transcription by mediating the interaction between L protein and the viral template (Keene et al., 1981).

Hybridization studies (Roy and Bishop, 1972) as well as studies on the inactivation of gene expression by UV irradiation have shown that the RNAs are synthesized sequentially in the order 3' leader RNA-N-NS-M-G-L 5' (Abraham and Banerjee, 1976; Ball and White, 1976; Ball, 1976) which is a reflection of the physical order of the

genes. A gradient of transcription exists in the same direction with a decreasing molar amount of transcripts encoded by the cistrons as a function of their distance from the 3' terminus of the genome (Villarreal et al., 1976). This implies that the 3' terminus of the genome plays an essential role either in the initiation of transcription or the binding of the transcriptase to the template. Methylation protection experiments have demonstrated that the VSV polymerase binds to the genomic RNA template in the middle of the leader gene approximately 16-30 nucleotides from the 3' terminus (Keene et al., 1981). This is an A+U rich sequence that resembles the Pribnow box or Goldberg-Hogness transcription promoter site (Rose, 1980) and supports the single entry model of VSV mRNA transcription.

The viral genome displays a very tight organization. Of the 11,000 nucleotides, only 117 are not represented in the mature mRNAs. Of these, 47 are in the leader RNA, 3 nucleotides at the leader-N gene junction, 2 at each of the four intergenic regions, and 59 between the end of the L gene and the 5' terminus of the genome (McGeoch, 1979, 1981; Rose, 1980).

The genome sequence $AUACU_7$ is common to all intercistronic genomic regions encoding the 3' termini of the mRNAs (Rose, 1980). This consensus sequence is thought to initiate poly A synthesis which is probably carried out by some part of the viral transcription apparatus. Since the entire poly A sequence is not encoded in the genome, it has been postulated that the addition of the 2-500

adenylate residues opposite the U_7 sequence occurs by repeated slippage of the polymerase at the U_7 sequence instead of progression of the enzyme with respect to the template (Ball and Wertz, 1981).

The five viral mRNAs have a common 5' terminal sequence complimentary to 3' UUGUCNNUAG 5', (Where N is variable) which may specify the capping/methylation site. The polyadenylation and capping sites are separated by dinucleotide spacers (CA or GA) which represent the intergenic regions since they are not present in the mRNA (Rose, 1980).

Several models have been proposed to account for the sequential order of transcription of the VSV genome. The progressive cleavage model postulates that each complete mRNA species is generated by putative cleavage process of a precursor RNA followed by post-transcriptional capping at the 5' end and polyadenylation at the 3' end (Abraham et al., 1975). Using ATP analogues to inhibit normal mRNA synthesis, Chinchar et al., (1982) were able to demonstrate the existence of linked transcripts of the type leader-N message and leader-N-M message, however, they were present in very low amounts. VSV mRNAs linked by poly A tracts have also been detected indicating that the polymerase can read through the intercistronic boundary, however, there has been no demonstration that these linked transcripts can be processed into monocistronic mRNAs (Herman et al., 1978, 1980).

A second model, known as the multiple entry stop/start model,

suggests that the mRNAs are generated by termination and reinitiation at each intercistronic region (Banerjee et al., 1977). This model is supported by *in vitro* studies that have demonstrated the existence of several short triphosphate initiated oligonucleotides (11 - 14 bases) representing the 5' uncapped sequences of the N mRNA that do not remain template bound. These were found in much larger amounts than the leader RNA providing evidence for efficient internal initiation at the N gene (Pinney and Emerson, 1982).

Recently, Testa et al., (1980) have postulated a third model of transcription (the simultaneous initiation model) suggesting that initiation of transcription occurs simultaneously at each gene resulting in the transcription of 5' RNAs of 30-70 nucleotides in length from the 3' end of each gene that remain template bound. These short RNAs, which bear di and triphosphate termini, are elongated into their corresponding full length mRNAs only after transcription of the preceding gene is complete. This model accounts for the sequential order of transcription as well as for the polarity observed. However, by examining the kinetics and stability of these triphosphate initiated mRNAs, it has recently been shown that most of them are products of abortive transcription (Lazzarini et al., 1982). It could not be completely ruled out, however, that a small subset of these RNAs can mature into functional mRNA.

In a related study examining the kinetics of appearance of 5' proximal RNase T1 oligonucleotides present at the 5' ends of the viral messages, it was found that the 5' ends of the mRNAs are

synthesized sequentially as is the case for the synthesis of full length message and that the synthesis of the oligonucleotides is dependent on the transcription of the preceeding gene (Iverson and Rose, 1982). This is inconsistent with transcription involving simultaneous initiation and presynthesis of 30-70 nucleotide long leader RNAs for each mRNA.

The single entry model of VSV mRNA transcription has been further supported by examining the appearance of short capped RNAs from the 5' terminus of both N and NS mRNA. In this case, the termini appear sequentially rather than simultaneously as predicted by the stop/start model (Piwnica-Worms and Keene, 1983).

Thus, the available evidence suggests that the VSV polymerase can gain access to the template by sequential readthrough starting from the 3' end or by internal initiation, however, it appears that internal initiation does not play a role in the normal transcriptive process. The precise method of VSV transcription remains to be elucidated.

1.3.3 Translation

Viral proteins are translated from the five monocistronic mRNA species produced from the primary and secondary transcription of the viral genome. That the mRNA species found in VSV infected cells have messenger function has been demonstrated in early studies where polyribosomes from VSV infected cells were isolated and translated *in vitro* into authentic viral proteins (Ghosh et al., 1973; Toneguzzo and Ghosh, 1975; Morrison et al., 1974; Knipe et al., 1975).

Viral proteins can be detected 1 hour after infection and are synthesized throughout the infectious cycle with protein synthesis peaking at approximately 4-5 hours post infection (Kang and Prevec, 1971). It appears that the relative amount of each viral gene product is a function of the relative level of mRNA from which it is translated (David, 1978). However, there also seems to be regulation at the translational level since only 30% of the total VSV mRNA in the infected cell is bound to polyribosomes at any one time despite the presence of a large pool of monomeric ribosomes (Huang et al., 1970; Morrison and Lodish, 1975).

N protein is made in the most abundant amount, representing approximately 40% of total viral protein synthesized (Kang and Prevec, 1971). L protein accounts for only 2-3% of total viral protein despite the fact that its mRNA represents approximately 30% of total VSV mRNA (Stampfer et al., 1969, Stamminger and Lazzarini, 1974; Stampfer and Baltimore, 1973; Kang and Prevec, 1971) suggesting that regulation of L protein synthesis occurs at the translational level. NS protein appears to be synthesized in large amounts early on in the infectious cycle but in decreasing amounts as infection progresses (Kang and Prevec, 1971). Phosphorylation of NS is an early event occuring before or shortly after its synthesis (Moyer and Summers, 1974b). G and M proteins increase in amount throughout the infectious cycle reaching peaks of 25 and 30% of total viral proteins respectively (Wagner, 1975).

The majority of viral mRNA becomes polysome associated by

3 hours after infection (Lynch et al., 1981). These polysomes fall into two classes - free in the cytoplasm and membrane bound. The mRNAs coding for L, NS, N, and M are found associated with free polyribosomes, however, the G mRNA is translated exclusively from membrane bound polyribosomes indicating that the various viral proteins have different pathways of maturation (Toneguzzo and Ghosh, 1975; Both et al., 1975b; Morrison and Lodish, 1975; Grubman et al., 1975). The biosynthesis of the membrane proteins G and M will be considered in more detail.

(a) G Protein

G protein undergoes a number of discrete modifications during synthesis and transportation to its final destination in the plasma membrane. G protein is always found in association with cellular membrane and never free in the cytoplasm (Wagner et al., 1970). Early studies on the synthesis of G demonstrated that G is associated with the rough endoplasmic reticulum (RER) during or shortly after its synthesis. The G protein can be chased into smooth cytoplasmic membranes and after about a 20 minute lag, it appears on the plasma membrane (Atkinson et al., 1976; Hunt and Summers, 1976; Knipe et al., 1977a, 1977b). During its maturation, G protein undergoes proteolytic processing, glycosylation, and fatty acid addition.

i) Membrane Insertion

The mechanism by which G protein is inserted into membranes is based on the "signal hypothesis" formulated by Blobel and

coworkers (Blobel and Sabatini, 1970; Blobel and Dobberstein, 1975a, 1975b). According to this model, secretory proteins are synthesized on ribosomes that are associated with the membrane of the RER which results in the cotranslational vectorial discharge of the polypeptide across the membrane and into the luminal space. The polypeptide then migrates through the intracellular membrane system, including the Golgi complex, by closed vesicles that bud from one organelle and fuse to the next whereupon it is secreted into the extracellular space (Palade, 1975). The specific information that determines polysome/ER membrane interaction is usually contained in an amino terminal transient domain of 15-30, principally hydrophobic amino acids known as the signal sequence. This amino terminal extension is said to initiate membrane association and cotranslational insertion of the nascent polypeptide by recruiting ER membrane receptor proteins in the formation of a hydrophilic multimeric, transmembrane tunnel that is stabilized by the direct association of the large ribosomal subunit with the receptor proteins. The existence of a protein tunnel would ensure that the complete polypeptide is sequestered into the cisternal space of the ER. The signal sequence is then cleaved by a "signal peptidase" during or shortly after it is inserted into the membrane (Blobel and Dobberstein, 1975a, 1975b; Bolbel, 1978).

Evidence in support of the central concepts of this model, that is, the existence of a transient hydrophobic amino terminal domain and cotranslational membrane penetration, are numerous

(Sabatini et al., 1982), however, different models, which do not envisage the participation of specific receptor proteins have been postulated (Inouye and Haegoua, 1980; von Hiijne, 1981; Wickner, 1980; Engleman and Steitz, 1981).

Recently, much information has accumulated in support of the existence of specific receptors. These include the observations that synthetic signal peptides can compete with proteins undergoing translocation (Prehn et al., 1981) and that translocation in vitro can be prevented by treating microsommal membranes with high salt concentrations or with proteolytic enzymes (Warren and Dobberstein, 1978; Walter et al., 1979; Jackson et al., 1980; Prehn et al., 1980; Meyer and Dobberstein, 1980). Protein translocation can be restored by the addition of a high molecular weight protein complex known as the signal recognition protein (SRP) that is removed from the salt extracted rough microsomes and has a high affinity only for ribosomes that are synthesizing secretory and membrane proteins (Meyer, 1982). In the absence of membranes, the SRP blocks translation as soon as the nascent chain is long enough to be recognized as a secretory or membrane protein (Walter and Blobel, 1981).

The current model is that protein synthesis begins on free ribosomes. The appearance of the signal sequence results in an interaction with the SRP thus blocking further translation. The translational block persists until the complex associates with a specific signal recognition protein receptor, known as the docking protein, which is present in the membrane of the ER (Meyer et al.,

1982a, 1982b; Gilmore et al., 1982a, 1982b). This interaction causes translation to resume and translocation to begin.

The steps involved in the synthesis, membrane insertion, and glycosylation of the VSV G protein have been well characterized by *in vitro* studies from a number of laboratories (Toneguzzo and Ghosh, 1977, 1978; Rothman and Lodish, 1977; Katz et al., 1977; Katz and Lodish, 1979).

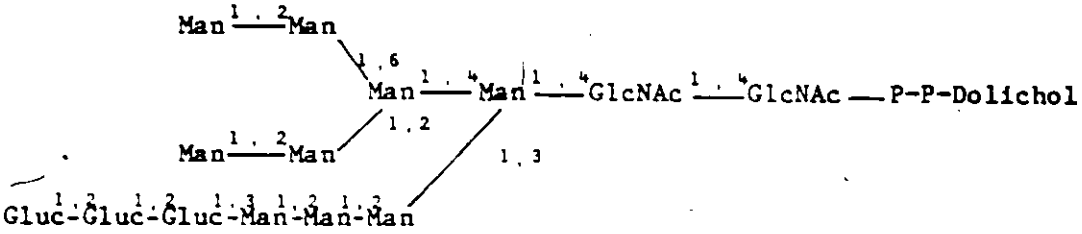
In an *in vitro* translation system in the absence of membranes, G protein is synthesized as a non glycosylated precursor G_1 having a molecular weight of 63,000 D. In the presence of membranes, G_1 is converted to G_2 (MW 67,000 D) by a cotranslational event involving membrane insertion, core glycosylation, and the concommitant removal of a 16 amino acid hydrophobic signal sequence from the amino terminus (Lingappa et al., 1978a; Irving et al., 1979). The requirement for membranes is an early event in that protein sequestration is prevented if the nascent chain is longer than about 80 amino acids (Rothman and Lodish, 1977).

Unlike secretory proteins, G protein is not completely extruded into the luminal space of the microsommal membrane but retains an absolute transmembranal orientation in which about 30 amino acids at the carboxy terminus is exposed to the cytoplasm while the remainder of the polypeptide, including the carbohydrate units and the amino terminus, faces the luminal side (Toneguzzo and Ghosh, 1978; Katz et al., 1977; Katz and Lodish, 1979). This same orientation is observed in microsommal vesicles isolated from

infected cells (Chatis and Morrison, 1979).

ii) Glycosylation

The initial glycosylation event of G protein occurs through the en bloc transfer of a preformed common oligosaccharide of the type



from a dolichol lipid carrier (Robbins et al., 1977).

The two core oligosaccharide chains are transferred to the G protein in sequence while the polypeptide is being translocated into the lumen of the ER (Rothman and Lodish, 1977). The core oligosaccharides are then processed in a series of steps to yield the complex oligosaccharide shown in section 1.2.2 (Tabas et al., 1978; Li et al., 1978; Kornfeld et al., 1978; Tabas and Kornfeld, 1978). The three terminal glucose residues are removed within a few minutes of transfer of the core oligosaccharide. This is followed by the removal of the first four mannose residues and the addition of the 3 N-acetylglucosamine, 3 galactose, 1 fucose, and 3 sialic acid residues in a stepwise fashion. Location of the processing enzymes suggest that core addition occurs in the RER, trimming occurs in the smooth membranes, and terminal sugar addition occurs in the Golgi

complex (Grinna and Robbins, 1979; Tabas and Kornfeld, 1979).

The biological role of the carbohydrates in the maturation of G protein is not clear. Glycosylation is not required for insertion of the G protein into the RER (Rothman et al., 1978), signal sequence cleavage (Irving et al., 1982) nor is it an absolute requirement for the transfer of G protein to the plasma membrane (Gibson et al., 1978). In addition, glycosylation has no effect on the polarized distribution of G protein in the plasma membrane of MDCK cells (Green et al., 1981) where G protein is normally found on the basolateral surface (Rodriquez-Boulan and Sabatini, 1978).

However, the conformational properties of G protein are affected by glycosylation. In the absence of glycosylation, G protein becomes aggregated at elevated temperatures which results in a decreased production of virus (Gibson et al., 1979, 1981).

iii) Fatty Acid Acylation

Schlesinger and coworkers have demonstrated that G protein contains 1-2 moles of tightly bound fatty acid which appear to be covalently linked to the polypeptide backbone (Schmidt and Schlesinger, 1979). Fatty acid acylation is a post-translational modification that appears to be an early event in the biosynthesis of G protein occuring shortly before the protein bound oligosaccharides become resistant to endoglycosidase H (Schmidt and Schlesinger, 1980). This suggests that fatty acid acylation occurs while G protein moves through the Golgi complex. This is supported by the observation that monensin, an ionophore that causes the accumulation

of G protein in the Golgi complex, does not prevent fatty acid attachment (Johnson and Schlesinger, 1980).

iv) Intracellular Transport of G Protein

G protein must migrate from its site of synthesis in the RER to the plasma membrane where it is incorporated into budding virions. Biochemical as well as cell fractionation studies indicate that G protein follows the secretory pathway described by Palade (Palade, 1975), that is, RER ⟶ Golgi complex ⟶ plasma membrane (Hunt and Summers, 1976; Tabas and Kornfeld, 1979; Knipe et al., 1977a, 1977b, 1977c; Johnson and Schlesinger, 1980; Fries and Rothman, 1980; Rothman and Fries, 1981). This morphological pathway of exocytosis has recently been confirmed by direct visualization of G protein migration using immunocytochemical techniques (Bergman et al., 1981) as well as electron microscope radioautography (Bergeron et al., 1982).

The actual mechanism by which G protein is transported is controversial. Röthman has suggested that G protein is transported through the intracellular membrane system in two successive stages by clathrin coated vesicles. The first wave transports G from its site of synthesis in the RER to the Golgi complex while a second wave transports G from the Golgi to the plasma membrane (Rothman and Fine, 1980; Rothman et al., 1980; Rothman, 1981). In contrast to this, a recent study using fluorescent electron microscopic localization techniques failed to detect any G protein in coated vesicle structures (Wehland et al., 1982). It was suggested that the G

protein containing vesicles observed by Rothman arose artifactually by the closure of coated pits of the Golgi-GERL system during cell homogenization.

(b) M Protein

In comparison to the G protein, much less is known about the biosynthesis and maturation of the M protein. M protein is predominantly synthesized on free ribosomes and is found as a soluble cytoplasmic protein soon after its synthesis (Morrison and Lodish, 1975).

In contrast to the lag observed in the synthesis of G protein, M protein progressively associates with membranes and can appear in the plasma membrane within 2 minutes in pulse chase experiments (Atkinson et al., 1976; Knipe et al., 1977c). Clearly then, the mechanism by which M protein attaches to membranes is entirely different to that of G protein.

Only a small amount of M protein is found associated with the plasma membrane in active budding sites at any one time (Knipe et al., 1977a, 1977c, 1977d). This is consistent with M protein acting as a nucleation site (for the RNP core) on the inner side of the membrane followed by rapid binding of the nucleocapsid.

The phosphorylation of M protein is heterogenous in that only 5-10% of the protein molecules contain up to 90% of the phosphate (Clinton et al., 1978a). The phosphorylation is a post-translational event that may be carried out by a VSV associated protein kinase or an endogenous cellular enzyme (Clinton et al., 1982).

1.3.4 Replication and RNP Assembly

The incoming 42S genome of VSV must serve as a template for two events; one is the production of viral mRNA, while the other is the synthesis of a full length 42S +ve copy. This +ve strand serves as the template for the synthesis of full length -ve sense progeny genomic RNA copies which become packaged into RNP structures and form new virus particles. The relationship between these two processes and the mechanism by which the mRNA transcription/ replication switch is accomplished during the infectious cycle is largely unknown.

Studies with temperature sensitive mutants defective in L protein synthesis have demonstrated that L protein is part of both the transcriptive and replicative enzymes (Perlman and Huang, 1973; Repik et al., 1976) suggesting that transcription and replication are interrelated events. Other studies have indicated that both L and NS are required for replication (Emerson and Yu. 1975; Hunt et al., 1976; Naito and Ishihama, 1976).

Both virion -ve strand and full length +ve strand 42S RNA can be isolated from intracellular RNP complexes (Simonsen et al., 1979). Nascent replicative RNA exists in RNP structures and is never found as naked RNA in vivo (Naeve and Summers, 1980).

In contrast to the transcriptive process, replication of the genome to produce full length -ve strand RNA is dependent on concurrent protein synthesis (Perlman and Huang, 1973; Wertz and Levine, 1973). It has been speculated that RNP protein synthesis and

its association with nascent RNA are required to suppress termination at a postulated strong stop signal present at the 3' end of the minus strand leader so that full length plus strand can be made (Leopert et al., 1979). This concept is derived from the belief that once RNA synthesis is initiated, it can proceed to yield mRNA or full length negative strand RNA.

Electron microscopic visualization of replicative RNPs has led to the suggestion that replication may occur in part on circular coiled RNP templates (Naeve and Summers, 1979). This is consistent with the finding that the termini of the genomic RNA are complementary (Keene et al., 1979).

Recently, 42S RNA of both +ve and -ve polarity has been synthesized in vitro and found to be assembled into nucleocapsid like structures (Ghosh and Ghosh, 1982; Hill et al., 1981, Davis and Wertz, 1982). N protein was the most abundant protein found in the nucleocapsid structures but there were also small amounts of L and NS.

In one case, the nucleocapsids were assembled from viral proteins and genomic RNA synthesized concurrently in vitro. Newly synthesized proteins could not associate with input nucleocapsid templates in the absence of RNA synthesis thus demonstrating the coupled nature of RNA synthesis and nucleocapsid assembly (Patton et al., 1983).

Kinetic analysis of nucleocapsid assembly in vivo has shown that it is a stepwise process consisting of an initial interaction

of the N protein with the viral RNA followed by attachment of L and NS (Hsu et al., 1979).

Recent evidence indicates that the switch from transcriptive to replicative RNA synthesis may depend on the intracellular levels of N protein (Blumberg et al., 1981; Lazzarini et al., 1981). It has been speculated that when N protein is abundant, nascent RNA transcripts are encapsidated. This results in the release of the polymerase from a site specific attenuation and allows it to synthesize full length genomic RNA. When the levels of N protein are low, as in the early stages of infection or in in vitro transcription, the nascent transcripts are released and new transcripts are initiated and elongated into monocistronic mRNAs. The mRNAs are not encapsidated because they do not contain the nucleation site which is believed to be found only at the 5' terminus of the +ve or -ve strand genomic RNA. This model is given support by several observations. In addition to the 47 nucleotide long leader RNA complementary to the 3' end of the -ve stranded genome, a 46 nucleotide long RNA complementary to the 3' end of the +ve strand genome RNA can be detected in infected cells (Schubert et al., 1978). The +ve and -ve leaders have homologous 5' termini, thus the termini of the genomic RNA are complementary (Keene et al., 1979). It has recently been demonstrated that N protein can encapsidate the leader RNAs in vitro into nucleocapsid like structures (Blumberg et al., 1983). Thus, the leader RNAs contain the nucleation site for RNP assembly. The selective sequence appears to be the group of five A

residues found at every third position from the 5' end of the leader chain. It is suggested that the RNP assembly proceeds co-operatively after the binding of the first two N protein molecules to the -ve leader strand. The encapsidation of the leader strand by N protein excess would then initiate replicative transcription. The mRNAs would not be encapsidated since they lack the nucleation site.

1.4 Virus Assembly and Morphogenesis

The assembly and maturation of the VSV virus particle consists of the association of two distinct components that are assembled and synthesized in different parts of the cell; the plasma membrane modified by the G and M proteins, and the cytoplasmically assembled RNP. Once this association is made, the completed virus particle egresses by a budding out mechanism in which the nucleocapsid becomes surrounded by the modified plasma membrane and is pinched off from the cell surface.

On the basis of crosslinking studies and the formation of mixed pseudotypes, it has been suggested that M protein serves as the link between the envelope and the nucleocapsid (Dubovi and Wagner, 1971; McSharry et al., 1971).

Pulse chase experiments have demonstrated that virions bud very soon after M protein associates with the region of the plasma membrane that contains G protein (Knipe et al., 1977a, 1977d). It is clear, however, that the presence of G protein is not entirely essential for budding to occur. Studies with temperature sensitive

mutants that are defective in the transportation of G protein to the plasma membrane have shown that virus particles are still produced, albeit in much lower amounts. These non-infectious particles were shown to contain no G protein but contain RNA and the other viral proteins in the normal amounts (Lodish and Weiss, 1979; Schnitzer et al., 1979). In addition, it has recently been demonstrated that G protein is not required for the ability of virions to package only -ve stranded RNA containing RNPs in spite of the fact that RNPs of both polarities are present in the cytoplasmic pool (Marnell and Wertz, 1982).

The G protein, however, may play a role in determining the actual site of plasma membrane envelopment. In polarized epithelial MDCK cells infected with VSV, G protein migrates exclusively to the basolateral and not the apical membrane surface. Consequently, virus budding was found to occur only from the basolateral membrane. (Rodriguez-Boulan and Pendegast, 1980).

In contrast to G protein, it is evident that the M protein plays a critical and essential role in the budding process. The synthesis of M protein seems to be rate limiting for virus particle formation (Kang and Prevec, 1971; Portner and Kingsbury, 1976; Atkinson, 1978). Studies on released virus have shown that while the ratio of M and G in the completed particle can vary greatly, the ratio between N and M is constant (Lodish and Porter, 1980a). In virus mutants defective in nucleocapsid assembly, G protein is transported to the cell surface, however, virus particles are not

formed (Knipe et al., 1977a, 1977d). In this case, it was found the M protein accumulates in the cytoplasm suggesting that M protein does not form a stable complex at the cell surface unless the nucleocapsid is present.

These data support the conclusion that there is a specific interaction between the M protein and the nucleocapsid and that M protein is the key polypeptide in virus formation.

1.5 Cellular Reactions to Infection

Infection of cells with VSV results in gross morphological and cytopathological changes which eventually result in cell death. Infection is accompanied by rapid inhibition of cellular DNA, RNA, and protein synthesis (Wagner, 1975).

The mechansim by which cellular synthetic functions are suppressed is controversial. Neither the replication of the genomic RNA nor the production of infectious virus is required to shut off host protein synthesis, however, some viral transcription is essential (McAllister and Wagner, 1976; Marcus et al., 1977; Marvaldi et al., 1978). The inhibitory function appears to be directed at the level of initiation of translation (Nuss and Koch, 1976) but neither the nature of the functions(s) nor their cellular targets are known.

Lodish and Porter suggested that the inhibition of host protein synthesis results from a competition for existing ribosomes between viral and cellular mRNAs thus obliviating the need for any specific viral inhibitory function (Lodish and Porter, 1980b, 1981). This mechanism is supported by the fact that viral infection results

in a two to three fold increase in total translatable mRNA within 4 hours of infection (Jaye et al., 1982; Lodish and Porter, 1980b, 1981).

This concept has been opposed on several grounds. In studies using a temperature sensitive mutant defective in the inhibition of host protein synthesis (Stanners et al., 1977), it was found that the level of viral mRNA bound to ribosomes was identical to that obtained from wild type infected cells (Schnitzlein et al., 1983). Furthermore, using UV inactivated virus where replication was inhibited and only N protein specific transcripts were made (therefore, no mRNA competition occurred), it was found that host protein synthesis was still inhibited (Marvaldi et al., 1978). The UV inactivation curve suggested that two transcription products, having UV target sizes of 375 and 42 nucleotides, were required for inhibition of host protein synthesis of (Dunnigan and Lucas-Lenard, 1983). However, the precise origin of these transcripts was not determined.

In addition to cellular protein synthesis, VSV infection results in the inhibition of cellular RNA synthesis (Weck et al., 1979; Wu and Lucas-Lenard, 1980). Thus, the inhibition of protein synthesis could be a direct consequence of the inhibition of RNA synthesis. Thomas and Wagner were able to detect a double stranded RNA molecule specific to VSV infected Hela cells that was able to inhibit <u>in vitro</u> protein synthesis in a reticulocyte lysate (Thomas and Wagner, 1982). In addition, the plus strand leader RNA of VSV

was found to inhibit DNA dependent transcription of adenovirus and SV40 genes in a soluble whole cell extract (McGowen et al., 1982). This latter observation is intriguing in light of the recent observation that the VSV leader RNA rapidly localizes in the nuclei of infected cells shortly after its synthesis (Kurilla et al., 1982).

Thus it appears that VSV mediated inhibition of cellular protein and RNA synthesis requires a viral specific transcription product(s); however, the cellular target of this product remains to be defined.

1.6 Scope of This Work

It is apparent from many studies over the past few years that the important biological functions carried out by membranes is reflected in the organization and structural asymmetry of its components (Rothman and Lenard, 1977). In the case of membrane proteins, this organization is imparted by structural and functional domains encoded in the polypeptide backbone that specify various types of membrane and protein interactions. Thus in the case of VSV, in order to understand the dynamic processes involved in membrane biogenesis and viral assembly, it is necessary to know not only how the viral components are arranged with respect to each other and with the viral envelope but also to comprehend the biological basis of these interactions.

The major portion of this thesis deals with the VSV membrane proteins G and M. Since the asymmetric orientation of G protein with respect to the viral membrane is thought to be essential for it to

carry out its biological functions, the topology of G protein with respect to the viral envelope was examined. The basis of this interaction was studied by partial amino acid sequence analysis and localization of fatty acid residues. This information was correlated with what is known about the biogenesis and maturation of this molecule.

The M protein is essential for viral morphogenesis; however, little is know about its interaction, if any, with the viral envelope and the RNP core. The membrane organization of both G and M protein was examined with the use of a chemically synthesized hydrophobic photoreactive probe. The nature of the interactions of M protein was studied by developing a reconstitution system with artificial lipid vesicles. This has allowed the direct examination of the interaction of M protein with the nucleocapsid protein N.

Finally, initial studies were undertaken to identify and examine the role of various domains or "topogenic sequences" present in membrane proteins that are thought to have important biological functions involved in biogenesis and maturation. This was accomplished through the construction and expression of a chimeric gene from cDNA clones containing the coding sequences for G and M protein.

MATERIALS AND METHODS

2.1 Materials

(a) Chemicals and Reagents

HPLC grade ethyl acetate, benzene, butyl chloride, methanol, acetonitrile, and tetrahydrofuran were purchased from Burdick and Jackson; PTH amino acid standards, sequence grade phenylisothiocyanate (PITC), heptafluorobutyric acid (HFBA), and 0.1 M Quadrol were obtained from Beckman or Pierce Biochemicals; Trifluoroacetic acid was from Pierce Biochemicals while ethanethiol heptane and sperm whale myoglobin were from Beckman. Polybrene and glycylglycine were obtained from Aldrich. Sodium deoxycholate and dithiothreitol (DTT) were purchased from Calbiochem; Sodium dodecyl sulphate and Na-ethylenediaminetetraacetic acid (EDTA) were obtained from BDH Chemicals Ltd./ Triton X100 was from Rhom and Hass; acrylamide, bis acrylamide, TEMED, Amnex-A5 were from Bio-Rad Laboratories, and agarose and CsCl were from Bethesda Research Laboratories. Bovine serum albumin (BSA), phenylmethylsulfonyl fluoride (PMSF), ethylene glycol bis (β-aminoethylether) N,N,N',N'-tetraacetic acid (EGTA), phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, ampicillin, tetracycline, chloramphenicol, and ethidium bromide were purchased from Sigma Chemicals. Soybean trypsin inhibitor was from Worthington. Trasylol (Aprotinin) was from Miles Laboratories;

and Ultrogel AcA54 was from LKB Industries. Joklik modified MEM, Dulbecco's modified MEM, fetal bovine serum (FBS), fetal calf serum (FCS), new born calf serum (NBCS), and dialyzed fetal calf serum (dFCS) were purchased from Gibco. Bacto-tryptone, Bactroagar, and yeast extract were from Difco Laboratories. Spectrapore dialysis membrane was obtained from Spectrum Medical Industries. Protein A Sepharose was purchased from Pharmacia, and nitrocellulose membrane was obtained from Schleicher and Schuell. The monosodium salt of GTP and the dipotassium salt of ATP was purchased from P-L Biochemicals. Enhance was from New England Nuclear. All other chemicals were of the highest analytical reagent grade available.

(b) Enzymes

TPCK-treated trypsin, α chymotrypsin, lysozyme, and deoxyribonuclease were obtained from Worthington; Thermolysin, bovine pancrease carboxypeptidase A (DFP) and carboxypeptidase B (DFP) were purchased from Sigma Chemicals; S. aureus V8 protease was from Miles Laboratories; hog kidney aminopeptidase M was purchased from Pierce Chemical Co. or from Boehringer, Mannheim; Creatine kinase was from Boehringer, Mannheim and micrococcal nuclease was purchased from P-L Biochemicals. The restriction enzymes Bam HI, Bgl II, Eco RI, Hind III, Pst I, and Sau3A were obtained from Bethesda Research Laboratories. T_4 DNA ligase was obtained from Bethesda Research Laboratories or New England Nuclear Corp. Phospholipase A_2 was obtained from Dr. G. E. Gerber, Department of Biochemistry, McMaster University.

(c) Radiochemicals

[^{3}H]-glycine, [^{3}H]-isoleucine, [^{3}H]-leucine, [^{3}H]-lysine, [^{3}H]-methionine, [^{3}H]-phenylalanine, [^{3}H]-proline, [^{3}H]-tyrosine, [^{3}H]-valine, [^{14}C]-lysine, and [^{35}S]-methionine were purchased from the New England Nuclear Corp. or Amersham Corporation, depending on the specific activity. [^{3}H]-palmitic acid and [^{125}I]-Protein A was from New England Nuclear Corporation. [^{3}H]-diazirinophenoxy nonanoate (160 mCi/mmole or 2 Ci/mmole) was provided by Dr. G. E. Gerber and P. Leblanc, Department of Biochemistry, McMaster University.

(d) Antiserum

Antiserum specific to whole VSV or monospecific to the various viral proteins was provided by Dr. L. Prevec, Department of Biology, McMaster University. Antibody specific to the SV40 large T antigen was provided by Dr. M. Breitman, Department of Medical Genetics, University of Toronto.

(e) Cell Culture

A continuous line of L cells designated L60 was obtained from Dr. L. Prevec, Department of Biology, McMaster University and maintained in suspension culture in Joklik modified MEM supplemented with 5% NBCS, 2% sodium bicarbonate, 1% penicillin and streptomycin, and 1% L-glutamine. Monolayer cultures were maintained in α MEM containing 7% NBCS, 2% sodium bicarbonate, 2% L-glutamine, and 1% penicillin and streptomycin.

Simian COS-1 cells, obtained from Dr. Y. Gluzman (Cold Spring

Harbor) were maintained in Dulbecco's modified MEM containing 7% fetal bovine serum. E. coli LE392 (F^-, hsd R514(r_k^- m_k^-) supE44, supS58, lac YI, or Δ(lacIZY)6, gal K2, gal T22, met B1, trp, R55,λ^-) was obtained from Dr. F. L. Graham, Department of Biology, McMaster University.

(f) Virus

Plaque purified VSV was of the Indiana serotype, strain HR-LT, originally obtained from Dr. L. Prevec, Department of Biology, McMaster University. Virus stock cultures were stored at -90°C.

(g) Plasmids

pG1 and pM309, pBR322 derivatives containing cDNA copies corresponding to the complete coding sequences of the VSV G and M mRNAs, respectively, were provided by Dr. J. Rose (Salk Institute, San Diego, California). pJS223, an expression vector containing the SV40 origin and full length cDNA copy of the VSV N protein mRNA was provided by Dr. R. Lazzarini (National Institutes of Health, Bethesda, Md). The pCVSVe expression vector was obtained from Dr. R. Kaufman (Massachusetts Institute of Technology, Cambridge, Mass).

2.2 Methods

2.2.1 Virus Infectivity

Viral infectivity was determined as plaque forming units per ml (PFU/ml) by the plaque assay procedure. 60 mm culture plates were seeded with 2×10^6 cells in medium containing 0.0044% $CaCl_2$. Confluent momolayers were infected with 0.2 ml of an appropriate dilution of VSV serially diluted with PBS (138 mM NaCl, 2.7 mM KCl,

8 mM Na_2HPO_4, 1.46 mM KH_2PO_4; Dulbecco and Vogt 1954). After viral absorption, the monolayers were overlaid with 6 ml of α MEM containing 10% NBCS and 0.9% agar. After the agar solidified, the plates were incubated at 37°C for 48 hours in a 5% CO_2, 95% air atmosphere. After the incubation period, the cells were fixed with 2 ml of Carnoy's solution and viral plaques were visualized and counted after staining with 1% crystal violet.

2.2.2 Preparation of VSV

L cells growing in suspension were harvested by centrifugation at 400g for 10 min at room temperature. Cells were resuspended with Joklik modified MEM containing 2% NBCS to a concentration of 2×10^7 cells/ml and infected with VSV to an m.o.i. of 5 PFU/cell. After virus absorption at 37°C for 45 min the culture was diluted with the above medium to 2×10^6 cells/ml and incubated overnight at 37°C until at least 80% of the cells had lysed. Cellular debris was removed by centrifugation at 5,000g for 10 min at 4°C. The clarified supernatant was centrifuged at 18,000 rpm for 3 hours at 4°C in a Beckman Type 19 rotor. The crude viral pellet was suspended in a small volume of NT buffer (10 mM Tris-HCl pH 7.6, 150 mM NaCl) and overlaid on a 10 ml gradient of 10-30% sucrose in NT buffer. The gradients were centrifuged at 22,000 rpm for 50 min at 4°C in a Beckman SW41 Ti rotor. The visible viral band was collected, diluted with NT buffer, and recovered by centrifugation at 48,000 rpm for 1 hour in a Beckman Type 50.2 Ti rotor. The purified viral pellet was suspended in a small volume of

NT buffer by brief sonication and stored at -90°C. Virus concentration was determined by optical density readings at 260 and 280 nm.

2.2.3 Preparation of Radioactive VSV

(a) Labeling of Virus with Radioactive Amino Acids

L cells were grown on 100 mm plates to near confluency and infected with VSV at an m.o.i. of 10. Virus absorption was for 1 hour at 37°C after which the culture was fed with 5 mls of αMEM containing 5% NBCS and incubation was continued for an additional 3 hours. At 4 hours post-infection, the medium was aspirated and replaced with αMEM containing 2% dialyzed fetal calf serum (dFCS) deficient in the amino acid(s) to be used for labeling. After a 15 min incubation at 37°C to deplete any endogenous stores of amino acid, the desired radioactive amino acid was added to a level of 10-20 uCi/ml and incubation was continued for an additional 16-18 hours. Released virus was purifed from the clarified culture fluid by differential and rate zonal centrifugation as described above. Alternatively, the virus was purified directly from the clarified supernatant by centrifugation through a 2 ml cushion of 20% sucrose in NT buffer for 2 hours at 40,000 rpm in the SW41 rotor.

Large scale production of [^{35}S]-methionine labeled VSV was carried out in suspension culture as described above. At 4 hours post infection, cells were recovered by centrifugation at 400 g for 10 min at room temperature at suspended to 2 x 10^6 cells/ml with methionine deficient Joklik modified MEM containing 2% dFCS and 2-4

uCi/ml of [^{35}S]-methionine (1000 Ci/mmole). Incubation was continued until 80% of the cells had lysed and the radioactive virus was purfied by differential and rate zonal centrifugation as above. Radioactive incorporation was determined by assaying hot insoluble TCA precipitable counts (Ghosh and Ghosh, 1972). An aliquot of labeled virus was spotted on a Whatman 3 MM filter paper and the sample was precipitated with 10% TCA. The filter paper was boiled in 5% TCA for 5 min and then washed with two changes of ice cold 5% TCA over a period of 20 min. The filter paper was rinsed with 95% ethanol:ether (1:1 v/v) followed by washing with ether. After drying under a heat lamp, the filter paper was placed in a precounted scintillation vial containing 10 ml of a PPO/toluene based scintillation fluid (4 gm of diphenyloxyzole in 1 liter of toluene). The radioactivity was determined in a scintillation counter.

(b) Labeling of Virus with [^{3}H] Fatty Acids

Labeling of VSV with [^{3}H]-palmitate was performed essentially as described by Schmidt and Schlesinger (1979). 9-10 [^{3}H]-palmitic acid, supplied in a benzene, was dried down under nitrogen and suspended in methanol to a final concentration of 5-10 mCi/ml. Labeled palmitate was added to L cell monolayers to 10 uCi/ml 12 hours prior to infection. Following infection and virus absorption, the medium was replaced with 5 ml of αMEM containing 2% NBCS and 20-40 uCi/ml of [^{3}H] palmitate. The final concentration of alcohol did not exceed 1%. The virus was purified as described above. Subsequently it was found that prelabeling the uninfected

cells did not appreciably enhance incorporation of radiocactivity into VSV so it was discontinued.

Labeling of VSV with the photoreactive fatty acid derivative [^{3}H]-diazirinophenoxy nonanoate was carried out as above, however, the fatty acid was saponified with an equimolar amount of filter sterilized NaOH prior to addition to the medium.

2.2.4 Radioactive Labeling of L Cells

L cell monolayers were infected with VSV to an m.o.i. of 10. At 4 hours post-infection, the medium was replaced with methionine deficient αMEM containing 2% dFCS and 20 uCi/ml of [^{35}S]-methionine. Incubation was continued for a further 1-2 hours after which the cells were harvested from the monolayer by gently scraping the cells into 1 ml of ice cold PBS with a rubber policeman. The cells were recovered by centrifugation, washed 3 times with cold PBS and finally suspended in a small volume of PBS. Cells were used immediately or stored at -20°C.

For labeling uninfected cells with [^{3}H]-palmitate or [^{3}H]-DAP nonanoate, L cells were grown to near confluency in 60 mm plates with αMEM containing 5% NBCS. The cells were rinsed with serum free αMEM and incubated at 37°C in serum free αMEM containing 20 uCi/ml of the [^{3}H] fatty acid for 2-8 hours. The cells were harvested and washed as described above.

Labeling of uninfected cells with [^{35}S]-methionine was carried out as described above, however, the labeling medium contained 2% dFCS and the labeling period was 1 hour.

2.2.5 Purification of Viral Components

(a) Purification of G Protein

Purified VSV was suspended at a concentration of 2mg/ml in 0.2% Triton X100, 10 mM Hepes (pH 7.6). The detergent concentration was increased to 2% and the suspension was incubated at room temperature for 1 hour with gentle mixing (Kelly et al, 1972). Alternatively, virus was disrupted by a 30 min incubation at room temperature in 50 mM octyl-D-glucoside, 10mM Hepes (pH 8.0), 1.5 mM DTT (Wilson and Lenard 1981). Nucleocapsids were removed from disrupted virus by centrifugation at 48,000 rpm in a Beckman Type 65 rotor at 4°C for 3 hours. The supernatant containing the G protein and the viral lipids was extracted with 2 vol of n-butanol for 25 min at 20°C. The phases were separated by low speed centrifugation and the interphase material was collected and precipitated with ethanol at -20°C. The precipitate was collected by centrifuging at 10,000 rpm for 30 min and the pellet was extracted 3 times with cold acetone and lyophilized. The purity of the material was determined by running an aliquot on a 10% polyacrylamide gel.

(b) Purification of M Protein

VSV (2 mg/ml) was disrupted with 50 mM octyl-D-glucoside, 10mM Hepes (pH 8.0), 250 mM NaCl for 30 min at room temperature. Nucleocapsids were removed by centrifugation as above and the supernatant containing G and M was applied to a 1.6 x 50 cm column of Ultrogel AcA 54 (LKB) and eluted with 50 mM Tris-HCl (pH 7.6), 0.4 M NaCl, 0.02% Na deoxycholate at 4°C. 2 ml fractions were collected.

The position of the M protein was determined by including [^{35}S]-methionine labeled VSV as a tracer prior to detergent disruption. Peak fractions were pooled and dialyzed extensively against water. The protein suspension was lyophilized and stored at -20°C.

(c) RNP Isolation

RNP cores, recovered from detergent disrupted virus as above, were further purified by suspending the crude RNP pellet in 10mM Tris-HCI (pH 8.0) 250 mM NaCl, 0.02% DTT and overlayihg on a preformed discontinuous gradient of 0.5 ml of 100% glycerol, 0.5 ml of 50% glyercol in 10 mM Tris-HCI (pH 8.0) and 1 ml of 10% glycerol in 10 mM Tris-HCI (pH 8.0), 100 mM NaCl in a 2 ml centrifuge tube (Carroll and Wagner, 1979). Centrifugation was at 45,000 rpm for 90 min in the Beckman Type 65 rotor. Purified RNP was collected from the interface between the 50 and 100% glycerol layers. The RNP was recovered by centrifugation and suspended in NT buffer.

(d) Isolation of Proteins from SDS Containing Polyacrylamide Gels

Purified VSV was fractionated on an SDS containing 10% polyacrylamide gel. Radioactive proteins of interest were localized on the wet gel by using an autoradiograph as a template. Nonradioactive proteins were localized by cutting a parallel strip from the gel and staining with Coomassie blue (0.25% Coomassie Blue, 50% methanol, 10% glacial acetic acid). The protein bands were excised from the gel and incubated at room temperature for 24 hours in 0.05M NH_4HCO_3 (pH 7.6) containing 0.01% SDS (approximately 0.5 ml of buffer was used per 1 cm^2 of gel). The elution buffer was

then passed through siliconized glass wool to remove particulate matter and the protein was precipitated with 9 vol of acetone at -20°C. The precipitate was recovered by centrifugation, dried under vacuum and suspended in an appropriate buffer.

2.2.6 SDS Polyacrylamide Gel Electrophoresis

SDS polyacrylamide gel electrophoresis was performed in slab gels utilizing the discontinuous buffer system as described by Laemmli (1970). Separating gels of the required acrylamide concentration were cast from a filtered stock solution of acrylamide (30:0.8 acrylamide: bis acrylamide, w/w). In addition, the separating gel contained 0.375 M Tris-HCl (pH 8.8), 0.1% glycerol and 0.025% TEMED. The solution was degassed prior to the addition of freshly prepared ammonium persulphate to 0.075% and SDS to 0.1%.

The stacking gel consisted of 5% acrylamide (30:0.8 acrylamide: bis acrylamide, w/w), 0.125 M Tris-HCl (pH 6.8), 0.1% SDS, 0.05% glycerol, 0.075% TEMED and 0.064% ammonium persulphate. The electrophoresis running buffer was Tris-glycine (pH 8.3), (3.02 gm Tris base, 14.4 gm glycine, 1 gm SDS per liter).

Samples were prepared for electrophoresis by the addition of an equal volume of twice concentrated electrophoresis sample buffer (0.125M Tris-HCl (pH 6.8), 6% SDS, 20% glycerol, 10% β - mercaptoethanol) and boiling for 2 min. A small amount of bromophenol blue was added to each sample as a tracking dye.

Following electrophoresis, the gel was immersed in staining solution (0.25% Coomassie blue, 50% methanol, 10% glacial acetic

acid) for 1 hour with shaking. The gel was destained with several changes of destaining solution (20% ethanol, 10% glacial acetic acid v/v) and dried on a sheet of Whatman #1 chromatography paper using a Bio-Rad dryer.

For the detection of [^{35}S] radioactivity, the dried gel was exposed to Kodak X-Omat XAR 5 film at room temperature for an appropriate length of time. [^{3}H] radioactivity and low levels of [^{35}S] radioactivity were detected by the fluorographic technique of Bonner and Lasky (1974). After electrophoresis, gels were soaked in destaining solution for 1 hour followed by two 30 min treatments with 100% DMSO. The gel was then treated for 3 hours with a solution of 22.2% PPO in DMSO (w/w). The PPO was precipitated by washing the gel in water for 1 hour and dried. Exposure was at -70°C for the appropriate length of time using prefogged Kodak X-Omat XAR 5 film (Laksy and Mills, 1975). In cases where radioactivity was determined directly by eluting material from gel slices, gel lanes were cut out from the slab gel and sectioned into 2-3 mm slices using a Bio Rad slicer. Radioactivity was eluted from each gel slice by incubating overnight in 0.5% SDS. The eluted radioactivity was determined by counting in Aquasol (3 gm PPO, 250 ml Triton-X114, 750 ml xylene).

2.2.7 Proteolytic Digestion of VSV

(a) Exopeptidase Digestion

Purified [^{35}S]-methionine labeled VSV (5 mg/ml) was incubated with aminopeptidase M (2 mg/ml) or a mixture of carboxypeptidases A and B (2 mg/ml) in 10mM Hepes (pH 7.6) in the

presence or absence of 2% Triton X100 for 4 hours at 30°C. The reaction was terminated by the addition of an equal volume of twice concentrated electrophoresis sample buffer and boiling for 2 min. Samples were analyzed by polyacrylamide gel electrophoresis.

(b) Preparation and Isolation of Spikeless Particles

Radiolabeled virus was digested with thermolysin (viral protein/enzyme 4:1, w/w), TPCK trypsin (viral protein/enzyme 2:1, w/w), or chymotrypsin (viral protein/enzyme 2:1, w/w) for 45 min at 35°C in a reaction mixture containing 10 mM Hepes (pH 7.6), 100 mM NaCl , and 10 mM $CaCl_2$. The reaction was stopped with 10 mM EGTA in the case of thermolysin and by the addition of 1 mM PMSF and incubation on ice for 10 min in the case of trypsin and chymotrypsin digestion. The reaction mixture was layered on a 2 ml cushion of 50% glycerol containing 10 mM Hepes (pH 7.6), 100 mM NaCl, 10 mM EGTA, 1 mM PMSF, and 10 mM DTT. Spikeless particles were isolated by centrifugation at 190,000g for 3 hours in a Beckman Type 65 rotor at 4°C. The pellet was washed with a buffer containing 10 mM Hepes (pH 7.6), 1 mM PMSF, and 1 mM EGTA and finally suspended in a small amount of this buffer.

2.2.8 Peptide Mapping

(a) Peptide Mapping by Limited Proteolysis

Peptide mapping by limited proteolysis in polyacrylamide gels was preformed essentially as described (Clevland et al, 1977). Bands containing proteins of interest were excised from dried gels using the autoradiogram as a guide. The gel slices were rehydrated in

0.125 M Tris-HCl (pH 6.8), 0.1% SDS, 1 mM EDTA, and placed in the sample wells of a second SDS polyacrylamide gel having an acrylamide concentration of 17.5 or 20%. The gel slice was overlaid with the above buffer containing 20% glycerol to fill in any spaces around the gel piece. Finally, 10 ul of 0.125 M Tris HCl (pH 6.8), 0.1% SDS, 1 mM EDTA, 10% glycerol containing varying amounts of Staphylococcus aureus V8 protease was added to each sample well. Electrophoresis was continued until the bromophenol blue tracking dye reached the interface between the stacking and separating gel at which point the current was turned off for 30 min. Electrophoresis continued in the normal manner after this and the gel was dried, fluorographed, and exposed to x-ray film as described.

(b) Tryptic Peptide Mapping

i) Ion Exchange Chromatography

Gel slices containing the protein to be analyzed were incubated in 0.2 M NH_4HCO_3 (pH 8.5) containing 100 ug TPCK trypsin per ml at 37°C for 16 hours. The solution was replaced with fresh ammonium bicarbonate and fresh trypsin and incubation was continued for an additional 4 hours. The solutions were combined, fresh trypsin was added, and incubation for a further 4 hours was carried out. At the end of the incubation, the solution was filtered and lyophilized. The tryptic peptides were resuspended in water and re-lyophilized. This was repeated three more times after which the tryptic peptides were suspended in 500 ul of 0.05 M pyridine-acetate pH 3.1 (8 mls pyridine, 139.2 ml glacial acetic acid, 1954 ml H_2O).

Tryptic peptides were separated on a water jacketed 0.9 x 22 cm column containing Aminex A5 cation exchange resin and maintained at 52°C (Toneguzzo 1977). Samples were loaded directly into the precolumn eluent with the aid of an Altex 4 way injection valve containing a 0.5 ml sample loop. Peptides were eluted at a flow rate of 1 ml/min with a Milton Roy mini pump using a linear gradient composed of 150 ml of 0.05 M pyridine-acetate, pH 3.1, and 150 ml of 1 M pyridine-acetate, pH 5.1, (40.3 ml pyridine, 37.6 ml glacial acetic acid, and H_2O to 500 ml). This was followed by elution with 50 ml of 2 M pyridine-acetate, pH 5.1, (32.5 ml pyridine, 27.8 glacial acetic acid, and H_2O to 200 ml). Fractions of 2 ml were collected directly into scintillation vials, dried, suspended in 200 ul of water and counted in 4 ml of Aquasol.

ii) Two Dimensional Peptide Mapping

Gel slices containing the protein to be analyzed were incubated with freshly prepared performic acid (formic acid/hydrogen peroxide 95:5, v/v) for two hours on ice. The gel pieces were removed, dried under vacuum and treated with TPCK-trypsin as above. Tryptic peptides were solubilized in 20% formic acid and spotted on a 20 x 20 Polygram Sil-N-HR silica thin layer plate. Phenol red was spotted in an opposite corner to serve as a tracking dye. Electrophoresis was in pyridine/acetic acid/water (1:10:100, v/v/v) and was discontinued when the phenol red had travelled approximately 15 cm. The plate was air dried and chromatographed at right angle to the direction of electrophoresis in butanol/pyridine/acetic acid/

water (5:4:1:5, v/v/v/v). After chromatography, the plate was dried in a chromatography oven at 90°C, sprayed with Enhance and exposed to Kodak X-Omat XAR film at -70°C for appropriate lengths of time.

2.2.9 Protein Sequencing

(a) Isolation of Peptides for Sequencing

Peptides to be sequenced were purified by polyacrylamide gel electrophoresis and excised from the wet gels by using the autoradiogram as a template. In cases where tritium labeled peptides were to be isolated, identical [^{35}S]-methionine labeled samples were run on the same gel in separate wells to facilitate visualization. Peptides were eluted from the gel slices by incubation at room temperature in 0.01 M NH_4HCO_3, 0.1% SDS, 1 mM PMSF, and 0.001 M β-mercaptoethanol. The eluted material was filtered through siliconized glass wool to remove particulate matter and dialyzed at 4°C for 48 hours against 0.1 M NaCl, 0.1% SDS followed by 24 hours against 0.05% SDS and finally for 24 hours against 0.01% SDS using a Spectrapore dialysis membrane (cut off, 3500; Spectrum Medical Industries). The dialysis buffer contained in addition 0.01 M NH_4HCO_3, 0.001 PMSF, and 0.001 M β-mercaptoethanol. The dialyzed material was lyophilized and suspended in 0.5 ml of 20% formic acid prior to sequence analysis. A small aliquot was dried down and checked for purity by re-electrophoresis.

(b) Automated Protein Sequencing

Automated protein sequencing was carried out in a Beckman

890C sequencer equipped with a cold trap modification and using the Beckman 0.1 M Quadrol program (No. 345801). The program was modified to exclude nitrogen delivery to the fraction collector during those steps in which the fraction collector was under vacuum.

Samples were applied to the cup and dried down with the Beckman sample application sub routine program (no. 345871). In cases where polybrene was used to assist retention of protein in the cup, 3-5 mg of polybrene (in water) along with 100 nmoles of glycylglycine were added to the cup and dried (Hunkapillar and Hood, 1978). The polybrene was precycled with three complete degradation cycles prior to the addition of sample. In each experiment, the protein film was conditioned by running an initial sham cycle in the absnece of PITC delivery.

For nonradioactive samples, the anilinothiozolinones (ATZ) were converted to their corresponding PTH derivatives by treatment with 0.2 ml of 25% trifluoroacetic acid containing 1 ul/ml of ethanethiol for 10 min at 80°C. After drying the sample under nitrogen, the PTH derivatives were suspended in ethyl acetate or HPLC grade methanol and identified by HPLC. A known amount of the PTH derivative of norleucine was included in each sample prior to HPLC analysis to serve as an internal control.

Microsequencing of proteins labeled with 1 amino acid or a mixture of [^{35}S]-methionine and a [^{3}H] amino acid was carried out as above except that hen egg white lysozyme was used as a carrier. The butyl chloride extracts were dried and suspended in 250 ul of

methanol. 80% of this was directly transferred into scintillation vials and the radioactivity was determined. The remaining 20% was converted with TFA to the corresponding PTH derivatives and used to identify the PTH-amino acids arising from lysozyme.

In cases where more than one [^{3}H] amino acid was present in a sample, the samples were converted to their PTH derivatives, mixed with the corresponding unlabeled PTH amino acids and analyzed by HPLC. The fractions containing the PTH amino acids were collected and the radioactivity was determined.

(c) PTH Amino Acid Analysis

PTH amino acids were separated on an Ultrosphere octadeclytrimethyloxysilane (ODS) column maintained at 45°C in an Altex model 322 high performance liquid chromatography system equipped with two model 110A pumps and a model 421 microprocessor controller. The column was equilibrated with buffer A (4.25 mM Na-acetate, pH 5.01, containing 5% tetrahydrofuran) and was eluted with a linear gradient of 0-40% B (10% tetrahydrofuran in acetonitrile) for 20 min followed by isocratic elution with 40% B for a further 20 min (Somack, 1980). Nonradioactive PTH amino acids were quantitated with an Altex model C-RIA integrator using the normalization mode. Aqueous buffers and organic buffers were filtered through Millipore HAWP (0.45 um) and FHUP (0.5 um) filters, respectively, and degassed prior to use.

2.2.10 Phospholipid analysis

Phospholipids were extracted from labeled L cells or VSV by

the Bligh-Dyer procedure (Bligh and Dyer, 1959). A sample of cells or virus was suspended in 1 ml of buffer and extracted with 3.75 ml of methanol/chloroform (2:1 v/v). After shaking at room temperature for 10 min the precipitated material was centrifuged in a bench top centrifuge and the supernatant was removed. The precipitated material was extracted twice more with 4.75 ml of methanol/chloroform/water (2:1:0.8 v/v/v) and the supernatants were combined (14.25 ml) and mixed with 3.75 ml of chloroform and 3.75 ml of water. After mixing, the phases were separated by a brief centrifugation and the lower chloroform phase was taken to dryness and suspended in chloroform/methanol (1:1 v/v). Aliquots were spotted on a Silica HIB5 thin layer plate (J.T. Baker) and chromatographed with chloroform/methanol/water (65:25:4 v/v/v; Kates, 1972) along with nonradioactive phospholipid standards. Radioactive spots were visualized by autoradiography while the positions of the phospholipid markers were determined by staining with I_2 vapour.

2.2.11 Analysis of Fatty Acids Attached to G Protein

(a) Hydroxylamine Treatment

VSV labeled with [^{3}H]-palmitate or [^{3}H] DAP-nonanoate was run in duplicate along with [^{35}S]-methionine labelled VSV on an SDS containing 10% polyacrylamide gel. One half of the gel was directly processed by fluorography while the other half was treated with hydroxylamine by soaking the fixed and washed gel in 200 mls of 1 M NH_2OH-HCl (titrated to pH 6.6 with NaOH) for 1 to 5 hours (Schlesinger et al., 1980; Omary and Trowbridge, 1981). The gel

was subsequently rinsed with water and processed by fluorography.

(b) Transesterfication

G protein was isolated from [^{3}H] fatty acid labeled virus by detergent extraction with Triton X100 as described except that the butanol extraction was omitted. The G protein was delipidized by series of sequential extractions with organic solvents according to the Bligh-Dyer procedure as described above. The G protein was dried under vacuum to remove traces of organic solvent and the covalently attached fatty acid residues were transesterfied by treatment with 1 ml of 0.1 M methanolic KOH for 10-20 min at 23°C (Schmidt and Schlesinger, 1979). The solution was acidified with HCl and the fatty acid methyl esters were recovered by four 1 ml extractions with petroleum ether. After drying down the petroleum ether, the residue was suspended in a small volume of petroleum ether and an aliquot was analyzed by thin layer chromatography on Silica HlB5 plates using petroleum ether/ether (1:1 v/v). Following chromatography, lanes corresponding to the sample were scrapped directly into scintillation vials, eluted with methanol, and the radioactivity was determined.

Alternatively, fatty acid methyl esters were suspended in methanol and analyzed by HPLC on an Ultrosphere ODS column. The methyl esters were eluted with a linear 25 min gradient of 80-95% B (100% methanol) followed by a 2 min rise to 100% B followed by 100% B for an additional 13 min. The flow rate was 2 ml/min and 0.5 min fractions were collected and the radioactivity determined.

2.2.12 Photolysis

Samples were purged with water saturated nitrogen for 45 min prior to photolysis. Photolysis was carried out in a 1 ml capacity quartz cell (1 cm path length) held in a metal cell holder located at the focal point of a 1000 watt xenon/mercury lamp (Schoeffel model LPS 225 HR). The light was first passed through a filter (Corning 7-51) contained in a water filled optical glass cell (5 x 5 cm); a second such filter was attached to the window of the metal cell housing. The resulting light has its maximum intensity near 360 nm, while all light below 300 nm is removed. Irradiation times were controlled with a manually operated metal shutter.

2.2.13 Immunoblotting

Immunoblotting of proteins on nitrocellulose was performed by the method described by Towbin et al., (1979). Proteins were electrophoresed on a 10% polyacrylamide gel. Following electrophoresis, the gel was soaked in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol (pH 8.3)) and then placed on a nitrocellulose membrane (0.45 um, Schleicher and Schuell) backed with 3 sheets of Whatman #1 filter paper. The gel and filter paper was sandwiched between electrode containing plastic grids with the nitrocellulose sheet facing the +ve electrode and the assembly was immersed in an electrophoretic chamber containing transfer buffer. Proteins were transferred to the nitrocellulose membrane by electrophoresis for 20 hours at 8 V/cm. Following electrophoresis, the blot was removed and soaked in saline (0.9% NaCl, in

10 mM Tris-HCl (pH 7.4)) containing 3% BSA for 1 hour at 40°C. The blot, or parts thereof, was rinsed with 500 ml of saline and incubated overnight at room temperature with shaking with antiserum appropriately diluted into saline containing 3% BSA. The sheet was then washed with 5 changes of saline over a period of 30 min and incubated overnight at room temperature with [^{125}I]-protein A diluted to 1 x 10^6 cpm/ml in 3% BSA-saline containing 0.01% NaN_3. The blot was then extensively washed with saline (5 changes over 30 min) and dried. Radioactivity was detected by autoradiography using Kodak X-Omat XAR5 film and an intensifying screen (Corning-Dupont).

2.2.14 Reconstitution of Purified Viral Proteins

Incorporation of M protein into synthetic lipid vesicles was carried out essentially as described for the influenza virus M protein (Gregoriades, 1980; Greogoriades and Rangione, 1981). Egg phosphatidylcholine was dried under a stream of nitrogen followed by vacuum dessication for 4 hrs. M protein, isolated from gels or by detergent extraction, was solubilized in buffer (0.05 M Tris-HCl (pH 7.6), 0.1 M NaCl, 0.005 M KCl, and 1% Na deoxycholate) and added to the dried film of phosphatidylcholine in a final ratio of 1:10 (protein:lipid, w/w). The mixture was vortexted for 5 min and dialyzed against several liter changes of dialysis buffer (0.05 M Tris-HCl (pH 7.6), 0.1 M NaCl, 0.005 M KCl, 0.001 M NaN_3) for 48 hrs at 4°C. The vesicles were made 40% with respect to sucrose and overlaid with a linear gradient of 5-30% sucrose (in 0.05 M Tris-HCl (pH 7.6), 0.1 M NaCl, 0.005 M KCl) and centrifuged for 18 hours at

40,000 rpm in a Beckman SW41 Ti rotor at 4°C. After centrifugation, fractions were collected from the bottom of the tube and aliquots were assayed for radioactivity. Peak fractions were pooled, diluted, and sedimented at 40,000 rpm in a Beckman SW41 Ti rotor for 1.5 hrs at 4°C.

2.2.15 Proteolytic Digestion of Vesicles

Reconstituted vesicles containing M protein were isolated from flotation gradients as described above and incubated with either TPCK-trypsin or thermolysin (protein:enzyme 3:1, w/w) in 50 mM NH_4HCO_3 (pH 8.0) at 37°C for 40 minutes. The samples were made 40% with respect to sucrose and isolated from digestion products by flotation in a sucrose gradient as described above.

2.2.16 Reconstitution with Preformed Vesicles

Preformed vesicles were made by suspending dried phosphatidylcholine to a concentration of 1 mg/ml in 0.05 M Tris-HCl (pH 7.6), 0.1 M NaCl, 0.005 M KCl, 1% Na deoxycholate and dialyzing out the detergent. Detergent free M protein (detergent removed by dialysis) was added to the preformed phospholipid vesicles, vortexed for 5 min, made 40% with respect to sucrose, and analyzed by flotation centrifugation as described above.

2.2.17 In Vitro Protein Synthesis

(a) Preparation of VSV Specific RNA

VSV specific RNA was isolated from infected L cells as described (Tonneguzzo and Ghosh, 1976). L cells growing in suspension culture were infected with VSV to an m.o.i. of 30. At

4.5 hrs post infection, the cells were quickly cooled for 10 min by the addition of frozen PBS. The cells were pelleted at 800 g for 10 min and washed 3 times with washing buffer (35 mM Tris-HCl (pH 7.5), 146 mM NaCl). The pellet was suspended in 2 vol of swelling buffer (10 mM Hepes pH 7.6, 15 mM KCl, 1.5 mM $MgAc_2$, 6 mM β - mercaptoethanol) and incubated on ice for 10 min. The swollen cells were dounced with 25 strokes in a 15 ml glass homogenizer and centrifuged at 4,000 g for 10 minutes at 0°C. The S-4 extract so prepared was made 0.5% with respect to SDS and an equal vol of mRNA extraction buffer was added (50 mM Tris-HCl pH 9.0, 100 mM NaCl, 0.1% SDS). An equal vol of chloroform/phenol (1:1, v/v) was added and the solution was vortexed at room temperature for 10 min. The phenol solution was saturated with mRNA extraction buffer. The phases were separated by centrifugation at 7,500g for 10 min at 4°C and the aqueous phase was re-extracted with an equal vol of chloroform/phenol (1:1, v/v). The aqueous phases were combined, made 0.4 M with respect to LiCl and the RNA was precipitated with 2.5 vol of 95% ehtanol at -20°C overnight. The precipitate was collected by centrifugation at 10,000 rpm for 45 min and re-precipitated with ethanol. The final RNA pellet was dried under vacuum and dissolved in 200 ul of sterile water and stored at -90°C.

(b) Preparation of Reticulocyte Lysate

Reticulocyte lysates were prepared as described by Villa-Komaroff et al., (1974). New Zealand white rabbits were made anemic by subcutaneous injection with 2 ml of 1.2% acetylphenylhydrazine on

each of six consecutive days. The rabbits were bled by canulation on the eighth day. The blood was collected in one fifth volume of chilled saline (140 mM NaCl, 1.5 mM $MgAc_2$, 5 mM KCl) containing 0.001% heparin, filtered through cheesecloth and washed three times with saline by centrifugation at 3,500 rpm for 5 min. The final centrifugation was at 7,000 rpm and the packed cells were lysed with an equal vol of cold H_2O at 0°C for 1 min. The lysate was centrifuged at 10,000 rpm for 40 min at 4°C and the supernatant was saved. The lysate was nuclease treated as described by Pelham and Jackson (1976). 1 ml of lysate was incubated with approximately 50 units of micrococcal nuclease for 10 min at 20°C. The reaction was stopped by the addition of 2.5 ug of rabbit tRNA and EGTA to 22 mM. The treated lysate was aliquoted and stored at -90°C.

(c) Protein Synthesis

Protein synthesis was carried out in a 50 ul reaction volume containing 30 ul of the nuclease treated reticulocyte lysate, 20 mM Hepes (pH 7.6), 90 mM KCl, 1 mM ATP, 180 mM GTP, 2 mg/ml creatine phosphate, 50 ug/ml creatine phosphokinase, 3.5 ug/ml hemine, 0.6 mM Mg acetate, 100 uCi/ml [^{35}S]-methionine and 1-5 ug of VSV RNA. Incubation was at 30°C for 90 min. Radioactive incorporation was determined by hot TCA precipitation and products were examined by polyacrylamide gel electrophoresis.

(d) Hybridization Arrested Translation

Hybridization arrested translation was carried out as described (Preston and McGeogh, 1981) using plasmid pM32 which

contains a cDNA copy of the 3' end of the mRNA sequence of M protein (Rose and Iverson, 1979). 1.5 ug of pM32 was linearized by digestion with 2 units of EcoRI endonuclease in a reaction mixture containing 50 mM Tris-HCl (pH 7.5), 10 mM $McCl_2$, 50 mM NaCl, and 1 mM DTT for 1 hr at 37°C. The linearized plasmid was phenol extracted, ethanol precipitated, and suspended in 2 ul of water. 80 ul of 95% deionized formamide containing 60 mM Pipes (pH 7.4) was added and the plasmid was heat denatured at 95°C for 5 min. 10 ul of the RNA preparation obtained from VSV infected L cells along with 8 ul of 5 M NaCl was added and the mixture was incubated at 58°C for 1 hr. After this incubation period, 100 ul of cold sterile water and 15 ug of wheat germ tRNA was added and the mixture was made 0.2 M with respect to Na-acetate and precipitated with 2.5 vol of ethanol at -20°C. The RNA was pelleted at 12,000g for 15 min, dried, suspended in a small amount of water and used for translation in the nuclease treated reticulocyte lysate.

2.2.18 Reconstitution of In Vitro Synthesized Proteins

Following the completion of protein synthesis *in vitro*, the reaction mixture was diluted to 1 ml with 50 mM Tris-HCl (pH 7.6), 0.1 M NaCl, 0.005 M KCl and dialyzed against the same buffer for 12 hr to remove unincorporated radioactive methionine. Reconstitution was accomplished by adding 100 ug of preformed phosphatidylcholine vesicles suspended in 50 mM Tris-HCl (pH 7.6), 0.1 M NaCl, 0.005 M KCl to the *in vitro* synthesized proteins and vortexing for 5 min. After making the mixture 40% with respect to sucrose, the vesicles

were overlaid with a discontinuous sucrose gradient composed of 10 ml of 35% sucrose in 50 mM Tris-HCl (pH 7.6), 0.1 M NaCl, 0.005 M KCl and 1 ml of this buffer without sucrose. The vesicles were floated through the sucrose by centrifugation at 40,000 rpm for 18 hr at 4°C in the Beckman SW41 Ti rotor. The vesicles were collected by taking fractions from the bottom of the tube or directly by collecting the lipid band at the interface between the 0 and 35% sucrose layers. The vesicles were diluted and concentrated by centrifugation at 40,000 rpm for 1.5 hrs at 4°C in the Beckman SW41 Ti rotor.

2.2.19 Protein Estimation

Protein concentrations were estimated by the method of Lowry et al., (1951) using bovine serum albumin as the standard and measuring the absorbance at 660 nm.

2.2.20 E. coli Transformation with Plasmid DNA

100 ml of L broth (10 gm Bacto-tryptone, 5 gm of Bacto-yeast extract, 10 gm NaCl (pH 7.5 with NaOH) per liter, containing 0.4% filter sterilized glucose) was innoculated with 1 ml of an overnight culture of E. coli LE392. The culture was incubated with vigorous shaking at 37°C until an OD_{660} of 0.6-0.7 was reached. The cells were collected by centrifugation at 4,000g for 5 min at 4°C and suspended in 35 ml of a sterile solution of 50 mM $CaCl_2$. After incubation on ice for 20 min, the cells were collected by centrifugation as above and suspended in 2 ml of sterile 100 mM $CaCl_2$. The cells were stored at 90°C in 200 ul aliquots following the addition of sterile glycerol to 15%.

For transformation, 100 ul of sterile 50 mM $CaCl_2$ containing 100-300 ng of plasmid DNA was added to 100 ul of competent cells, prepared as above, and incubated on ice for 20 min. The suspension was heat shocked at 42°C for 2 min and placed on ice for a further 20 min. The cells were then diluted with 3 ml of prewarmed L broth and incubated at 37°C without shaking for 1-3 hours. Aliquots of an appropriate dilution, in L broth, of the cells were spread on agar plates (L broth containing 1.85 gm/ml of Bacto-agar) containing the selective antibiotic (tetracycline at 10 ug/ml or ampicillin at 20 ug/ml) and incubated at 37°C. Colonies were visualized 16-20 hr later.

2.2.21 Large Scale Purification of Plasmid DNA

A colony isolate of E. coli, harbouring the plasmid to be purified, was grown overnight in 10 ml of L broth at 37°C with vigorous shaking. The next morning, the culture was diluted to 1 liter with L broth containing the selective antibiotic (usually 10 ug/ml tetracycline) and incubated at 37°C with shaking until the culture reached late log phase (OD_{660nm}=0.6). Chloramphenicol was added to 170 ug/ml and incubation continued for a further 16-20 hr. Cells were harvested by centrifugation at 4,000g for 10 min at 4°C, washed with PBS, and re-pelleted. The pellet was resuspended in 3.33 ml of mix K (50 mM Tris-HCl (pH 8.0), 25% sucrose) and 0.66 ml of a solution of 45 mg/ml of lysozyme was added. After incubation on ice for 10 min, 1.66 ml of 0.2 M EDTA (pH 8.0) was added and incubation on ice continued for a further 10 min. Cells were lysed by the

addition of 5.33 ml of lysis buffer (50 mM Tris-HCl (pH 8.0), 0.1% Triton X100, 50 mM EDTA) and incubated on ice for 10 min. The lysate was cleared by centrifugation at 45,000 rpm for 50 min at 4°C in the Beckman 50.2 Ti rotor and the supernatant was collected. CsCl and ethidium bromide was added to the supernatant such that the final density of CsCl was 1.59 g/ml and the final concentration of ethidium bromide was 600 ug/ml. The solution was centrifuged at 36,000 rpm for 60 hours at 20°C in the Beckman 65 rotor. The DNA bands were visualized by long wave length UV illumination and the lower band, which contains the covalently closed circular plasmid DNA, was collected by puncturing the side of the tube with a syringe. The plasmid containing solution was extracted 3 times with isoamyl alcohol saturated with CsCl to remove the ethidium bromide. The aqueous phase, containing the DNA, was dialyzed against several changes of TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA). The solution was made 0.25 M with respect to sodium acetate (pH 5.2) and precipitated at -20°C after the addition of two volumes of cold ethanol. The precipitate was collected by centrifugation at 10,000g for 30 min, washed with 70% cold ethanol, dried briefly under vacuum, and finally suspended in a small volume of TE buffer. DNA concentration was estimated by absorbance at 260 nm (an A_{260} of 1 is equivalent to approximately 50 ug duplex DNA/ml).

2.2.22 Restriction Enzyme Digestion

Restriction enzyme digestions were carried out in 25-50 ul reaction volumes containing 1-3 ug of plasmid DNA, 100 ug/ml

nuclease free BSA, and an appropriate amount of restriction enzyme as recommended by the supplier. The following buffers, diluted from 5 times stock solutions were used for the various enzymes:

Bam HI; 20 mM Tris-HCl (pH 7.0), 100 mM NaCl, 7 mM $MgCl_2$, 2 mM 2-mercaptoethanol.

Bgl II; 100 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 60 mM NaCl.

Hind III; 20 mM Tris-HCl (pH 7.4), 7 mM $MgCl_2$, 60 mM NaCl.

Pst I; 20 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 50 mM $(NH_4)_2SO_4$.

Sau3A; 6 mM Tris-HCl (pH 7.5), 50 mM NaCl, 6 mM $MgCl_2$.

The reaction mixtures were incubated at 37°C for 1 hour and terminated by the addition of loading buffer (5 times stock = 4M urea, 50% sucrose, 50 mM EDTA (pH 8.0) and 0.1% bromophenol blue) and heating at 56°C for 2 min. In the case of double enzyme digestion, digestion was first carried out with the enzyme requiring lower salt concentration. After incubation at 37°C for one hour, the enzyme was heat inactivated and the reaction was supplemented with NaCl and other components required by the second enzyme. The reaction was incubated for a further 60 min after the addition of the second enzyme and terminated as above by the addition of loading buffer. Samples were then analyzed by gel electrophoresis.

2.2.23 Separation of DNA Fragments by Gel Electrophoresis

(a) Agarose gel Electrophoresis

Agarose was added to TBE electrophoresis buffer (89 mM Tris 89 mM boric acid (pH 8.3), 1 mM EDTA) to the desired concentration (usually 1-2%) and placed over a bunsen burner until the agarose

dissolved. The solution was cooled to 50°C and poured onto a 16 x 20 cm glass plate contained in a horizontal gel apparatus. A sample well comb was placed at one end of the molten gel (close to the -ve electrode), and the gel was allowed to polymerize for 30-45 min. After the gel had solidified, the comb was removed and samples loaded into the preformed wells. The buffer chambers were filled with TBE until the buffer contacted the edges of the gel. Electrophoresis was at 2-3 V/cm and was continued until the bromophenol blue tracking dye reached the end of the gel. The gel was removed from the apparatus and stained for 30 min with 3 gel vol of a solution of 0.5 ug/ml of ethidium bromide. The gel was photographed as described below.

(b) Polyacrylamide Gel Electrophoresis

Polyacrylamide gel solution was made up in a 50 ml volume containing 6% acrylamide, 0.17% bis acrylamide in TBE electrophoresis buffer. After the addition of 0.4 ml of freshly prepared 10% ammonium persulphate and 15 ul of TEMED, the solution was poured between two glass plates and allowed to polymerize. The plates were clamped into a vertical slab gel electrophoresis system and the samples were loaded. Electrophoresis was at 2 V/cm until the bromophenol blue tracking dye reached the bottom of the gel. The gel was removed and stained with ethidium bromide as described above.

(c) Photography

Agarose or polyacrylamide gels were photographed under short wave UV illumination from a Chromo-Vue model 0-63 transilluminator (Ultra-Violet Products Inc., San Gabriel, Calif.). Photographs were

taken with a Polaroid MP-3 land camera containing Type 107 land film and equipped with a 39 mm SR60-2 filter.

2.2.24 Recovery of DNA fragments from Gels

DNA bands, separated by agarose or polyacrylamide gel electrophoresis, were visualized by long wavelength UV illumination of the ethidium bromide stained gels and the desired fragment was excised with a razor blade. The gel piece containing the DNA was placed in a dialysis bag containing 0.5 X TBE and placed in an electrophoresis chamber containing this buffer in an orientation parallel to the electrodes such that the buffer just covered the dialysis bag. The DNA was electroeluted out of the gel by applying a current of 100 V for 3-4 hours. The polarity of the current was reversed for 2 min to release the DNA from the inner wall of the dialysis membrane and the DNA containing solution was recovered. The solution was filtered through siliconized glass wool, extracted twice with phenol (saturated with 10 mM Tris-HCl (pH 7.6), 1mM EDTA)/ chloroform/isoamyl alcohol (25/24/1, v/v/v) and once with an equal volume of chloroform/isoamyl alcohol (24/1, v/v). The aqueous phase was extracted twice with an equal vol of water saturated ether. The last traces of ether were removed with a gentle stream of N_2 and the DNA was precipitated at -20°C overnight after the addition of 2 vol of cold ethanol and sodium acetate (pH 5.2) to 0.3 M. The DNA was recovered by centrifugation at 10,000g for 30 min at 4°C. The alcohol precipitation step was repeated and the final DNA pellet was washed with 70% ethanol, dried briefly, and suspended in a small

volume of TE buffer or 0.1X SSC (1 X SSC = 0.15 M NaCl, 0.015 M sodium citrate).

2.2.25 Subcloning DNA Fragments into Vector Plasmid

Ligation of vector DNA with the desired DNA fragments was carried out in a 20 ul reaction volume containing 1-200 ng of vector DNA cleaved with the appropriate restriction enzymes, 2-400 ng of the DNA fragments to be subcloned, 2 ul of 10X ligation buffer (0.1 M $MgCl_2$, 0.2 M DTT, 0.5 M Tris-HCl (pH 7.7), 0.01 M ATP, 500 ug/ml nuclease free BSA) and 0.1 - 1 units of T_4 DNA ligase. The mixture was incubated at 4°C for 16-20 hr and the reaction was terminated by heating at 65°C for 10 min. The reaction was diluted with sterile $CaCl_2$ to 100 ul to give a final $CaCl_2$ concentration of 50 mM and used to transform competent LE392 as described earlier.

2.2.26 Indentification of Recombinant Plasmids

Positive clones (exhibiting the proper antibiotic resistent phenotype) were screened by isolating plasmid DNA from individual colonies by the following rapid procedure. Single colony isolates were inoculated in 5 ml of L broth containing 10 ug/ml of tetracycline and allowed to grow to saturation overnight at 37°C. 1.5 ml of the culture was poured into an Eppendorf centrifuge tube and the cells were recovered after a brief centrifugation in an Eppendorf centrifuge. The medium was removed by aspiration and the pellet was suspended in 350 ul of a solution containing 10 mM Tris-HCl (pH 8.0), 8% sucrose, 0.5% Triton-X100, and 50 mM EDTA (pH 8.0). The cells were lysed by the addition of 25 ul of freshly

prepared lysozyme (10 mg/ml in 10 mM Tris-HCl, (pH 8.0)) and vortexing for 3 sec. The tube was placed in a boiling water bath for 45 sec and immediately centrifuged for 10 min at room temperature in an Eppendorf centrifuge. The supernatant was decanted into a new tube (as an optional step, 35 ul of 0.2 mg/ml DNase free RNase was added and the reaction incubated at 37°C for 10 min to remove contaminating RNA). 200 ul of 5 M NH_4-acetate and 1 ml of isopropanol was added to the supernatant and the DNA was precipitated at -20°C for 10 min. The DNA was recovered by centrifugation at 12,000g for 15 min. The precipitation was repeated and the DNA pellet was washed with 70% ethanol, dried, and suspended in 100 ul of TE buffer. 20-40 ul of this solution was digested with the desired restriction enzyme and analyzed by gel electrophoresis.

2.2.27 DNA Mediated Gene Transfer

DNA mediated gene transfer was carried out by the calcium-phosphate transfection procedure (Graham and van der Eb, 1973; Wigler et al., 1979). COS-1 cells were plated in 60 mm culture dishes at a confluency of 50-60% 1 day prior to transfection. For each plate, approximately 5 ug of plasmid DNA along with 10 ug of salmon sperm carrier DNA was added to 0.25 ml of sterile 0.25 M $CaCl_2$. The DNA/$CaCl_2$ solution was added dropwise to an equal volume of 2X HBS (280 mM NaCl, 50 mM Hepes, 1.5 mM sodium phosphate, (pH 7.1)). Bubbles were gently blown into the solution using a 1 ml cotton plugged sterile plastic pipette while the DNA was being added. The calcium phosphate/DNA precipitate was allowed to form without

agitation for 20 min at room temperature and was added directly to the culture dish containing 4 mls of Dulbecco's modified MEM supplemented with 7% fetal bovine serum. After a 4 hr incubation at 37°C, the medium was removed and the cells were shocked with 1 ml of media containing 25% glycerol for 1 min at room temperature. The shock medium was removed and the cells were washed twice with 2 ml of warm saline (0.9% NaCl, 0.01 M Tris-HCl (pH 7.4)) and then incubated with 4 ml of fresh growth medium for 40-48 hrs. After the incubation period, the medium was removed and saved, and the cells were harvested by scrapping with a rubber policeman and washed three times with cold PBS prior to further processing.

For metabolic labeling of the transfected cells, the cells were washed 40 hr post transfection, and the medium was replaced with 2 ml of methionine deficient Dulbecco's modified MEM containing 2% dialyzed fetal calf serum. After a 30 min incubation at 37°C to deplete endogenous methionine, [^{35}S]-methionine was added to a level of 50 uCi/ml and the cells were incubated for a further 8-20 hrs. The labeling medium was removed and the cells were harvested as above.

Cell extracts, to be analyzed directly by gel electrophoresis, were prepared by the addition of SDS polyacrylamide gel electrophoresis sample buffer to the packed cells, boiling for 2 min at 100°C, and removing the particulate matter by centrifugation at 12,000g for 3 min before electrophoresis. For the direct analysis of unlabeled proteins present in the growth medium, the

medium that was collected 40-48 hrs post transfection was clarified by centrifugation at 10,000 g for 10 min and precipitated with 9 vol of cold acetone at -20°C. The precipitated material was collected by centrifugation at 10,000g for 30 min, washed with cold ethanol, dried briefly, and solubilized in SDS polyacrylamide gel sample buffer as above, and analyzed by polyacrylamide gel electrophoresis.

2.2.28 Immunoprecipitation

Packed, washed cells, labelled with [^{35}S]-methionine, were lysed by the addition of 0.5 ml of RIPA buffer (50 mM Tris-HCl (pH 7.2), 150 mM NaCl, 1% Triton-X100, 0.1% SDS, 1% Na-deoxycholate, and 100 Kallikrein units of Trasylol). Cell nuclei and insoluble material was removed by centrifugation at 12,000g for 5 min at 4°C and 5-10 ul of antiserum was added to the supernatant. After incubation at 4°C for 60 min under continuous rotation, 0.2 ml of a 10% suspension of Protein A Sepharose (in RIPA buffer) was added and incubation continued for an additional 3 hrs. The protein A Sepharose beads were collected by brief centrifugation in an Eppendorf centrifuge, washed 4 times with RIPA buffer and 3 times with TBS (50 mM Tris-HCl (pH 7.2), 150 mM NaCl, 100 units Trasylol). The antibody-antigen complexes were released by the addition of SDS polyacrylamide gel sample buffer and boiling for 2 min. The beads were removed by centrifugation and the supernatant subjected by gel electrophoresis.

RESULTS

SECTION 3.1: TOPOGRAPHY OF G PROTEIN IN THE VIRAL ENVELOPE

Membrane proteins play critical roles in a number of diverse and essential cell surface related biological processes such as cell-cell recognition, surface mediated transmembrane signaling, regulation of cell growth and division, ligand receptor interaction, and metabolite transport (Sergest and Jackson, 1977). Since these functions are correlated with an absolute structural asymmetry between the external and internal surfaces of the membrane, a molecular understanding of the topological organization of membrane proteins relative to the bilayer is an important prerequisite to a understanding of their function as well as the processes involved in membrane assembly.

As described in more detail in the Introduction, the glycoprotein G of VSV is inserted into microsommal vesicles in a co-translational event involving the cleavage of an amino terminal transient signal sequence (Toneguzzo and Ghosh, 1977, 1978; Irving et al., 1979; Lingappa et al., 1978a; Rothman and Lodish, 1977; Katz et al, 1977; Katz and Lodish, 1979). Unlike secretory proteins, G is not completely extruded into the luminal space of these membrane vesicles but exists as a transmembrane protein with approximately 30 amino acids at its carboxy terminus exposed to the cytoplasmic face of the membrane. The same general topology is observed in cytoplasmic membrane preparations containing G protein obtained from VSV infected

cells (Morrison and McQuain, 1979).

Thus, there are at least two distinct regions or domains on the G molecule which are important in the biogenesis of this protein; the amino terminal extension sequence which is required for membrane insertion but is removed in the completed polypeptide, and a domain near the COOH terminus which may serve to stop the transfer of the polypeptide into the membrane vesicles and thus insure a proper orientation of the G protein during its intracellular maturation.

Since the intravesicular space of membrane vesicles and the rough endoplasmic reticulum is topologically equivalent to the extracytoplasmic surface of the plasma membrane (Palade, 1975), one would expect that the G protein would be orientated in the viral envelope with the COOH terminus on the inside of the virus and the amino terminus exposed to the outside. It is not known, however, whether the asymmetry of G protein as observed in the microsommal vesicles is maintained in the viral envelope. This question is of importance not only from the point of view of examining which domains of G are responsible for insuring its correct and ultimate disposition, but also in understanding its mode of migration from its initial point of synthesis to its final destination in the completed virion.

It was, therefore, decided to examine in detail the topography of G protein in the viral envelope. The primary technique used was controlled proteolysis which is based on the fact that the lipid bilayer serves as a permeability barrier to exogenously added

proteases. Thus, under the proper conditions, only those components that are external to the bilayer are susceptable to proteolytic degradation.

One of the first questions asked was whether the NH_2- and COOH-termini of G protein are protected by the lipid bilayer in intact virions. If the topology of G found in the intracellular membrane system is the same as that in the viral envelope, then one would expect at least the COOH terminus to be protected from proteases. To test this, the two exoproteases, aminopeptidase, and carboxypeptidase were used. These enzymes sequentially degrade proteins from their NH_2- and COOH-termini, respectively. Aminopeptidase M will remove all amino acids from the amino terminus of a polypeptide chain, however, it will cleave sequences of the type X-Pro very slowly, if at all (van Beynum et al., 1977). A mixture of carboxypeptidases A and B, which was routinely used to ensure broader substrate specificity, will remove all amino acids from the C-terminus of a polypeptide chain with the exception of a C-terminal proline residue (Narita, 1970).

When [^{35}S]-methionine labeled VSV was incubated with either aminopeptidase or a mixture of carboxypeptidases A and B, no reduction in the size of G protein was observed (Fig. 3.1.1, lanes b and d). Thus, both the NH_2 - and COOH - termini of G are unavailable as substrates for these proteases in the intact virion. However, when the nonionic detergent Triton X100 was included in the incubation mixture, a definite reduction in the size of G was

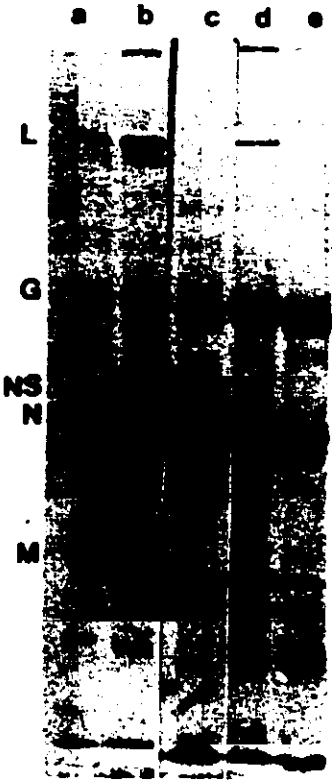


Figure 3.1.1: Polyacrylamide gel analysis of VSV treated with aminopeptidase or carboxypeptidase.

[^{35}S]-Methionine labeled VSV was treated with aminopeptidase M or a mixture of carboxypeptidase A and B in the absence or presence of 1% Triton X100 as described in Methods and analyzed on a 10% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV marker; lanes b and c, VSV treated with a mixture of carboxypeptidase A and B in the absence and in the presence of Triton X100, respectively; lanes d and e, VSV treated with aminopeptidase M in the absence and in the presence of Triton X100, respectively. L, G, NS, N, and M, represent the VSV viral proteins.

observed with both carboxypeptidase and aminopeptidase (Fig. 3.1.1, lanes c and e). Triton X100 releases integral membrane proteins by disrupting membrane structures and forming mixed micelles with phospholipids and proteins (Helenius and Simons, 1972). These results indicate that the intact viral envelope protects the NH_2- and COOH-termini of G protein from proteolytic attack. This can be interpreted to mean that the membrane bilayer confers a structural conformation on G protein such that the termini are unavailable for proteolytic attack or, alternatively, the termini are protected by virtue of being buried in the membrane or by being in close proximity to the membrane surface. Both these interpretations are valid since Triton X100 may expose susceptible sites by altering tertiary conformation or expose buried polypeptide regions by disrupting the lipid bilayer.

The molecular weight reduction of G protein in the presence of detergent was found to be 2-3000D in the case of carboxypeptidase digestion and approximately 3500D in the case of digestion with aminopeptidase. The nucleocapsid protein N is completely resistant to proteolysis either in the presence or absence of detergent. This may be due to the face that N is tightly complexed with nucleic acid in the RNP core and thus highly resistant to proteolytic attack (Wagner, 1975). In the presence of detergent, the viral proteins L, NS, and M are completely degraded with either carboxypeptidase or aminopeptidase (Fig. 1, c and e). The molecular weight of G could not be furthe reduced by increasing the enzyme concentration

(Fig. 3.1.2). The reason G is not completely hydrolyzed may be due to the primary structure of the protein or due to detergent binding to selective regions on the molecule. It is known that Triton X100 and other nonionic detergents bind preferentially to hydrophobic regions in proteins (Helenius and Simons, 1972). If the detergent is binding to a hydrophobic region near the COOH terminus, the carboxypeptidase enzyme may not be able to continue beyond that point. A molecular weight reduction of 2-3000D agrees with the finding that 3000D at the COOH terminus of G is exposed on the outside of microsommal vesicles when G protein is synthesized and inserted into membranes in vitro (Toneguzzo and Ghosh, 1978; Rothman and Lodish, 1977; Katz and Lodish, 1979). This reasoning could also be applied to the limited reduction in size observed with aminopeptidase M. In addition, the primary structure of G or the presence of modified amino acids could have an effect on the extent of the degradation.

Since the digestion experiments with the exopeptidases provide no knowledge concerning which regions of the G protein interact with the membrane, it was decided to undertake a more detailed analysis of G protein topology by treating intact virions with a variety of proteolytic enzymes.

That the G protein is present on the external surface of the viral membrane and forms the peplomers or spikes observed by electron microscopy of negatively stained viral preparations has been shown by a variety of techniques (Cartwright et al., 1969, 1971; McSharry,

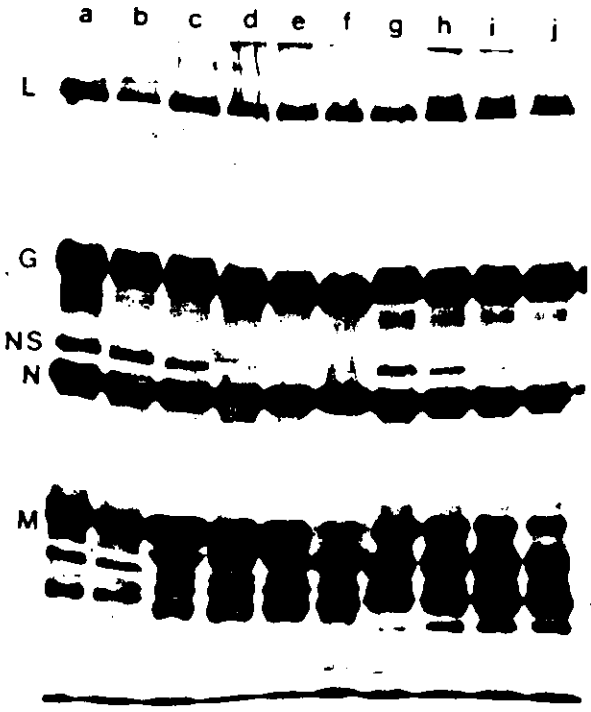


Figure 3.2.2: Polyacrylamide gel analysis of VSV treated with different amounts of carboxypeptidase and aminopeptidase.

[^{35}S]-Methionine labeled VSV (80 ug protein, 80,000 cpm) was treated with different amounts of carboxypeptidase or aminopeptidase and analyzed on a 10% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV marker; lane b, VSV incubated with 1% Triton X100; lane c, VSV incubated with 50 ug of carboxypeptidase in the absence of Triton X100; lanes d, e, and f, VSV treated with 25, 50, and 100 ug of carboxypeptidase in the presence of 1% Triton X100, respectively; lane g, VSV treated with 50 ug of aminopeptidase in the absence of Triton X100; lanes h, i, and j, VSV treated with 25, 50, and 100 ug of aminopeptidase in the presence of 1% Triton X100, respectively.

1971; Feuer et al., 1978). Surface labeling studies utilizing membrane impermeable reagents have shown that G is essentially the only component which becomes labeled in intact virions (Egar et al., 1975; McSharry, 1977). More convincingly, treatment of intact virions under the proper conditions with a variety of proteases such as trypsin, pronase, thermolysin, chymotrypsin, or bromelain results in the selective degradation of G protein while leaving the internal proteins intact (Mudd, 1974; Schloemer and Wagner, 1975a; Bishop et al., 1975).

The possibility of G protein being an integral membrane protein was suggested by the fact that the removal of G protein from virus requires the presence of a detergent (Kelly et al., 1972). Its ampipathic nature has been demonstrated from studies showing that delipidated, detergent free G protein forms rosette like structures (Petri and Wagner, 1979). In addition, a small molecular weight peptide remained associated with virions following proteolytic treatment (Mudd, 1974; Schloemer and Wagner, 1975a). This small peptide was presumed to represent the membrane interacting region of G since degradation of the internal viral proteins was not observed. In addition, this peptide was shown to be rich in hydrophobic amino acids. Evidence was provided that this peptide was derived from the amino terminus of G (Schloemer and Wagner, 1975a). This was based on the erronous assignment of alanine as the NH_2-terminal amino acid of G. It has subsequently been demonstrated that lysine is the amino terminal amino acid residue of G (Irving et al., 1979; Rose and

Gallione, 1981). A membrane interacting domain of G located at the NH_2-terminus would be inconsistent with the orientation of G that is observed in the microsommal vesicles and one would thus have to postulate a "flip-flop" of G as it migrates to the plasma membrane. These studies were based on analysis of viral degradation products on polyacrylamide gels of poor resolving power and thus the homogeneity of the peptide remaining with the spikeless virions following proteolysis was not certain. This was evidenced from the fact that the size of the peptide remaining following enzymatic treatment exhibited a broad range of molecular weight and was dependent upon the enzyme used and the conditions of digestion.

In order to examine more precisely the orientation of G protein in the viral envelope and define the region or regions of the molecule that are directly interacting with the lipid moiety of the virus, the following experimental approach was taken.

In preliminary experiments, various proteases were tested for their ability to hydrolyze the G protein while leaving the internal viral proteins intact. These included trypsin, chymotrypsin, and thermolysin. It was found that in the case of chymotrypsin, a two fold excess of enzyme by weight was required to completely digest G (Fig. 3.1.3, lane c). On the other hand, an enzyme/substrate ratio of 1:5 for thermolysin was sufficient to remove G (Fig. 3.1.3, lane d). Thermolysin was chosen for further studies for a number of reasons. Its activity in low concentrations would minimize degradation of internal structural proteins, it has a broad substrate

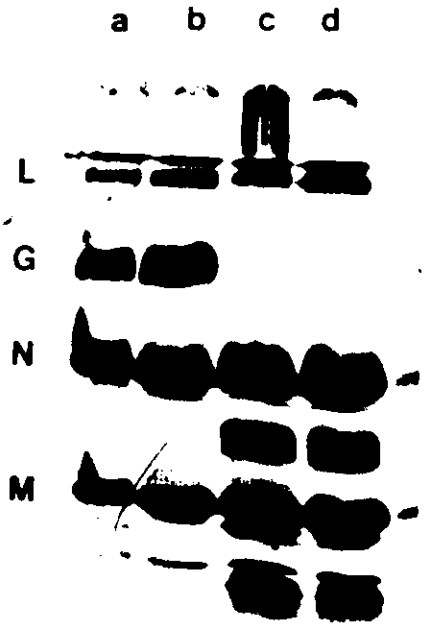


Figure 3.1.3: Polyacrylamide gel analysis of VSV treated with chymotrypsin and thermolysin.

100 ug of [^{35}S]-methionine labeled VSV (720,000 cpm) was digested with either 200 ug of chymotrypsin or 20 ug of thermolysin for 45 minutes at 37°C. The spikeless virions were isolated by centrifugation through a glycerol chushion as described in Methods and analyzed on a 17.5% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV marker; lane b, untreated VSV; lanes c and d, VSV treated with chymotrypsin and thermolysin, respectively.

specificity and is, therefore, more likely to cleave G protein closer to the membrane surface. Since the activity of thermolysin is dependent upon the presence of Zn^{2+}, it can be inactivated by divalant chelators such as EGTA.

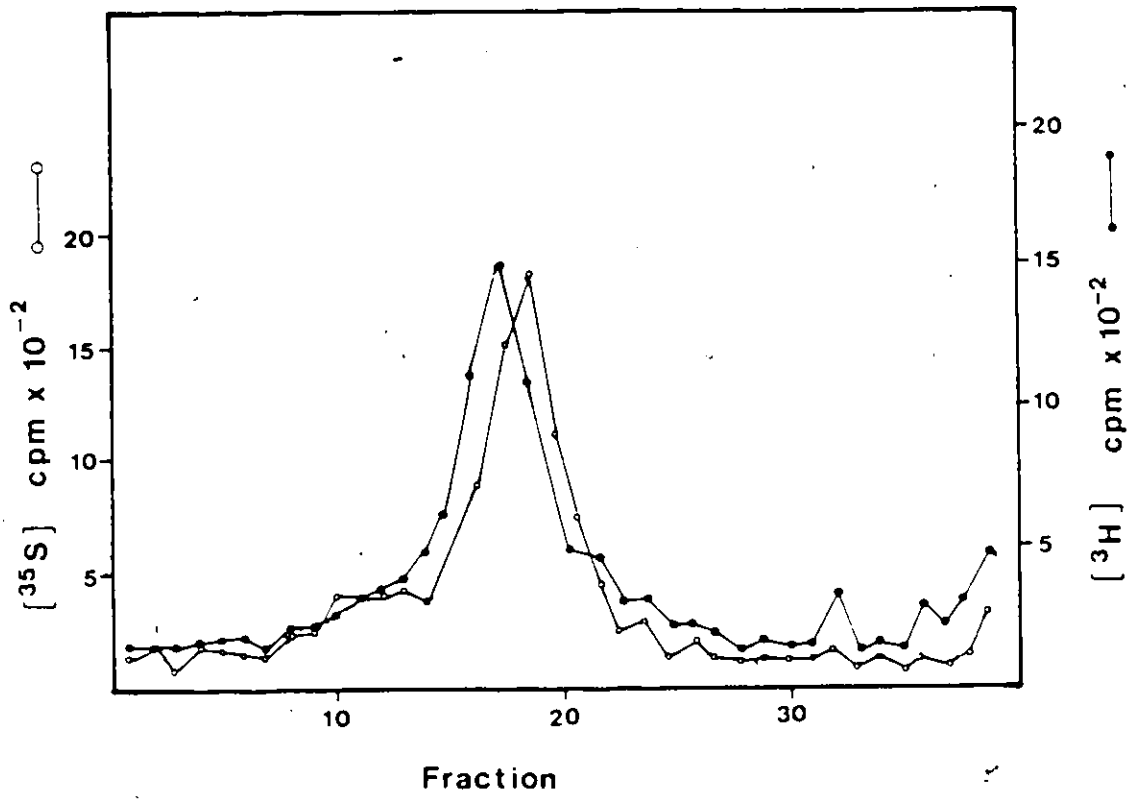
Analysis of thermolysin treated virions by sucrose density gradient centrifugation shows that the digested virus sediments at a buoyant density that is slightly higher than untreated virus (Fig. 3.1.4). This is consistent with the removal of the surface projections of the virus and has been observed by others (Mudd, 1974).

In order to examine the association of low molecular weight peptides with spikeless virions, 17.5% polyacrylamide gels containing SDS were used to obtain better resolution. Spikeless virions were purified from protease degradation products by centrifugation through a 50% glycerol or 50% sucrose cushion. Nonspecific degradation of viral proteins was minimized by the addition of EGTA and the protease inhibitors PMSF and trasylol to the reaction mixture as well as to the glycerol cushion following incubation. Nonspecific electrostatic association of degraded peptides with the spikeless virions was prevented by the inclusion of 0.1M NaCl to the 50% glycerol pad. PMSF was also routinely added to the sample buffer that was used to solubilize the viral pellet prior to electrophoresis. Generally, 50–60% of the starting material was recovered in the viral pellet following protease treatment and centrifugation.

Figure 3.1.5 shows the electrophoretic pattern obtained from

Figure 3.1.4: Sucrose density gradient profile of spikeless virions.

[^{35}S]-Methionine labelled VSV was digested with thermolysin and the spikeless virions were isolated by centrifugation through a glycerol cushion. The pellet was solubilized in 100 ul of NT buffer (10 mM Tris-HCl (pH 7.6), 100 mM NaCl). A sample of the spikeless virus, containing 25,000 cpm was mixed with 60,000 cpm of [^{3}H]-methionine labeled VSV and overlaid on a 12 ml 10-30% sucrose density gradient and centrifuged for 1 hour at 22,000 cpm in an SB283 rotor. Fractions were collected from the bottom of the tube and assayed for radioactivity in a scintillation counter preset for ^{3}H/^{14}C discrimination. (•) ^{3}H radioactivity; (o) ^{35}S radioactivity.


[35S] cpm x 10-2
[3H] cpm x 10-2
20
15
10
5
20
15
10
5
10
20
30
Fraction

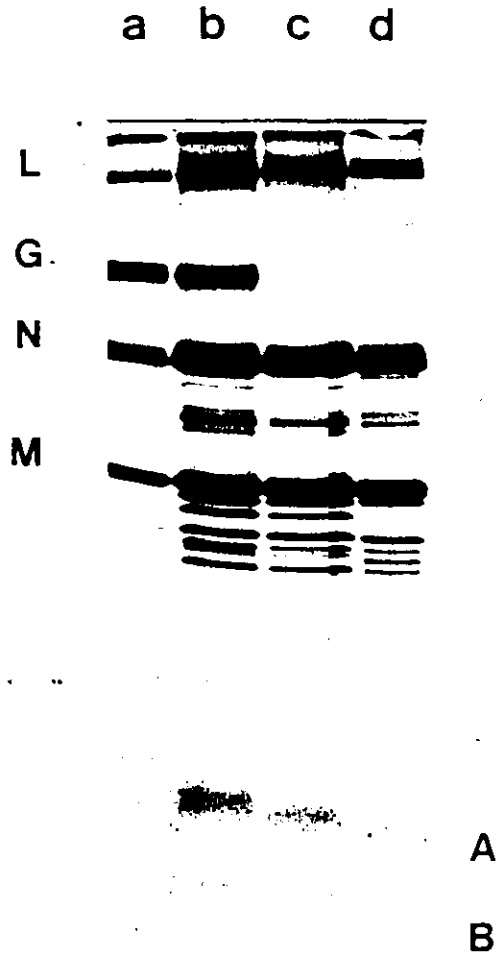


Figure 3.1.5: Polyacrylamide gel analysis of proteins remaining associated with VSV following proteolytic digestion.

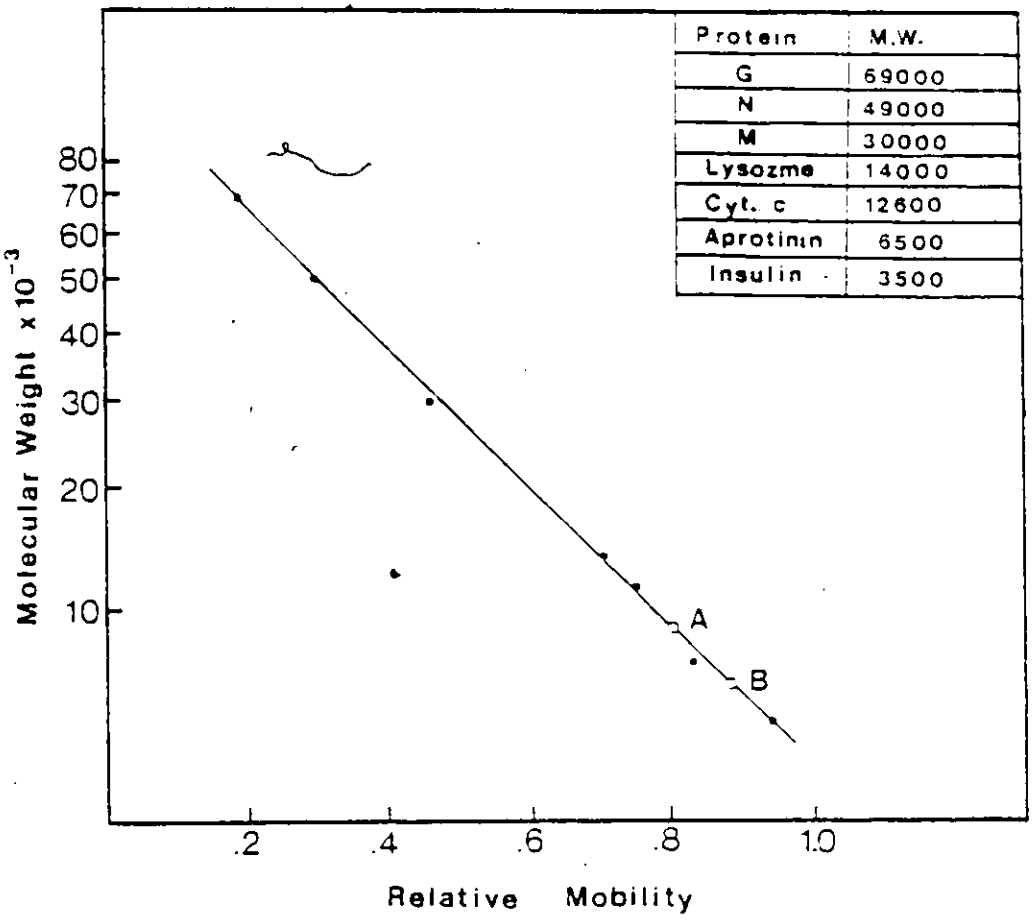
[^{35}S]-Methionine labeled VSV was digested with protease and the spikeless virions were purified and analyzed on a 17.5% polyacrylamide gel. Lane a, untreated VSV; lane b, VSV treated with TPCK-trypsin (viral protein/enzyme ratio of 4:1, w/w); Lane c, VSV treated with TPCK-trypsin (viral protein/enzyme ratio of 2:1, w/w); lane d, VSV treated with thermolysin (viral protein/enzyme ratio of 4:1, w/w). A and B indicate low molecular weight peptides which remain associated with VSV following proteolytic digestion.

VSV treated with thermolysin or trypsin. As can be seen, this treatment results in the degradation of G protein but not the other viral proteins L, NS, N, and M. However, two new low molecular weight peptides were observed in the spikeless virus preparations. The appearance of these low molecular weight peptides was dependent on prior treatment of VSV with proteolytic enzymes. These protease generated peptides are designated Fragments A and B, corresponding the slower and faster migrating species, respectively. The molecular weight of these fragments were estimated by comparing their mobility against appropriate molecular weight standards (Fig. 3.1.6). In the case of the thermolysin digestion, Fragments A and B were estimated to have apparent molecular weights of 9000D and 6000D, respectively, while in the case of trypsin, digestion, they were found to be 9000D and 7500D. In subsequent experiments, the molecular weight of Fragment A was found to be slightly higher in the case of thermolysin treatment as opposed to trypsin treatment.

Fragments A and B are end products of exhaustive proteolytic digestion of VSV since their molecular weights could not be further reduced by increasing enzyme concentration or time of incubation. The presence of 0.1 M NaCl in the gradient would tend to rule out the possibility that these represent degradation products co-sedimenting with the spikeless virions. Since G is the only protein completely degraded by proteolytic treatment, the conclusion is that one or both of these fragments represent that portion of G protein that is directly interacting with the membrane and, thus, protected from further

Figure 3.1.6: Molecular weight determination of peptide fragments remaining associated with VSV following proteolytic digestion.

[^{35}S]-Methionine labeled VSV was digested with thermolysin and the spikeless particles were run on a 17.5% polyacrylamide gel in parallel with reference proteins which included [^{35}S]-methionine labeled VSV, lysozyme, cytochrome c, aprotinin, and insulin. Nonradioactive proteins were detected by staining with Coomassie Blue while the radioactive proteins were detected by autoradiography. The apparent molecular weights of Fragments A and B, derived from thermolysin digestion of VSV, were estimated according to Weber and Osborn (1969) and were found to correspond to approximately 9,000D and 6,000D, respectively.


Protein | M.W.
G | 69000
N | 49000
M | 30000
Lysozme | 14000
Cyt. c | 12600
Aprotinin | 6500
Insulin | 3500
Molecular Weight x10^-3
80
70
60
50
40
30
20
10
A
B
.2
.4
.6
.8
1.0
Relative Mobility

proteolytic attack by virtue of being embedded in the viral envelope.

G protein can be selectively extracted from virions after solubilization with nonionic detergents under conditions of low ionic concentration (Schloemer and Wagner, 1975a). In order to determine whether the two low molecular weight fragments remaining associated with spikeless virions were derived from G protein, and hence associated with the membrane, spikeless virions were extracted with 2% Triton X100 under low ionic concentration as described in Methods. As can be seen in Figure 3.1.7, lane d, both Fragments A and B were extracted from spikeless virions under conditions where only G protein is extracted from intact virus. Thus, both Fragments A and B appear to be membrane associated.

As demonstrated in Figure 3.1.5, proteolytic treatment of VSV also results in the appearance of new protein species migrating between N and M protein and below M protein. These bands were consistently observed following proteolysis of VSV, however, they were not extracted with detergent (Fig. 3.1.7, lane d), suggesting that these protease generated bands are not partial cleavage products of G protein that remain embedded in the membrane. They do not represent degradation products adhering to the membrane since they too would be solubilized with the detergent. In all likelihood, they represent digestion products of internal structural proteins. This could arise from a population of leaky virions present in the purified virus preparation and their appearance may be compounded by the fact that a high concentration of enzyme was needed to completely

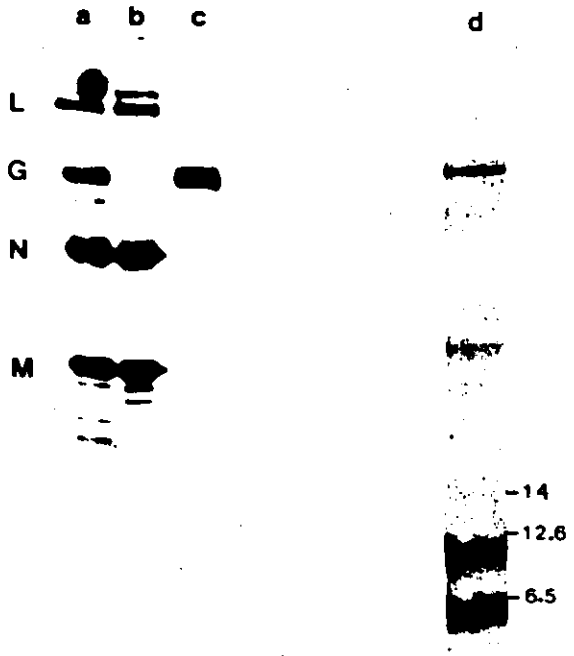


Figure 3.1.7: Isolation of membrane associated fragments by detergent extraction.

[^{35}S]-Methionine labeled VSV or spikeless virions was incubated in 10 mM Hepes (pH 7.6) containing 2% Triton X100 at room temperature for 1 hour. Nucleocapsids were removed by centrifugation and the supernatant was lyophylized, extracted with acetone, solubilized in electrophoresis sample buffer and analyzed on a 17.5% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV marker; lane b, VSV pellet following detergent solubilization; lane c, VSV supernatant following detergent solubilization; lane d, supernatant from detergent solubilization of thermolysin treated VSV. Lane d was from a separate gel which included lysozyme (14,000D), cytochrome c (12,600D) and aprotinin (6,500D) as molecular weight markers.

hydrolyze G protein. Alternatively, there may be some degradation occurring during solubilization of the spikeless virions and subsequent electrophoresis. This may be a possibility since it has been reported that purified VSV has some associated proteolytic activity (Holland et al., 1971) and that some proteases remain active in SDS containing gels (Cleveland et al., 1978). These possibilities were minimized by carrying out all manipulations following proteolysis in the cold, sedimenting the spikeless virions soon after treatment as possible, including BSA in the reaction mixture following enzyme digestion, and including protease inhibitors in all buffers. The anamolous bands migrating between N and M remained associated with the spikeless virions even after repurification on a sucrose density gradient followed by pelleting through sucrose although their intensity was reduced (Fig. 3.1.8). It was found that their intensity could be further reduced by including 10 mM DTT in the glycerol cushion used to purify the spikeless virions (Fig. 3.1.9). DTT may act by disrupting protein interactions in leaky virions to an extent that they no longer co-sediment with the spikeless virions.

In order to examine the rate of appearance of Fragments A and B, VSV labeled with methionine, leucine, valine, and lysine was digested with increasing amounts of thermolysin and run on a 10-20% polyacrylamide gel. The regions of the gel corresponding to Fragments A and B and G and N proteins were excised from the gel, eluted with 0.05% SDS, and assayed for radioactivity. The results

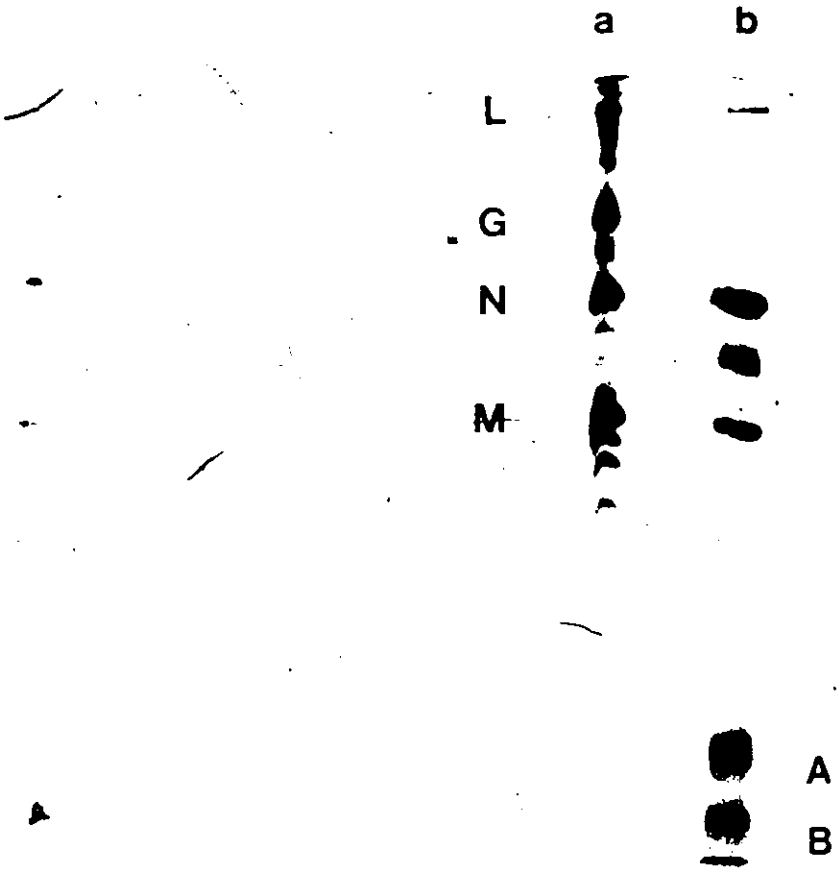


Figure 3.1.8: Isolation of spikeless virions by sucrose gradient centrifugation.

[^{35}S]-Methionine labeled VSV, digested with thermolysin, was pelleted through 50% glycerol, suspended in NT buffer, and overlaid on a 10-30% sucrose density gradient. After centrifugation at 22,000 rpm for 1 hour, fractions were collected from the bottom of the tube and aliquots assayed for radioactivity. The peak fraction was pooled, diluted with NT buffer, sedimented at 48,000 rpm for 1 hour, and the pellet was analyzed on a 17.5% polyacrylamide gel. Lane a, untreated VSV; lane b, VSV digested with thermolysin.

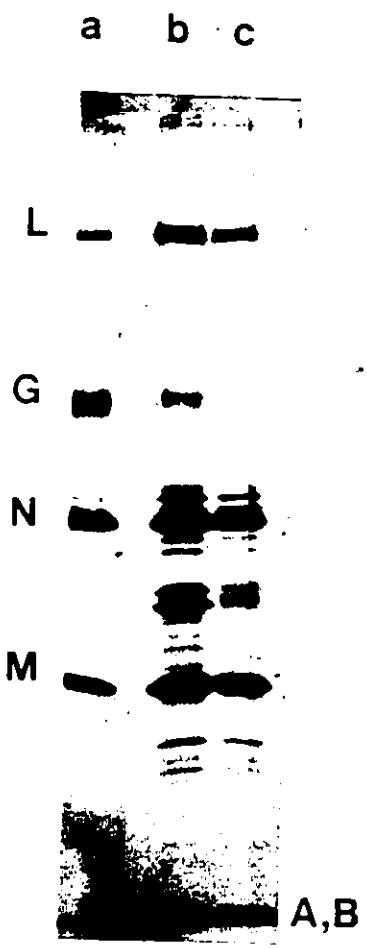


Figure 3.1.9: Effect of dithiothreitol on the association of proteins with spikeless virions.

[^{35}S]-methionine labelled VSV was digested with thermolysin. The reaction was divided into two equal portions and pelleted through a 50% glycerol cushion in the absence or presence of 10 mM DTT. The resulting spikeless viral pellets were solubilized in sample buffer and analyzed on a 15% polyacrylamide gel. Lane a, untreated VSV; lanes b and c, spikeless virions purified in the absence or presence of 10 mM DTT, respectively.

shown in Figure 3.1.10 indicate that the rate of appearance of A and B are similar and coincide with the rate of disappearance of G protein. Very little loss of N protein occurred. Since both A and B show a similar rate of appearance it appears that they are distinct peptides and that Fragment A is not further digested to a peptide(s) that migrates with B.

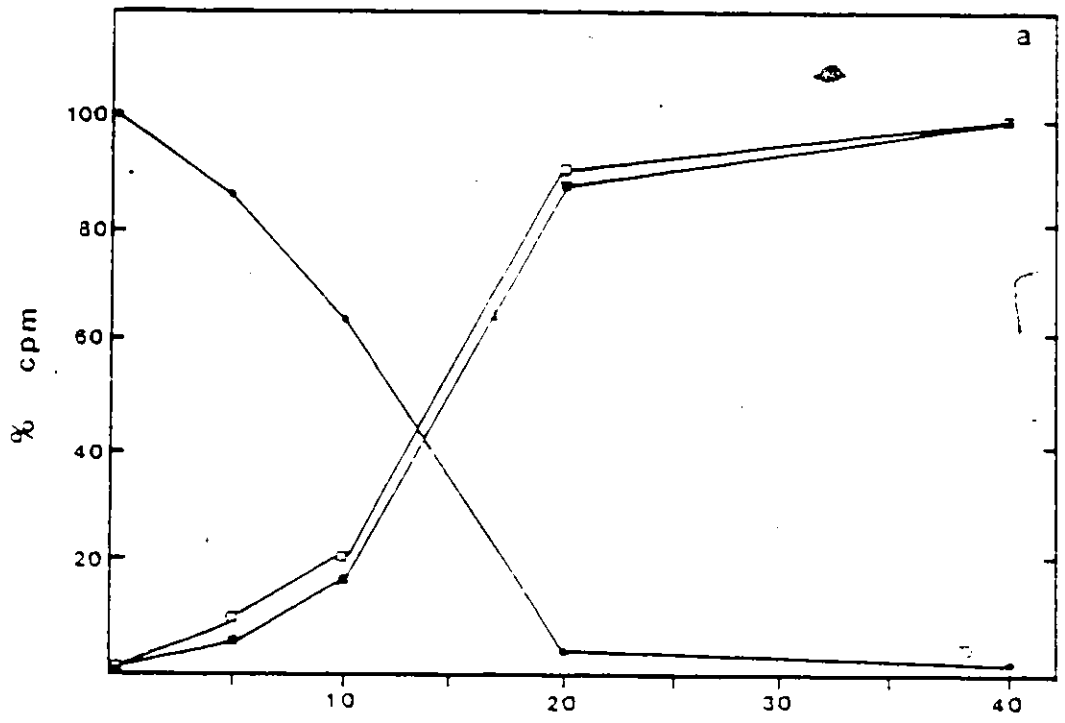
In order to quantitate what percentage A and B make of the total viral protein, [^{35}S]-methionine labeled VSV was exhaustively digested with thermolysin and the viral proteins were eluted from the gel and counted. As demonstrated in Table 1, A and B represent 2.7 and 1.7% of the methionine radioactivity remaining associated with the spikeless virions.

To determine the origin of the minor species appearing between N and M proteins and below M protein following proteolytic digestion of VSV, a one dimensional peptide mapping procedure was used (Cleveland et al., 1978). Staphylococcus aureus V8 protease was used to generate peptide fragments for the comparative studies. In a direct comparison with G protein, it was found that the band migrating between N and M proteins had no homology with G protein (Fig. 3.1.11). However, when this analysis was carried out with N protein, a clear homology was observed (Fig. 3.1.12). This was confirmed by a two dimensional mapping procedure whereby tryptic peptides were generated by digestion of proteins embedded in gel slices with trypsin and separating them on thin layer plates by electrophoresis in one direction and chromatography in the other.

Figure 3.1.10: Effect of increasing thermolysin concentration on the appearance of Fragments A and B.

Top: 100 ug of VSV labeled with [^{35}S]-methionine, [^{3}H]-lysine, [^{3}H]-leucine, and [^{3}H]-valine, was digested with increasing amounts of thermolysin and the spikeless virions were run on a 10-20% polyacrylamide gel. Regions of the gel corresponding to G protein and Fragments A and B were excised, eluted with 0.05% SDS, and the radioactivity was determined. The results were plotted as a percentage of radioactivity remaining after proteolytic treatment. (●) G protein; (□) Fragment A; (■) Fragment B.

Bottom: Autoradiogram of the region of the gel containing Fragments A and B which was used to determine the values shown in the top figure.


a
100
80
60
40
20
% cpm
10
20
30
40
µg Thermolysin

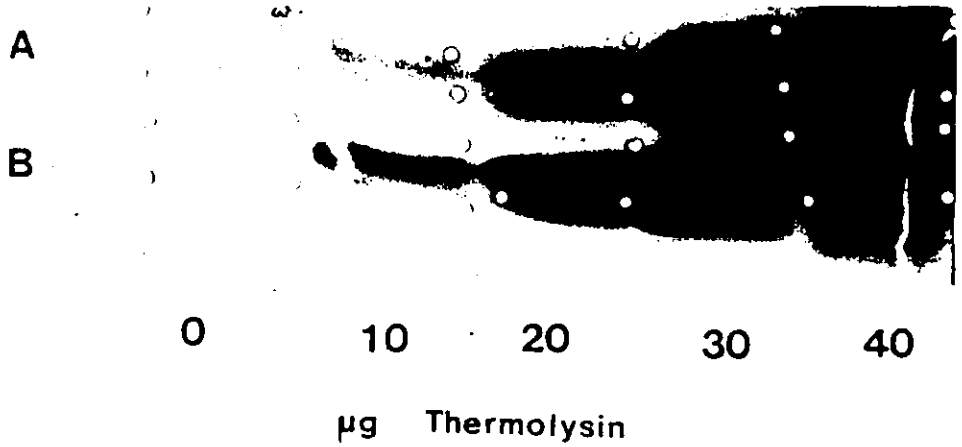

A
B
0
10
20
30
40
µg Thermolysin

Table I: Quantitation of proteins present in VSV before and after digestion with thermolysin.

Viral protein	% of total radioactivity	
	Untreated VSV	VSV + thermolysin
L	6.3	9.3
G	17.7	-
NS/N	40	39.8
M	36	40.1
Fragment A	-	2.7
Fragment B	-	1.7

Protein bands corresponding to [^{35}S]-methionine labeled VSV untreated or treated with thermolysin were excised from a gel, eluted, and the radioactivity was determined. The amount of radioactivity present in each protein species is expressed as a percentage of the total present in all protein species. NS and N proteins are counted as one species since these two proteins comigrate on a 17.5% polyacrylamide gel. Approximately 6% of the radioactivity present in the spikeless virus preparation was present in the minor species migrating between N and M protein and below M protein.

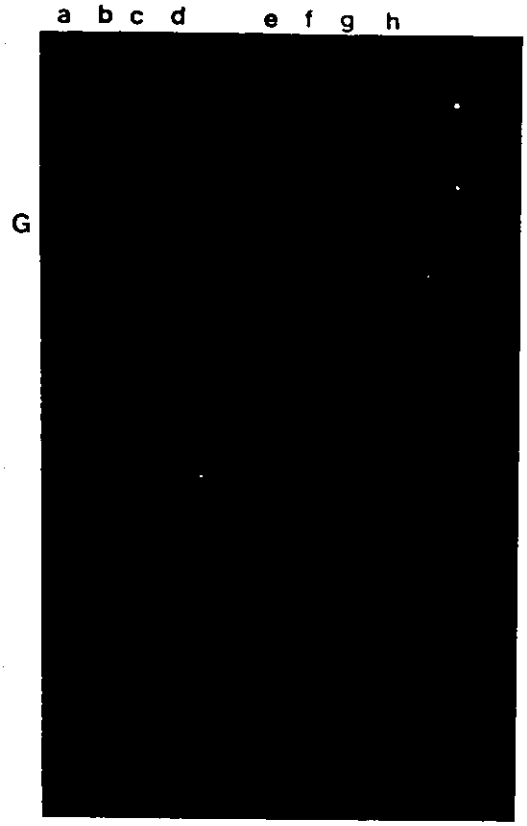


Figure 3.1.11: Comparison of G protein and the species migrating between N and M in spikeless virions by limited proteolysis.

[^{35}S]-Methionine labeled G protein and the band migrating between N and M following thermolysin digestion of VSV were analyzed by limited proteolysis with V8 protease as described in Methods. Lanes a, b, c, and d, G protein treated with 0, 2, 0.5, and 0.15 ug of V8 protease, respectively; lanes e, f, g, and h, band migrating between N and M treated with 0, 0.15, 0.5, and 2 ug of V8 protease, respectively.

The results of this analysis are shown in Figure 3.1.13. As can be seen, this protein shows little or no homology with G protein, however, there is a great deal of similarity with N protein. Thus, this fragment is likely derived from partial degradation of N. Minor bands migrating just below M protein were occasionally observed following proteolysis. A comparison of these bands by the Cleveland mapping technique indicates that they were derived from partial degradation of M protein (Fig. 3.1.12).

Thus, it appears that even under the most stringent conditions used, there is some degradation of internal structural proteins. This observed proteolysis of internal viral proteins was probably due to the presence of a population of virus which allowed the leakage of protease inside the virions. Use of a lower enzyme concentration or reducing the time of incubation left a large fraction of G protein undigested and Fragments A and B were not consistently produced.

Although it was clear that A and B were end products of extensive proteolytic digestion of VSV, it could not unequivocally be stated that they represent the membrane interacting regions of G since some degradation of internal proteins was occurring. If the orientation of G that is observed in the internal membrane structures of the infected cell is maintained in the viral envelope, then both A and B could be candidates for the membrane interacting domain. The membrane anchoring region of G should have a minimum molecular weight of 5000D since approximately 30 amino acids would be exposed on the

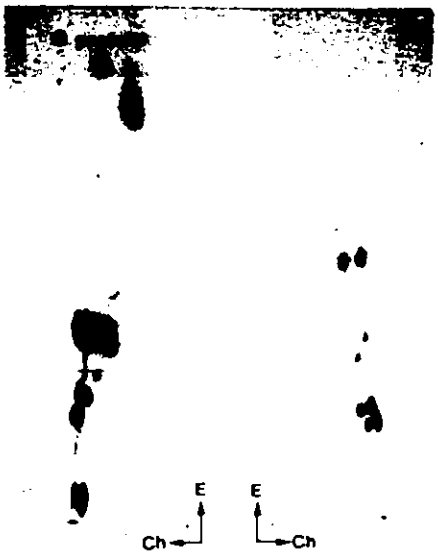
Figure 3.1.12: Analysis of proteolytic degradation products remaining associated with spikeless VSV by limited proteolysis.

[^{35}S]-Methionine labeled proteins were excised from polyacrylamide gels and analyzed by limited proteolysis with V8 protease as described in Methods. Lane a, N protein treated with 1 ug of enzyme; lanes b and c, band migrating between N amd M proteins in spikeless virus treated with 0.2 and 1 ug of enzyme, respectively; lane d, [^{35}S]-methionine labeled VSV marker; lane e, M protein treated with 1 ug of enzyme; lane f, protein species migrating below M protein in spikeless virus treated with 1 ug of enzyme. Lanes e and f were from a separate gel.


a
b
c
d
e
f
L
G
N
M

Figure 3.1.13: Two dimensional tryptic peptide analysis of the protein species migrating between N and M protein in spikeless VSV.

[^{35}S]-Methionine labeled proteins were excised from polyacrylamide gels, digested with trypsin, and the tryptic peptides were analyzed by TLC as described in Methods. For the direct comparison of tryptic peptides, samples were spotted on the same plate on opposite sides of the center line of the plate. Following electrophoresis, the plate was cut along the center line and the two halves were chromatographed at right angles to the direction of electrophoresis. The figure shows the tryptic peptides derived from (A) G protein, (B) N protein, and (C and D) the protein species migrating between N and M protein present in VSV treated with thermolysin. The tryptic peptides were not oxidized in this analysis. The direction of electrophoresis (E) and chromatography (Ch) is indicated.



B D

N protein

DOCUMENT SLANTED
Document incliné

internal surface of the membrane (Toneguzzo and Ghosh, 1978; Rothman and Lodish, 1977; Katz and Lodish, 1979) and at least 20 amino acids in an helical conformation would be required to span the membrane (Sergest et al., 1971).

The origin of the small molecular weight, protease resistant fragments were characterized as follows. VSV was digested with a mixture of carboxypeptidases A and B in the presence of Triton X100 such that G protein was only hydrolyzed to a limited extent. This truncated form of G was then used to compare with the peptide patterns derived from Fragments A and B. The rationale behind this was that if A or B was derived from the carboxy terminus of G, then they should have peptides in common with intact G protein but that are absent from the truncated form of G.

As seen in Figure 3.1.14, limited digestion of [^{35}S]-methionine labeled G protein with V8 protease generates six distinct fragments. One of these fragments is missing in G protein treated with carboxypeptidase A and B prior to V8 protease digestion (Fig. 3.1.14, lane b). Digestion of [^{35}S]-methionine labeled Fragment A with V8 protease results in the generation of a band corresponding to the one missing from carboxypeptidase treated G. Treatment of [^{35}S]-methionine labeled Fragment B with V8 on the other hand does not result in the generation of a band corresponding to the one missing from carboxypeptidase treated G.

The tryptic peptides generated from Fragments A and B were compared with the tryptic peptides generated from G protein

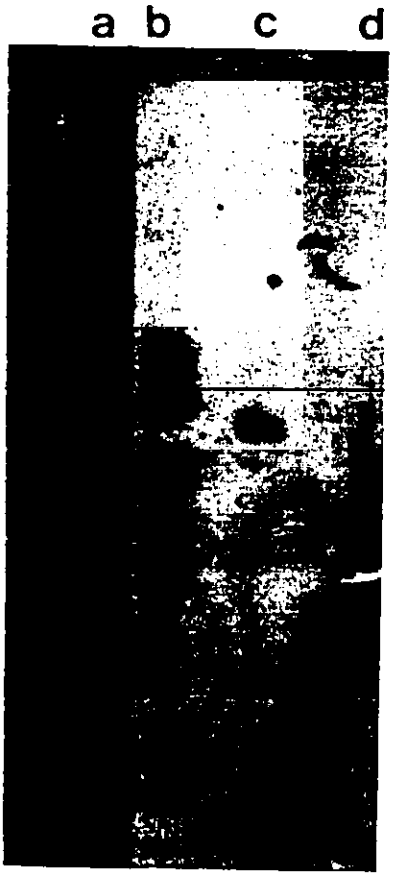


Figure 3.1.14: Comparison of carboxypeptidase treated G protein and Fragments A and B by limited proteolysis.

[^{35}S]-Methionine labeled proteins were excised from polyacrylamide gels and analyzed by one dimensional peptide mapping on a 20% polyacrylamide gel after limited digestion with V8 protease. Lane a, G protein; lane b, carboxypeptidase treated G protein; lane c, Fragment A; lane d, Fragment B.

untreated with carboxypeptidase by using ion exchange chromatography. As can be seen in Figure 3.1.15, analysis of the tryptic peptides derived from [^{35}S]-methionine labeled G on an Aminex A5 ion exchange resin separated 9-10 distinct peaks of radioactivity. The peak eluting in the first few fractions probably represents undegraded or heterogenous material not binding to the column since its intensity varied. The minor peaks observed result from sulfoxide derivatives of methionine residues since the peptides were not oxidized prior to chromatography.

To examine which tryptic peptides correspond to the carboxy terminal end of the molecule, [^{35}S]-methionine labeled G was truncated by approximately 3000D by carboxypeptidase treatment and co-digested with intact [^{3}H]-methionine labeled G protein. As shown in Figure 3.1.16, carboxypeptidase treatment of G protein results in the loss of two peptides. These are numbered 2 and 5 and are indicated by the arrows in the figure. These are the same two peptides that are present in the cytoplasmic domain of G when G is translated and inserted into microsommal vesicles in vitro (H. P. Ghosh, personal communication; Irving and Ghosh, 1982; Toneguzzo and Ghosh, 1978). When tryptic peptides of [^{35}S]-methionine labeled Fragment A were compared with tryptic peptides from G, it was found that Fragment A contained both the peptides that were removed in carboxypeptidase treated G (Fig. 3.1.16b). Thus, Fragment A is derived from a region at or near the COOH-terminus of G. Since A is obtained only after exhaustive

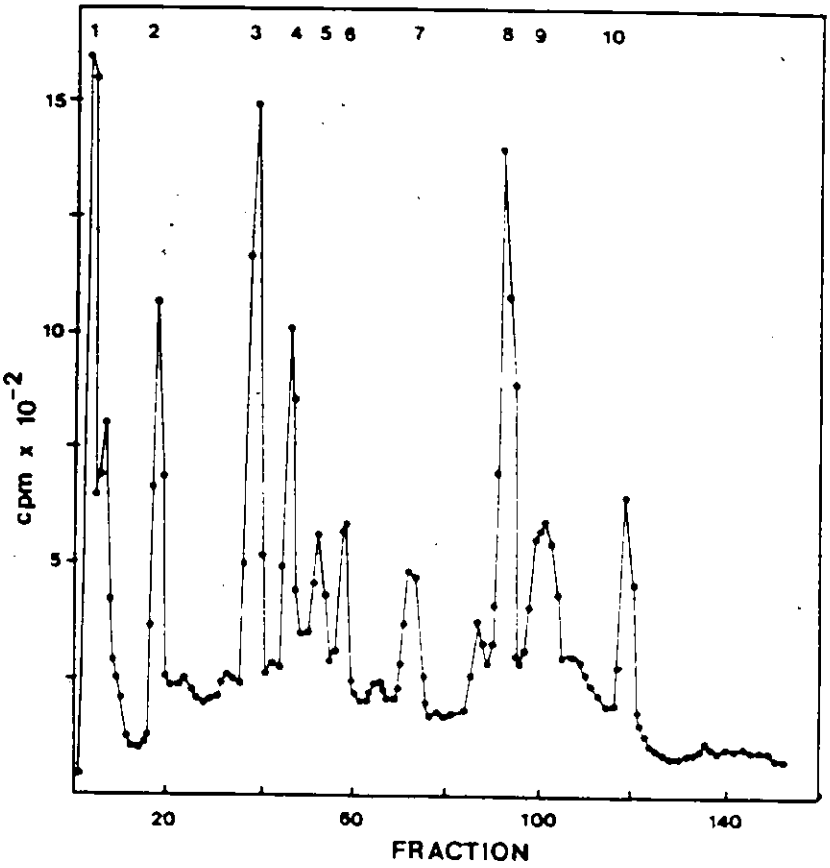
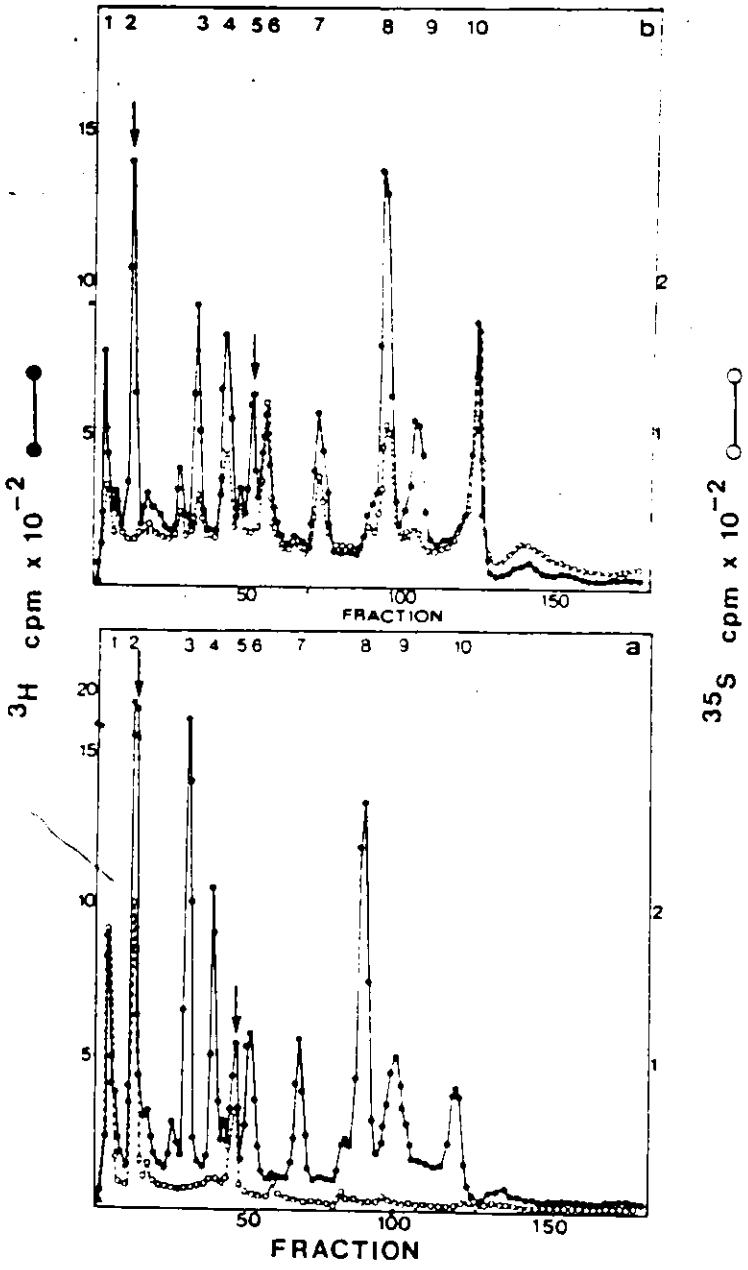


Figure 3.1.15: Tryptic peptide analysis of [^{35}S]-methionine labeled G protein.

[^{35}S]-Methionine labeled G protein was isolated from a 10% polyacrylamide gel, digested with trypsin, and analyzed by ion exchange chromatography as described in Methods. Approximately 50,000 cpm was applied to the column. The major methionine containing tryptic peptides are indicated by the numbering system at the top of the figure.

Figure 3.1.16: Tryptic peptide analysis of [^{35}S]-methionine labeled Fragment A and carboxypeptidase treated G protein.

The figure shows the chromatographic profile obtained from (a) codigestion of [^{35}S]-methionine labeled Fragment A (10,000 cpm) and [^{3}H]-methionine labeled G protein (90,000 cpm); (b) codigestion of [^{3}H]-methionine labeled G protein and [^{35}S]-methionine labeled G protein which was partially digested with a mixture of carboxypeptidases A and B. Carboxypeptidase digestion in this case reduced the apparent molecular weight of G protein by approximately 3,000D. The arrows designate the tryptic peptides that are present in Fragment A but which are absent in carboxypeptidase treated G protein. (•) [^{3}H] radioactivity; (o) [^{35}S] radioactivity.


1 2 3 4 5 6 7 8 9 10
b
a
^3H cpm x 10^{-2}
^{35}S cpm x 10^{-2}
FRACTION
50
100
150
5
10
15
20
1
2

proteolytic digestion of VSV and remains associated with the spikeless virions during subsequent purifications, it represents that region of G protein that is embedded in the viral envelope.

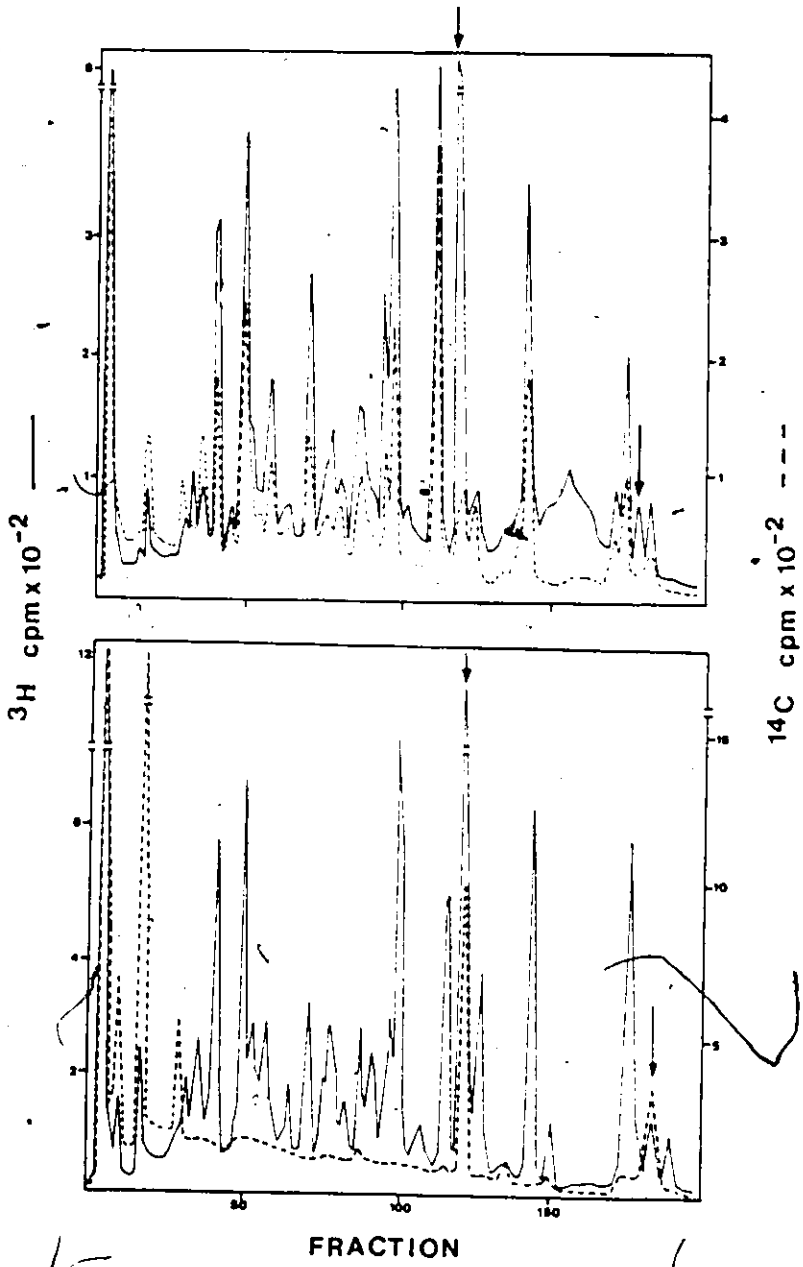
The COOH terminal localization of Fragment A was further confirmed by repeating the experiment with Fragment A or carboxypeptidase treated G that was metabolically labeled with [^{14}C]-lysine (Fig. 3.1.17). In this case, carboxypeptidase treatment of G results in the loss of 1-2 peptides (indicated by the arrows) that are present in trypsinized Fragment A.

That Fragment A was derived from a protected region near the amino terminus of G was ruled out by comparing tryptic maps from [^{14}C]-lysine labeled G protein that had been truncated with aminopeptidase M such that the molecular weight was reduced by approximately 1500D. As indicated in Figure 3.1.18, aminopeptidase M treatment results in the loss of one peak. This peak was distinct from the peaks removed by carboxypeptidase digestion and from the tryptic peptides derived from Fragment A.

The data presented in Figure 3.1.9 would tend to indicate that Fragment B was not derived from a degradation product of Fragment A since the kinetics of appearance of A and B were similar and they were both stable with increasing enzyme concentrations. Tryptic peptide analysis of [^{35}S]-methionine labeled B revealed that most of the radioactivity was eluted with 1 M pyridine (Fig. 3.1.19a). This indicated that a large fraction of B was tightly bound to the column suggesting that B contains a high

Figure 3.1.17: Tryptic peptide analysis of [^{14}C]-lysine labeled Fragment A and carboxypeptidase treated G protein.

Panel a, codigestion of [^{14}C]-lysine labeled G protein partially digested with carboxypeptidase and [^{3}H]-lysine labeled G protein; Panel b, codigestion of [^{14}C]-lysine labeled Fragment A and [^{3}H]-lysine labeled G protein. Arrows indicate tryptic peptides present in Fragment A but absent in G protein treated with carboxypeptidase. (——) [^{14}C] radioactivity; (——) [^{3}H] radioactivity.


^3H cpm x 10^{-2}
^{14}C cpm x 10^{-2}
FRACTION

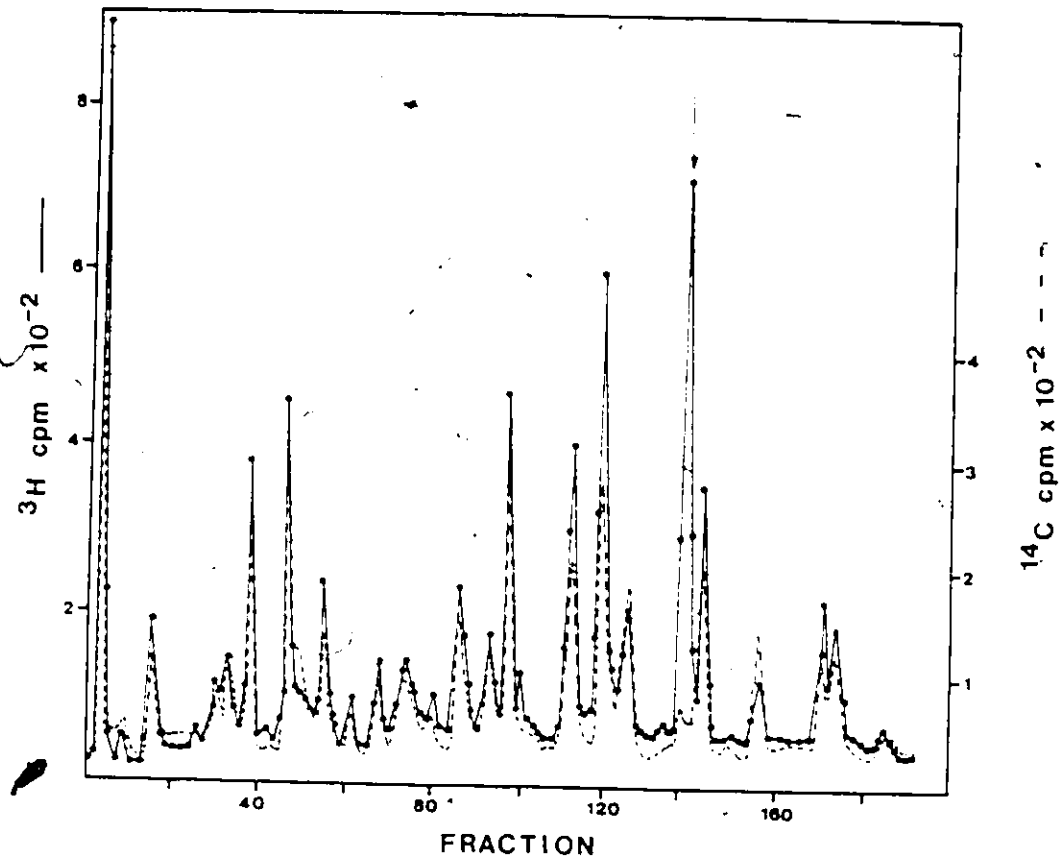


Figure 3.1.18: Tryptic peptide analysis of G protein treated with aminopeptidase M.

[^{14}C]-Lysine labeled VSV was treated with aminopeptidase M in the presence of Triton X100 as described in Methods. The truncated G protein (molecular weight reduction of approximately 1,500D) was isolated and codigested with [^{3}H]-lysine labeled G protein. The arrow indicates a peak which is present in G protein but is absent in aminopeptidase treated G protein.
(----) [^{14}C] radioactivity; (——) [^{3}H] radioactivity.

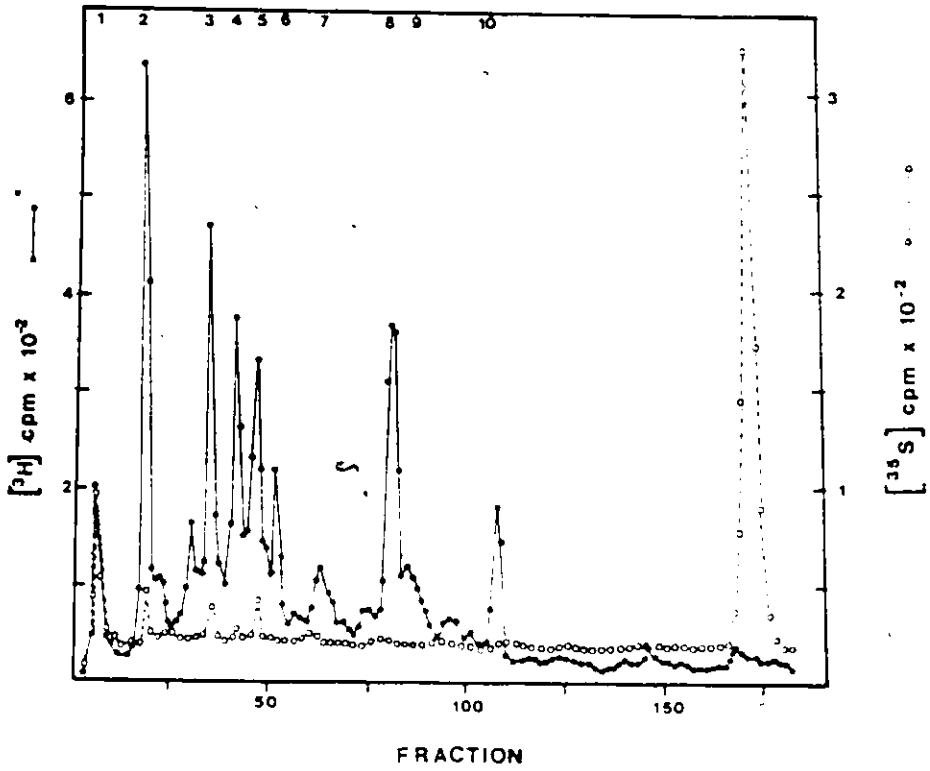
proportion of charged amino acids or that there is a considerable degree of aggregation resulting in retarded elution. This was not observed with Fragment A.

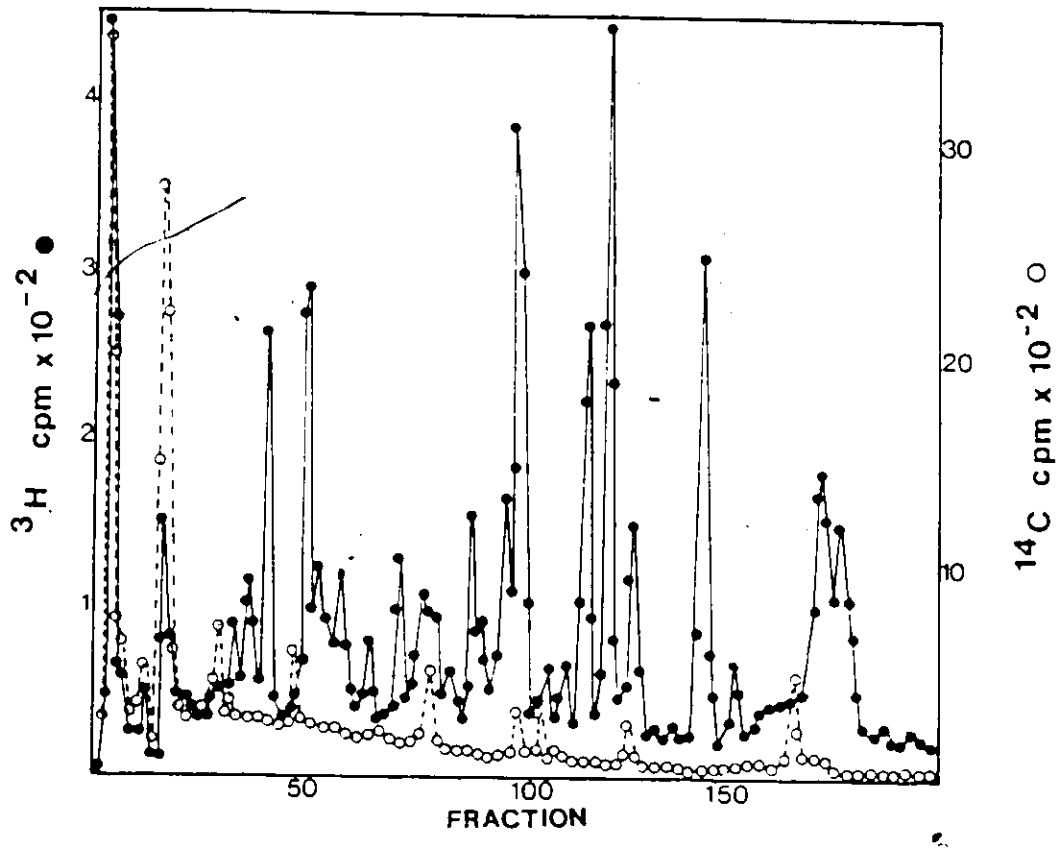
Analysis of [^{14}C]-lysine labeled Fragment B revealed one major peptide eluting at fraction 15 and six other minor peptides, none of which co-migrated with lysine labeled tryptic peptides from G (Fig. 3.1.19b). Thus, Fragment B does not appear to be derived from G protein. Fragment A shares a common lysine labeled peptide with B eluting between fractions 15-20 (compare Fig. 3.1.19b and 3.1.17). It is possible that the material eluting at this position represents aggregated or undigested material since the radioactivity in this peak is disporportionate to the other lysine labeled peaks. Tryptic digestion of lysine labeled proteins should generate peptides with similar levels or radioactivity. This was not observed with either Fragments A, B, or G protein labeled with lysine. This discrepancy in the amount of radioactivity in each peak has been observed by other workers examining different lysine labeled proteins (Kinney and Trent, 1982). It is possible that two or more tryptic peptides have the same elution profile. Although limited protease digestion cannot be ruled out completely, all digestions were carried out with a vast excess of trypsin relative to substrate. Selective losses of various peptides cannot be completely discounted but steps to minimize such losses, such as siliconizing all glassware, were routinely undertaken. It is also possible that certain peptides may be

Figure 3.1.19: Tryptic peptide analysis of Fragment B.

Top: Codigestion of [^{35}S]-methionine labeled Fragment B and [^{3}H]-methionine labeled G protein. (•) [^{3}H] radioactivity; (o) [^{35}S] radioactivity.

Bottom: Codigestion of [^{14}C]-lysine labeled Fragment B and [^{3}H]-lysine labeled G protein. (•) [^{3}H] radioactivity; (o) [^{14}C] radioactivity.


1 2 3 4 5 6 7 8 9 10
[3H] cpm x 10-3
[35S] cpm x 10-2
FRACTION


3H cpm x 10-2 ●
14C cpm x 10-2 ○
FRACTION

preferentially retained in the gel matrix during and after trypsinization.

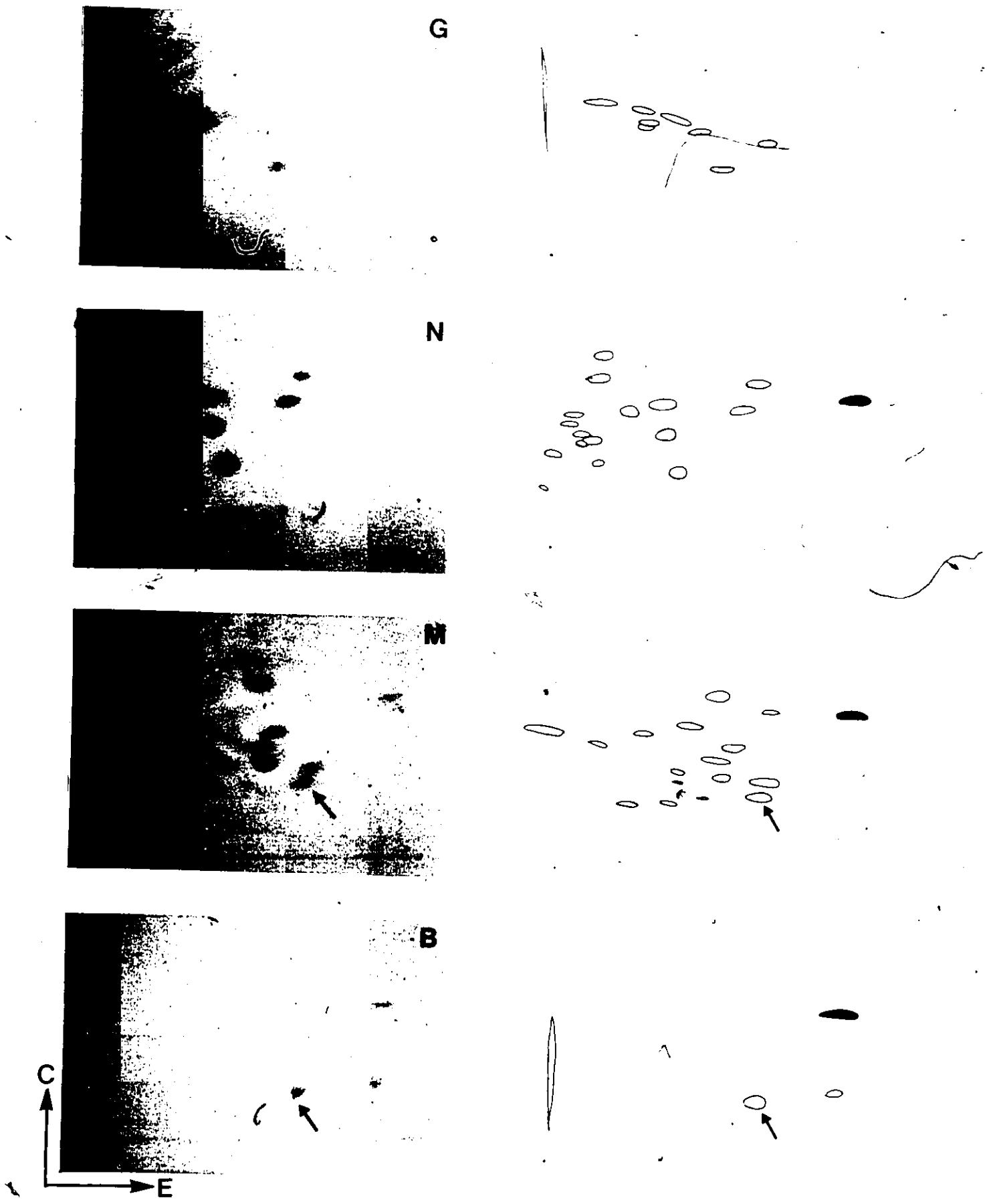
In order to get a clearer idea of the origin of Fragment B, two dimensional tryptic maps of [^{35}S]-methionine labeled Fragment B and viral proteins were prepared.

Trypsinization of [^{35}S]-methionine labeled Fragment B resulted in the appearance of three distinct peptides (Fig. 3.1.20). None of these peptides showed any homology with G protein, however, two of the spots corresponded to peptides generated from M protein. When tryptic maps were compared with N protein, one of the spots present in Fragment B that showed homology with M protein was also present in N. Thus, it appears that Fragment B is derived from degradation products of both N and M. M protein would be a likely candidate for proteolytic attack since it is in close proximity to the inner surface of the viral envelope. Thus, in the fraction of virions in which proteases could attack internal proteins, M would be the most susceptible. It appears that B Fragment is generated under conditions found necessary to completely hydrolyze the external portion of G protein. Under conditions where G protein was not completely removed, Fragment B was not consistently observed.

Although Fragment B is a degradation product or possibly an artifact of the isolation and preparation procedures, its generation must result from a specific cleavage since its size was always consistent through many experiments. Its size was also dependent upon the enzyme used to digest the virus. The fact that Fragment B

Figure 3.1.20: Two dimensional tryptic peptide analysis of G, N, M proteins, and Fragment B.

[^{35}S]-Methionine labeled proteins were excised from polyacrylamide gels and digested with trypsin. The tryptic peptides were oxidized with performic acid and analyzed by TLC as described in Methods. Panels a, b, c, and d, represent the tryptic peptides obtained from G, N, M proteins, and Fragment B, respectively. Approximately 20,000 cpm each of G, N, and M, and 5,000 cpm of Fragment B were spotted. The left part of the figure is a schematic representation of the peptide maps. The filled in spot indicates a common tryptic peptide which is present in N, M, and Fragment B. The arrow indicates a common peptide present in M protein and Fragment B. The direction of electrophoresis (E) and chromatography (C) is indicated.


G
N
M
B
C
E

could be extracted with Triton X100 under conditions where only G protein is normally liberated indicates that B is digested free from internal proteins since Fragments A and B were the only species found in the detergent supernatant (Fig. 3.1.7).

Previous workers investigating G protein topology have reported that only one low molecular weight peptide is present after proteolytic digestion of VSV (Mudd, 1974; Schloemer and Wagner, 1975a). However, in both cases this peptide exhibited a very broad range of molecular weight and from its characterization it is likely that more than one peptide species was present in this region of the gel.

Recently, Rose et al., have reported results in agreement with those reported here (Rose et al., 1980). Using a similar digestion protocol, they observed only one low molecular weight fragment following enzymatic digestion of VSV. However, they also observed the characteristic bands migrating between N and M and below M protein. They assumed that these were partial cleavage products of G protein co-terminal with the smallest fragment. Though it is possible that a fragment corresponding to Fragment B was not generated under the conditions they used, it is also possible that a smaller molecular weight fragment was produced but that it electrophoresed off the gel.

The major conclusions from this chapter are that: 1) both the amino and carboxy termini of G protein are protected from proteolytic attack by the integrity of the viral envelope and 2) that

the G protein itself is anchored in the membrane by a domain at or near the COOH terminus. This is the same domain that initially stopped transfer of the protein through the RER. The implication of this finding is that the orientation of G protein that is assumed in the endoplasmic reticulum is maintained during its transport to the plasma membrane and to its ultimate destination in the viral envelope.

The fact that the amino terminus of G protein was protected from aminopeptidase attack but that no protected fragment corresponding to the amino terminus was detected following proteolytic digestion of VSV indicates that the amino terminus of G may be in close proximity to the membrane surface or that the viral envelope maintains G in a conformation such that the amino terminus is protected. Alternatively, the amino terminus of G could be buried in the membrane but the protected region may be too small to observe on the gel system used.

SECTION 3.2: PRIMARY STRUCTURAL STUDIES ON THE PROTEASE RESISTANT FRAGMENTS ASSOCIATED WITH VSV

The results obtained in Section 3.1 indicate that G protein exists in the viral envelope as an integral membrane protein with approximately 90% of its bulk exposed on the surface and anchored in the membrane by a region at or near the COOH terminal end of the molecule. Thus a carboxy terminal domain plays a role in maintaining protein asymmetry by directly interacting with the lipid bilayer. In terms of protein maturation, the same domain may act to interrupt the translocation process initiated by the signal sequence.

The results on the proteolysis of VSV do not provide any information on the properties of this domain that may endow it with these specialized functions. It has been demonstrated that glycophorin A, a red cell membrane glycoprotein, contains a 23 amino acid uninterrupted hydrophobic domain (Tomita and Marchesi, 1975). This region presumably represents the membrane spanning segment of the protein due to its highly lipophilic properties. It was, therefore, decided to attempt to obtain primary structural information on the carboxy terminal membrane anchoring region of G in order to get a more precise localization of this region with respect to the G molecule and with the hope that this information could be correlated to some functional characteristic. Partial amino acid sequencing was also carried out on Fragment B in order to determine its origin in a more definitive manner.

In attempting to sequence the membrane associated fragments of VSV, two considerations had to be taken into account; a) the small amount of material available, and b) intrinsic properties of the peptide such as hydrophobicity. This latter point is important since peptide losses may be so severe due to solubility in the solvents used in the degradation that any sequencing information becomes impossible. This is well noted in the few membrane proteins that have been sequenced by conventional means (Furthmayr, 1977).

There have been several approaches developed over the past few years to overcome the first point, that is the ability to sequence proteins and peptides available in only trace amounts (Walsh et al., 1981). These have included i) redesign of instrument to improve vacuum and reduce the level of chemical impurities in the reaction vessel, ii) use of radioactive, coloured or fluorescent isothiocynates to increase the sensitivity of detecting the thiohydantion, and iii) use of radiolabeled amino acids incorporated into the protein. This latter approach was used in the sequencing of protease resistant fragments of VSV since it is the most convenient way of preparing the samples and provides the highest analytical sensitivity. In addition, assignments made based on single amino acid labels are unambiguous.

The standarization of the Beckman sequencer and the detection of PTH amino acids by HPLC is described in the Appendix. The peptides to be sequenced were resolved by SDS polyacrylamide gel electrophoresis following proteolytic digestion of labeled virus.

Detection of [^{3}H] labeled peptides was enhanced by inclusion of [^{35}S]-methionine labeled peptides in the sample. When [^{35}S]-methionine was not desired in the peptide, a sample of methionine labeled VSV digested with protease was run on the same gel in separate wells and used as a guide. Isolation of the peptides was accomplished by cutting out the appropriate band from the gel and either eluting passively by diffusion or electrophoretically into dialysis casing. The recoveries in the case of electrophoretic elution were poor. This was probably due to absorption of peptides onto the dialysis bag, even though lysozyme was included in the buffer to prevent this. All sequencing data reported here was carried out with peptides isolated by diffusion.

To prevent nonspecific losses, plasticware was used wherever possible and all glassware was siliconized. Lysozyme (such that the final amount after elution and lyophilization was 3-5 mg) was included in the elution buffer to prevent losses and to serve as an internal nonradioactive control during the sequencing. Various buffer systems were tried to maximize recoveries. The diffusion buffer used is described in Methods. The inclusion of a reducing agent facilitated recoveries while the SDS insured solubility. The final product that was subjected to sequence analysis contained 1-2 mg of SDS. SDS is not detrimental to the sequencing reactions, in fact, it is beneficial in certain cases since it aids in the retention of nonpolar peptides in the cup (Bailey et al., 1977).

Recovery of proteins from the gel was much improved if the

gel was not fixed or dried prior to autoradiography. The purity of the eluted protein was determined by running an aliquot on a gel prior to sequence analysis.

Initial attempts to sequence the membrane associated fragments were unsuccessful due to extensive losses during the degradations. Most of the radioactivity was extracted by chlorobutane in the first few cycles.

There are various modifications that are available for sequencing short peptides. Most of these concern replacing the Quadrol buffer with more volatile coupling buffers such as DMBA to avoid the need for the ethyl acetate washing step (Hermodson et al., 1972). These programs have been generally replaced by the use of dilute Quadrol (Bauer et al., 1975; Crewther and Inglis, 1975). The most significant advance, however, has been the use of synthetic carriers, in particular polyquaternary amines such as polybrene (Hunkapillar and Hood, 1978; Tarr et al., 1978; Klapper et al., 1978). Polybrene serves to enhance retention of the protein or peptide in the cup and prevent excessive loss during the washing steps. It is particularly useful for sequencing hydrophobic peptides. All the sequencing studies reported in this section, except where indicated, were done in the presence of polybrene.

Partial Amino Acid Sequence of Fragment A

In initial studies, sequencing attempts with thermolysin generated Fragment A were unsuccessful because of extensive peptide losses during the sequencing. This was the case with the

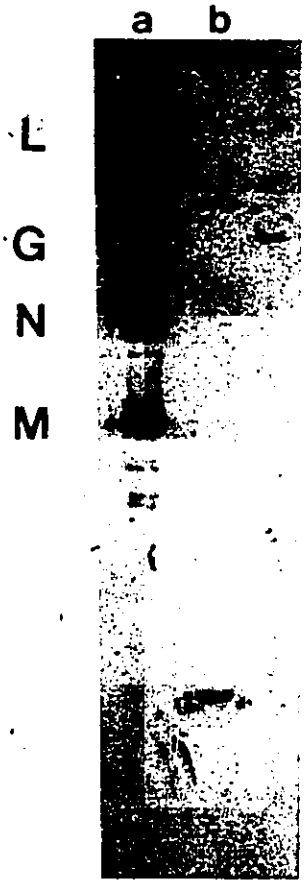


Figure 3.2.1: Polyacrylamide gel analysis of purified Fragment A.

Fragment A was purified from spikeless virions as described in Methods and an aliquot was analyzed on a 17.5% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV; lane b, purified [^{35}S]-methionine labeled Fragment A.

0.1 M Quadrol program as well as with the DMBA program.

The addition of polybrene to the cup reduced the degree of peptide loss, however, in various sequences performed with a variety of labeled amino acids, no radioactive amino acids were released in the cycles examined. In all the cases, the sequence of the internal lysozyme was correct and the yields were good. It was thought that the inability to sequence Fragment A may be due to the thermolysin cleavage site in that the amino terminus may become blocked during the purification or during the actual sequence by some side reaction. Before sequencing, Fragment A was generally solubilized in 50% formic acid. It has been reported that formate buffers tend to produce a variable amount of NH_2-terminal blockage, depending on the length of exposure and concentration of formic acid and the particular amino terminal amino acid (Shively et al., 1982). In addition, during the sequencing reactions, glutaminyl residues have a tendency to cyclize to pyroglutamyl derivatives (Edman, 1970). This possibility is reduced due to the mild cleavage conditions employed in the 0.1 M Quadrol program, however, it cannot be completely discounted.

In attempts to overcome these potential problems, sequencing studies were carried out with Fragment A that was generated by trypsin digestion of VSV. As shown previously (Fig. 3.1.5), both trypsin and thermolysin digestion of VSV produce a Fragment A of approximately similar molecular weight. In addition, trypsin has the advantage of very high substrate specificity as compared with thermolysin. Thus, the use of trypsin would probably produce much

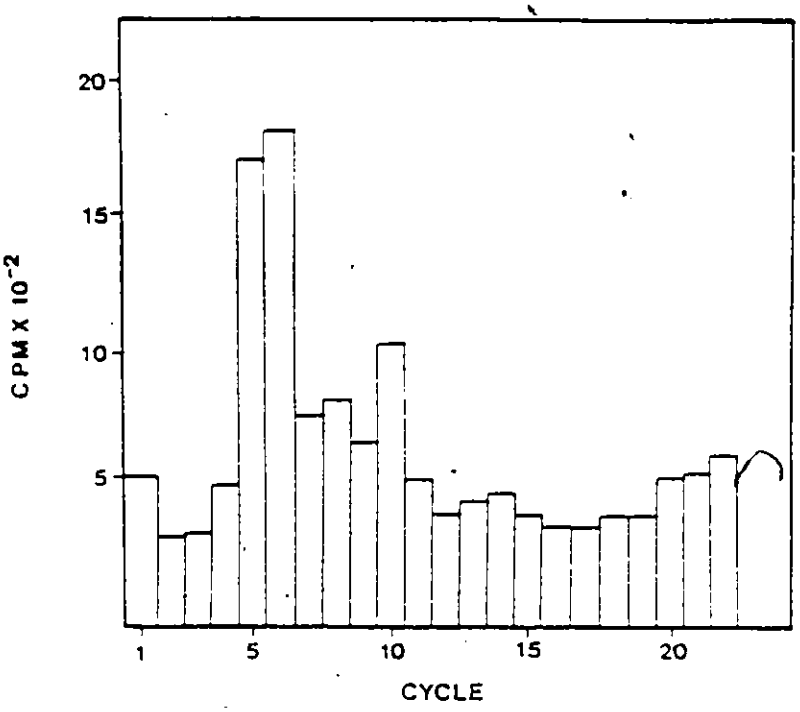
less amino terminal heterogeneity at the cleavage site and consequently reduce the possibility of producing asynchronous sequencing data.

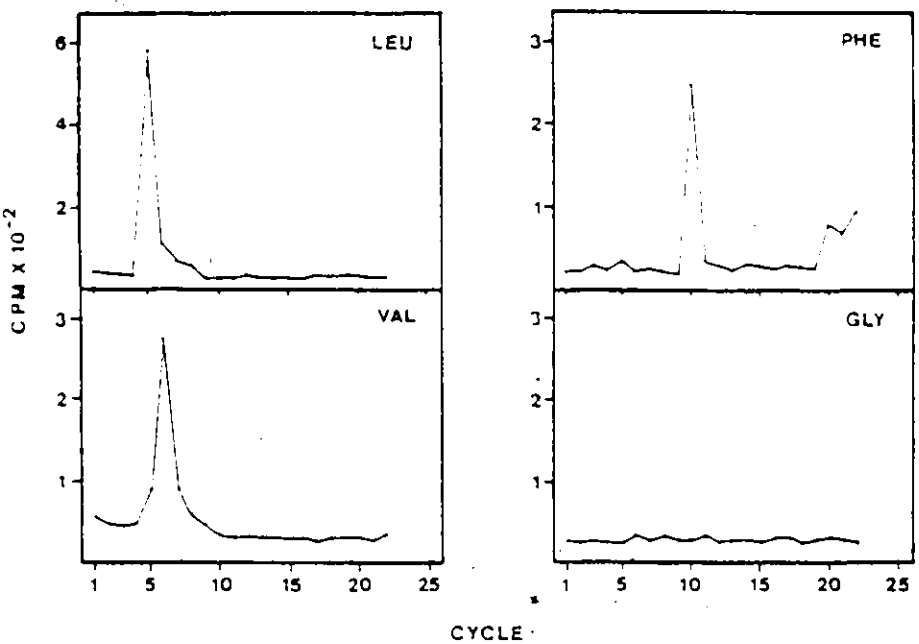
VSV was labeled with a variety of tritiated amino acids, digested with trypsin and the Fragment A was isolated. The radioactive amino acids chosen for incorporation included valine, leucine, phenylalanine, and glycine. These amino acids were chosen because all are hydrophobic or uncharged and, with the exception of glycine, available at a high specific activity. This is important since if Fragment A is membrane bound it is expected to have a high proportion of hydrophobic or nonpolar amino acids. The other important consideration, however, is that the PTH derivatives of these amino acids are well separated by HPLC (see Appendix) so that there is no possibility of incorrect assignment.

Thus, Fragment A labeled with [^{3}H]-Gly, Leu, Val, and Phe was highly purified and subjected to sequencing. In this case, double coupling in the first cycle was performed by closing the HFBA delivery valve to insure a high initial yield. As demonstrated in Figure 3.3.2, radioactivity above background was detected in cycles 5, 6, 10, 20, 21, and 22. The PTH amino acids arising from each cycle were separated by HPLC, fractions were collected, and the radioactivity determined. As shown in Figure 3.2.2, leucine was found in position 5, valine in 6, and phenylalanine in 10, 20, 21, and 22. Although recoveries for Phe in cycles 20, 21 and 22 were relatively low, they could be confidently assigned since the peaks

Figure 3.2.2: Radiosequence analysis of Fragment A labeled with various amino acids.

Fragment A was purified from VSV labeled with [^{3}H]-leucine (134 Ci/mmole), [^{3}H]-valine (58.6 Ci/mmole), [^{3}H]-glycine (12.8 Ci/mmole), and [^{3}H]-phenylalanine (58 Ci/mmole). The purified peptide (80,000 cpm) was sequenced using the 0.1 M Quadrol program with double coupling in the first cycle. The top panel shows the radioactivity released at each cycle. The PTH amino acids recovered from each cycle were mixed with unlabeled carrier PTH amino acids and separated by HPLC. Fractions were collected and the radioactivity was determined. The bottom panels show the radioactivity recovered in the peaks corresponding to the respective PTH amino acid. The repetitive yield from the internal lysozyme was 93%.


20
15
10
5
CPM X 10^{-2}
1
5
10
15
20
CYCLE


LEU
PHE
VAL
GLY
CPM X 10^{-2}
1
5
10
15
20
25
CYCLE

were above background and the repetitive yield (90-93%) was consistent with that of the internal lysozyme. The sequence derived for trypsin generated Fragment A is thus;

```
     1        5              10         15          20
NH2 - - - - LeuVal- - - Phe - - - - - - - - - - PhePhePhe -
```

While this work was in progress, the entire amino acid sequence of G protein as deduced from cloned cDNA was determined (Rose and Gallione, 1981). The mature form of G protein is predicted to contain 495 amino acids. The partial amino acid sequence determined for trypsin generated Fragment A shows no homology with the predicted sequence of the viral proteins NS, N, or M (Rose and Gallione, 1981; Gallione et al., 1981), however, there is a perfect alignment with the predicted amino acid sequence of G protein starting at Asn_{433} (Fig. 3.2.3). It is also evident then that the trypsin cleavage that generated Fragment A occurred between Lys_{432} and Asn_{433}. This localization is in complete agreement with the partial characterization of the protease resistant fragment carried out by Rose et al., (1980).

The Gly_{440} was not detected in the amino acid sequence analysis. This was probably due to the low specific activity (12.3 Ci/mmole) and consequent inefficient incorporation of glycine. Alternatively, this could be due to the difference between the viral strains used by Rose and the one used in this study. Primary structural studies on the G protein from different strains of VSV

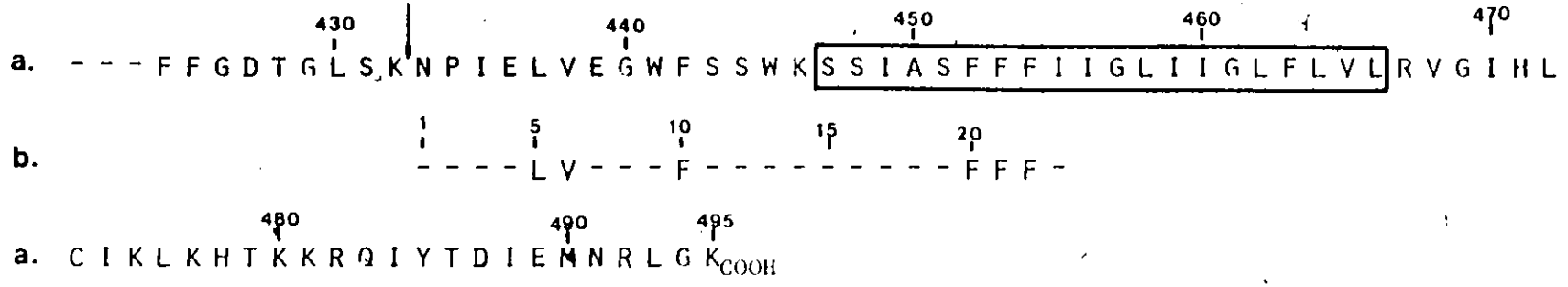


Figure 3.2.3: Comparison of the amino acid sequence of Fragmant A and the amino acid sequence of the COOH terminal region of G protein.

a) carboxy terminal sequence of G protein as deduced from the cDNA sequence data (Rose et al., 1980; Rose and Gallione, 1981). The numbering starts from the amino terminus of the mature protein. The boxed in region represents a continuous hydrophobic domain.

b) partial amino acid sequence of Fragment A as derived from Figure 3.2.2. The arrow indicates the site of trypsin cleavage.

have shown slight differences in amino acid sequences present at the amino terminus of the different glycoproteins (see Appendix, Fig. A-9).

Fragment A appears to encompass the entire carboxy terminal domain of G protein starting at Asn_{433}. The predicted molecular weight of this region is 7311D and agrees quite well with the apparent molecular weight of Fragment A (9000D) predicted from its mobility on polyacrylamide gels. The differences could be due to anamolous migration behaviour, for instance, due to the variable binding of SDS, or due to some type of protein modification (Section 3.3.). This localization also indicates that the carboxy terminus is not exposed and susceptible to proteolytic digestion since the size of the protected fragment, if this were the case, would be smaller than 9000D.

As pointed out by Rose (Rose et al., 1980) and discussed further in the Discussion, the COOH terminal domain has important features that can be related to function. The most important of these is the presence of an uninterrupted stretch of 20 consecutive hydrophobic or uncharged amino acids, which presumably represents the region of G protein that actually spans the membrane. The location of the hydrophobic domain may also explain why only a 3000D reduction in the size of G protein was obtained after digestion with carboxypeptidase A and B (Section 3.1). Presumably, the binding to Triton X100 to the hydrophobic domain halts the action of this exoprotease.

The size of the protected region of G protein as deduced from the cDNA data indicates that it corresponds to approximately 13% of the G protein molecule. From the gel analysis reported in Section 3.1, the protected region also corresponds to approximately 13% (9000D/69000D) when fully glycosylated G is used for comparison. Since carbohydrate makes up approximately 10% of the glycoprotein weight (Patzer et al., 1980) then the membrane protected species would actually represent approximately 14.5% of the polypeptide backbone of G protein.

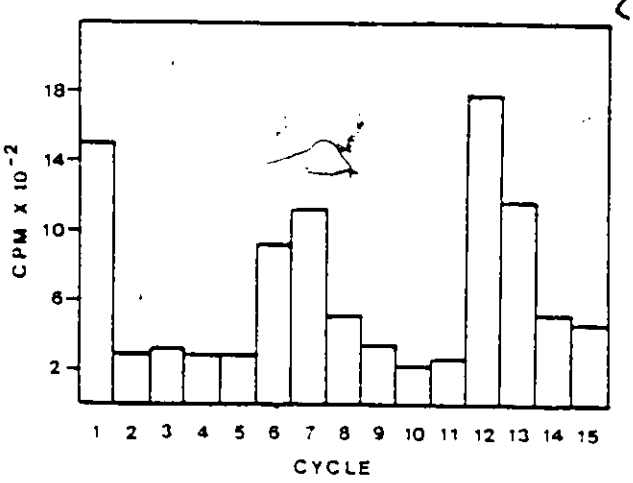
Partial Amino Acid Sequence of Fragment B

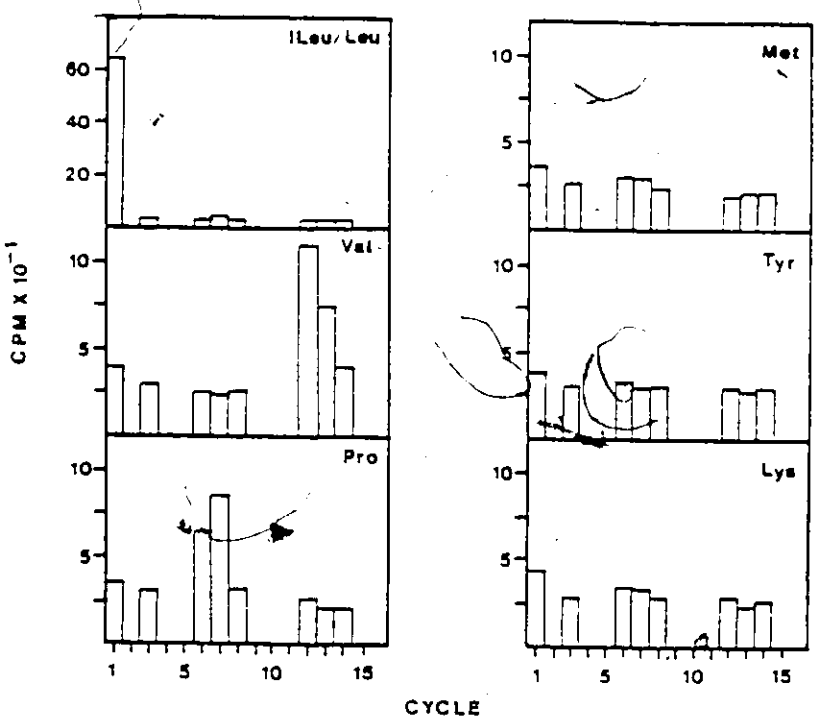
The initial sequencing run on Fragment B was performed at the University of Toronto on a Beckman 890B sequencer using a volatile buffer system containing N-N- dimethylamino benzylamine in the absence of polybrene. VSV labeled with [^{35}S]-methionine and [^{3}H]-Ile, Leu, Val, Pro, Tyr, and Lys was digested with thermolysin and Fragment B was isolated as described. Of the 15 cycles performed, radioactivity above background was detected in cycles 1, 6, 7, 12, and 13 (Figure 3.2.4).

The PTH amino acids released in these cycles were identified by analysis on thin layer paltes (Habener et al., 1978; Jacobs et al., 1974) since an HPLC instrument was not available at the time. Portions of cycles 3, 8, and 14 were analyzed in a similar manner to serve as controls. The results displayed in Figure 3.2.4 indicate the presence of an Ile/Leu in cycle 1, Pro in positions 5 and 6, and Val in positions 12 and 13. The amino acids, Met, Tyr, and Lys, were

Figure 3.2.4: Radiosequence analysis of Fragment B labeled with various amino acids.

Fragment B was purified from VSV which was labeled with [^{35}S]-methionine (1,000 Ci/mmole), [^{3}H]-leucine (59 Ci/mmole), [^{3}H]-isoleucine (40 Ci/mmole), [^{3}H]-tyrosine (52 Ci/mmole), [^{3}H]- valine (29 Ci/mmole), [^{3}H]-lysine (40 Ci/mmole), and [^{3}H]-proline (65 Ci/mmole). The purified peptide (45,000 cpm) was sequenced using the DMBA program. The PTH amino acids were dried, suspended in methanol, and aliquots were counted directly as indicated in the top panel of the figure. The remaining sample from cycles 1,3,6,7,8,12,13, and 14, was spotted along with nonradioactive PTH amino acid markers on a Silica Gel GF thin layer plate (Analtech). The plate was developed in heptane/proprionic acid/ethylene chloride (3.4/1/1.5, v/v/v) and the positions of the PTH amino acid markers were visualized by UV illumination. The areas of the sample lanes corresponding to the markers were scrapped off the plate, extracted 5 times with acetone, and the radioactivity was determined. The panels at the bottom of the figure show the radioactivity recovered for each respective amino acid from the cycles examined.


CPM X 10⁻²
18
14
10
6
2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
CYCLE


CPM X 10⁻¹
ILeu/Leu
Val
Pro
Met
Tyr
Lys
60
40
20
10
5
1 5 10 15
CYCLE

not detected in the first 15 positions. The Ile/Leu in cycle one could not be discriminiated since these two PTH derivatives comigrated on the TLC system used.

This particular run was the only one performed on the Toronto sequencer. The remaining data reported here was derived from analyses carried out on the McMaster sequencer using the 0.1 M Quadrol program and polybrene precycling.

The presence of Leu as the amino terminal amino acid of thermolysin generated Fragment B was confirmed by sequencing Fragment B metabolically labeled with only [^{3}H]-leucine. In this case, leu was found in residue number 1 as well as residue number 16 (Fig. 3.2.5a). The repetitive yield in this case was 95% between Leu_1 and Leu_{16}. Portions of cycles 1, 8, 13, and 17 were examined by HPLC to determine the amino acids released from the internal lysozyme. The internal lysozyme sequenced correctly and had a repetitive yield of 93-95%. The Leu assignments were further confirmed by repeating the sequence of [^{3}H]-Leu labeled Fragment B, however, in this case [^{35}S]-methionine labeled Fragment B was included. As demonstrated in Figure 3.2.4b, Leu was detected in position 1 and 16 and, additionally, methionine was present in positions 30, 31, and possibly 32.

The absence of isoleucine in the first position was demonstrated by sequencing Fragment B labeled with [^{3}H]-Ile and [^{35}S]-Met (Fig. 3.2.5b). No tritium counts were released in the first 33 cycles examined, however, methionine was present in

Figure 3.2.5: Radiosequence analysis of Fragment B labeled with [^{3}H]-leucine and [^{35}S]-methionine.

a) sequence of Fragment B labeled with [^{3}H]-leucine (40,000 cpm).
b) sequence of Fragment B labeled with [^{3}H]-leucine (30,000 cpm) and [^{35}S]-methionine (15,000 cpm).

The butyl chloride exracts were dried down, suspended in a small volume of methanol, and 80% counted directly. The remainder, after drying and conversion with 25% TFA was analyzed by HPLC to check the sequence of the internal lysozyme. (•) [^{3}H] radioactivity; (o) [^{35}S] radioactivity.

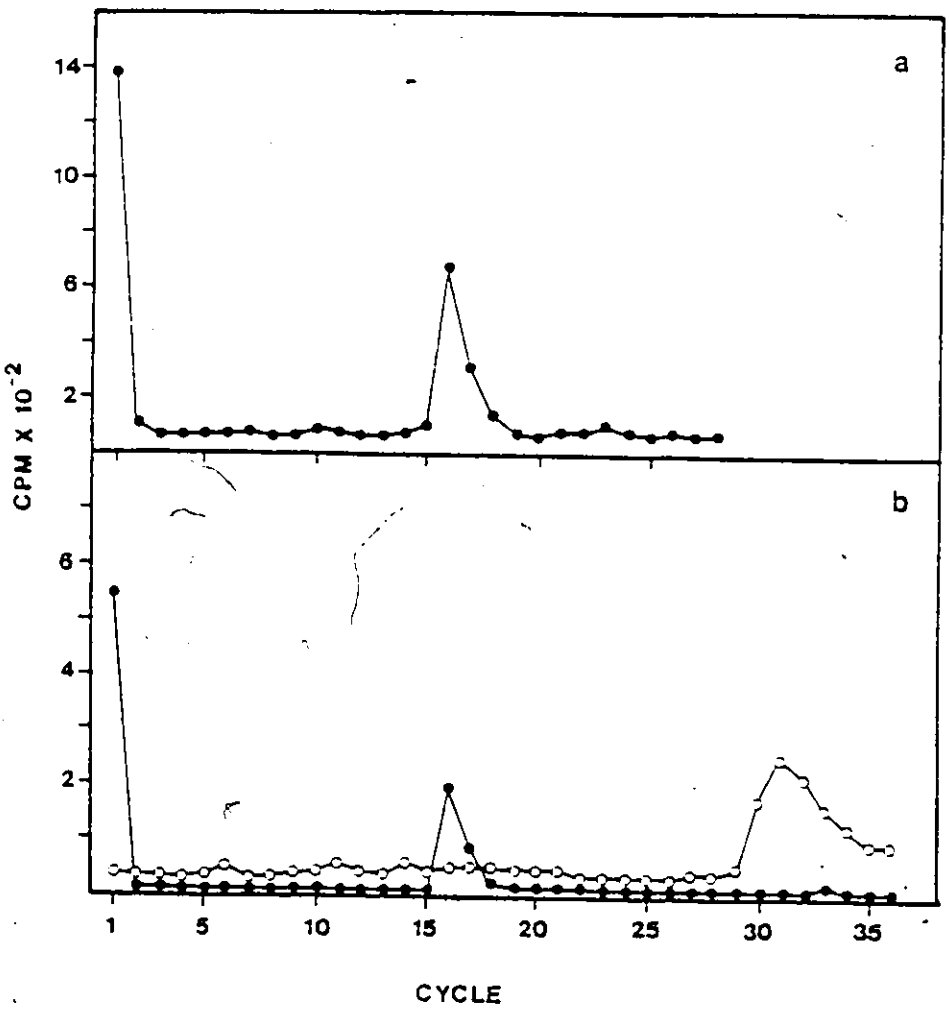
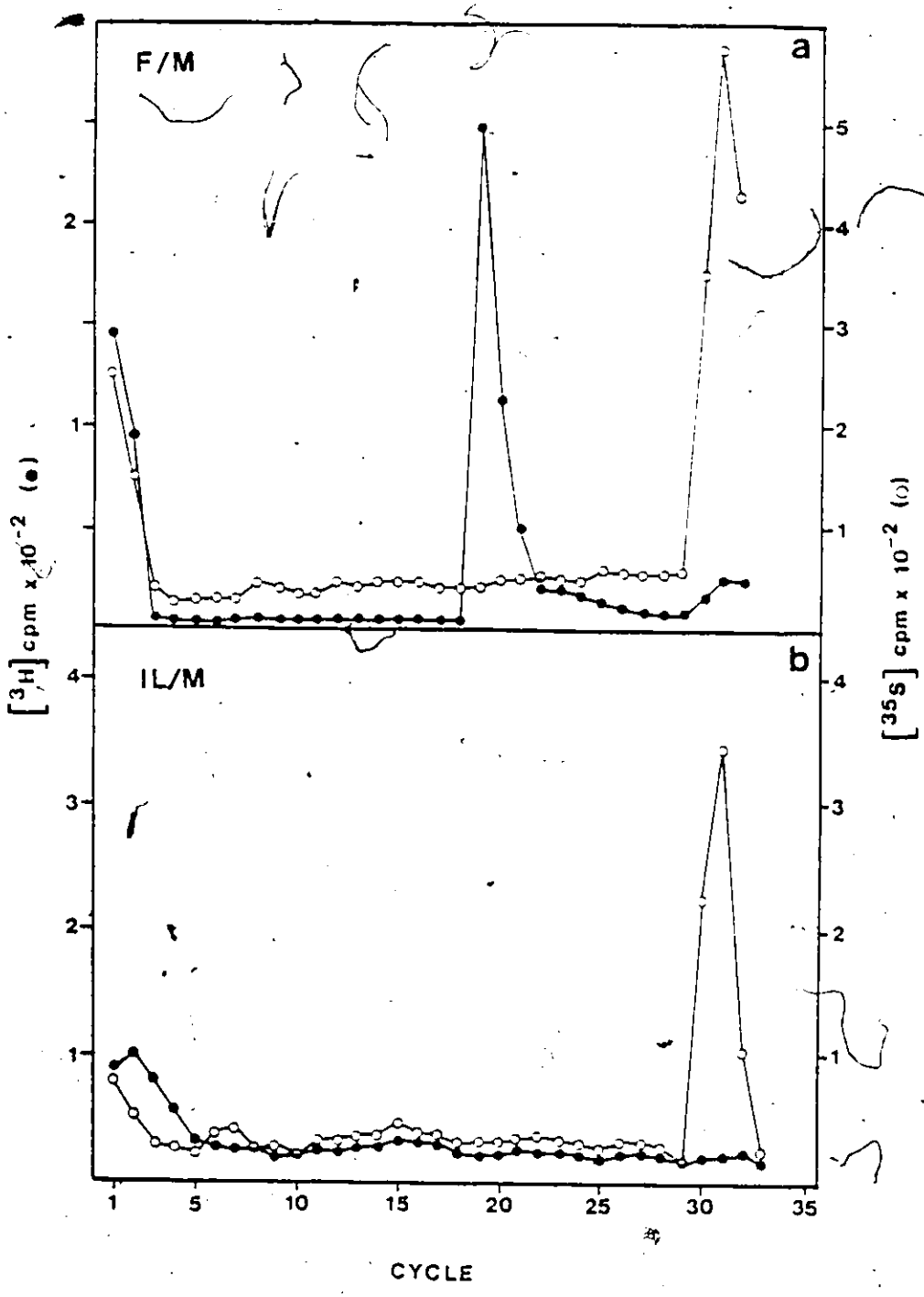

a
b
14
10
6
2
6
4
2
CPM X 10^{-2}
1
5
10
15
20
25
30
35
CYCLE

Figure 3.2.6: Radiosequence analysis of Fragment B labeled with [^{3}H]-phenylalanine or [^{3}H]-isoleucine.

a) Fragment B labeled with [^{3}H]-phenylalanine (50,000 cpm) and [^{35}S]-methionine (25,000 cpm).
b) Fragment B labeled with [^{3}H]-isoleucine (40,000 cpm) and [^{35}S]-methionine (15,000 cpm).
(•) [^{3}H] radioactivity; (o) [^{35}S] radioactivity.


F/M
a
IL/M
b
[3H] cpm x 10-2 (●)
[35S] cpm x 10-2 (○)
CYCLE
1
5
10
15
20
25
30
35

positions 30 and 31. Additional sequencing information was obtained from the analysis of Fragment B labeled with [^{3}H]-Phe and [^{35}S]-Met (Fig. 3.2.6a). In this case, Phe was detected in cycle 19 while again the methionines were present in positions 30, 31, and possibly 32. The high background present in cycles 1 and 2 in Figure 3.2.6a represents protein washout since the $^{3}H/^{35}S$ ratio was the same as the starting material.

The partial amino acid sequence of thermolysin generated Fragment B was thus found to be:

```
     1          5          10         15         20         25
NH2 Leu - - - - ProPro - - - - ValVal - - Leu - - Phe - - - - - -
          30
- - - - MetMet(Met)- - -
```

The assignment of methionine in position 32 is tentative based on the following considerations. There is always a certain amount of carry over of ATZ amino acids from one cycle into the following cycle (see Appendix). This can be seen in the sequence of [^{3}H]-Phe labeled Fragment B in Figure 3.2.6a. Cycle 20 contains 35-40% carry over from the [^{3}H] radioactivity present in cycle 19. This is a normal occurrence in radiosequencing and in general it can easily be detected as carry over by examining the repetitive yield. The problem is compounded, however, when there are sequential amino acids of the same type as is the case for methionine in cycles 30 and 31. Thus, the methionine radioactivity detected in cycle 32 could represent carry over from the previous two cycles. This is probably

the case since in the sequence of [^{3}H]-Ile labeled Fragment B (Fig. 3.2.6b), the relative amount of [^{35}S] radioactivity in cycle 32 compared to 30 and 31 is much closer to the background level as compared to the [^{3}H]-Phe analysis.

As indicated in Section 3.1 and shown more clearly in Figure 3.2.7, the size of Fragment B was smaller in the case of thermolysin digestion as opposed to the case where trypsin was used to generate the membrane associated fragments. The molecular weight difference was approximately 1000 - 1500D as judged by SDS polyacrylamide gel electrophoresis.

A sequence analysis was performed on Fragment B generated from trypsinization of VSV labeled with [^{3}H]-Leu. In this instance, Leu radioactivity was detected only in cycle 10 (Figure 3.2.8). Since there is only one definitive assignment, one cannot overlap the sequence with the data derived from thermolysin generated Fragment B.

The data derived from the amino acid sequence of Fragment B was used to compare with that predicted for the viral protains from cDNA data (Rose and Gallione, 1981; Gallione et al., 1981). No homology was observed with either G, NS, or M proteins, however, a perfect alignment was observed with the N protein. According to the cDNA data, N protein contains 422 amino acids. The sequence deduced from thermolysin generated Fragment B matches a region of N protein starting at Leu_{364} (Fig. 3.2.9). The cDNA sequence data predicts a Gln residue in place of the methionine that was tentatively assigned

Figure 3.2.7: Purification of Fragment B from VSV digested with trypsin or thermolysin.

[^{35}S]-Methionine labeled VSV was digested with trypsin or thermolysin and Fragment B was purified as described in Methods. Aliquots of the purified material were analyzed on a 17.5% polyacrylamide gel. Lane a, trypsin generated Fragment B; lane b, thermolysin generated Fragment B.

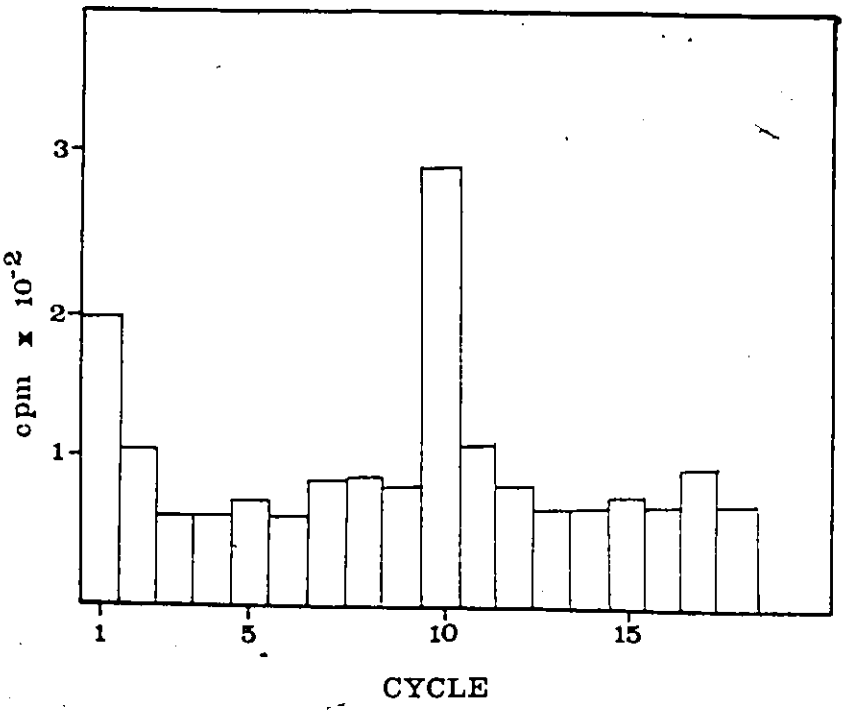
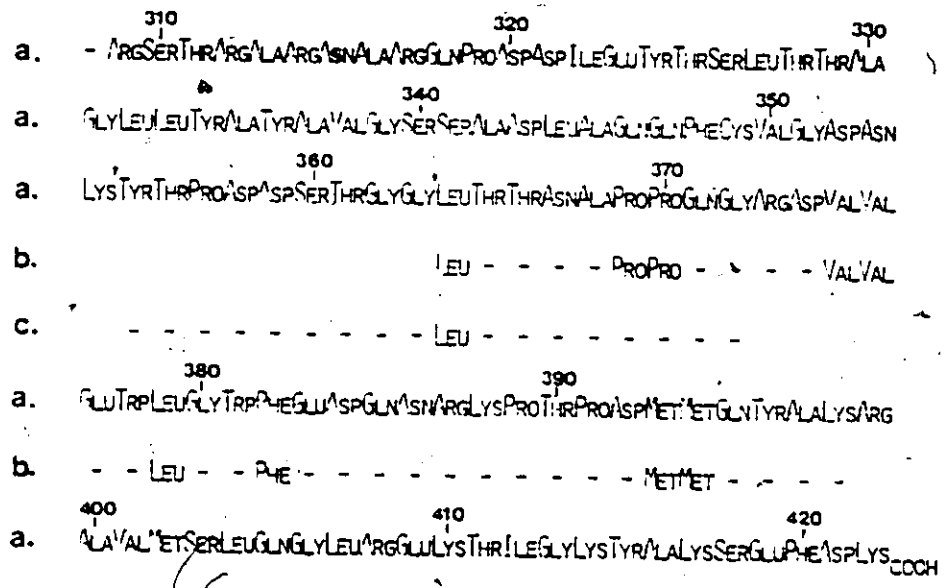


Figure 3.2.8: Radiosequence analysis of [^{3}H]-leucine labeled Fragment B purified from trypsin digested virus.

Figure 3.2.9: Comparison of the amino acid sequence of fragment B and the amino acid sequence of the N protein from VSV



a) amino acid sequence of N protein of VSV as deduced from cDNA (Gallione et. al. 1981). The numbering is from the amino terminus and only the sequence starting from residue 309 is displayed.
b) amino acid sequence determined for thermolysin generated fragment B;
c) amino acid sequence determined for trypsin generated fragment B. The arrows indicate the sites of enzyme cleavage.

at position 32 from the amino acid sequence of Fragment B. The site of thermolysin cleavage, indicated by an arrow over Gly_{363}-Leu_{364}, is consistent with the enzyme specificity. In addition, the data derived from the sequence of [^{3}H]-Leu trypsin generated Fragment B could be aligned. The amino acid data predicted a Leu at position 10 from the amino terminus in trypsin generated Fragment B (Fig. 3.2.8). If this residue is taken as the amino terminal Leu found in thermolysin derived Fragment B, then the nearest lysine or arginine residue amino terminal to this point is found at Lys_{354}. If this is taken as the trypsin cleavage site, then the difference in molecular weight between trypsin and thermolysin generated Fragment B would be 1055D. This is in excellent agreement with the differences in molecular weight of the fragments as determined by gel electrophoresis. The next closest trypsin sensitive site is located at Arg_{317}. Although there is also a leucine residue located 10 amino acids downstream from this site (at Leu_{327}), this is an unlikely cleavage site since the molecular weight differences of the species generated by the two proteases would be approximately 5000D. In addition, the leucines present at position 332 and 333 would have been detected by the sequencing.

Thus, Fragment B is derived from a region at the COOH terminus of N protein by proteolytic cleavage at a very specific site. The specificity of cleavage is evidenced from the fact that in sequencing experiments performed, Fragment B was prepared from different viral preparations, digested at different times, and

purified at different times. The presence of methionine in position 30 and 31 in all cases indicates that the thermolysin cleavage stie is always located between Gly_{363} and Leu_{364}. This is all the more intriguing since thermolysin has a broad substrate specificity and there are several thermolysin sensitive amino acids flanking this site. The fact that Leu_{364} is the preferred site of thermolysin attack may be a reflection of the tertiary conformation of N protein or how it is complexed in the RNP. For instance, the proline residues located five and six amino acids downstream from the thermolysin cleavage site could conceivably cause a bend in the polypeptide chain thus exposing this site to enzymatic attack. N protein complexed in the RNP structure is highly resistant to proteolytic attacked, however, a product equivalent to Fragment B would have previously gone undetected because of its small size.

Whether Fragment B actually represents the entire carboxy terminal region of N protein is not known. The predicted molecular weight of this region from thermolysin digestion (Leu_{364}-Lys_{422}) is 7363 D and from the trypsin cleavage site (Tyr_{355}-Lys_{422}) it is 8418D whereas the molecular weights determined by gel electrophoresis are 6000 and 7500D, respectively. The differences between the predicted and observed sizes could be due to anamolous electrophoretic behaviour on polyacrylamide gels. This is a well known phenomena for low molecular weight species. Alternatively, the enzymes may also be cleaving N at an additional site very close to the carboxy terminus which could account for the lower observed

molecular weight of B. Examination of the predicted sequence of N protein from Met_{402} to Lys_{422} reveals a number of amino acids that are potential cleavage sites for either thermolysin or trypsin. Secondary cleavage at these sites would generate fragments of molecular weight closer to those observed on the gels.

Cleavage at a secondary site could also be used to explain some of the possibly conflicting results obtained with the fingerprinting data shown in Figure 3.1.20. Tryptic peptide analysis of [^{35}S]-methionine labeled Fragment B generated from thermolysin digestion of VSV generated three tryptic peptides, only one of which was in common with N protein. From the sequence data, only two methionine containing tryptic peptides should be present (Pro_{389} - Lys_{398} and Ala_{400} - Arg_{408}). The peptide present that shows no homology with N protein could arise if thermolysin also cleaved within the second methionine containing peptide, for instance between Gly_{406} and Leu_{407}. A secondary cleavage at this site would generate a Fragment B more consistent with the molecular weight as determined by gel electrophoresis.

An examination of the sequence of Fragment B and Fragment A also indicates why there seems to be a relatively large amount of Fragment B even though it arises from degradation of N protein. As stated in Section 3.1, A and B represent 2.7 and 1.7% of the methionine radioactivity remaining associated with the spikeless virions following proteolytic digestion. The amino acid sequences of Fragment A and B show, however, that while A contains only one

methioine, B contains three. Thus, the ratio of radioactivity between A and B is actually closer to 5:1. When one considers that N protein comprises approximately 40% of the total viral protein as opposed to less than 20% for G protein (Wagner, 1975), then one can estimate that degradation of only 10% of N protein present in the virus preparation would be sufficient to produce a Fragment B of the intensity observed. A preparation of purified VSV in which 10% of the viral particles are susceptible to degradation of internal structural components is not unexpected.

The sequencing work presented in this section in conjunction with published nucleic acid sequences of the VSV viral genes has allowed the unambiguous localization of the protease resistant, small molecular weight fragments of VSV. It has also been demonstrated that G protein is anchored in the membrane by a hydrophobic domain located at the COOH terminal end of the molecule.

SECTION 3.3: LIPID MODIFICATION OF THE VSV G PROTEIN

The VSV glycoprotein G undergoes a number of discrete modifications during its maturation. These include signal sequence cleavage, and addition and processing of oligosaccharide chains (Ghosh, 1980). Recently, a new type of modification has been described for G protein as well as for a number of other viral membrane proteins, namely, the attachment of fatty acid residues (Schmidt et al., 1979; Schmidt and Schlesinger, 1979; Schmidt, 1982a). In the case of the VSV glycoprotein, it was determined that 1-2 moles of fatty acid were covalently bound per mole of protein. The covalent nature of the linkage was deduced from the observations that the protein bound fatty acids could not be removed with denaturants such as SDS, urea, or guanidine hydrochloride, or by exhaustive extraction with organic solvents. The fatty acids, however, could be quantitatively released by mild alkaline hydrolysis in methanol and recovered as the methyl ester derivatives. Since the presence of phospholipid or glycerol was not detected, the data indicated that the fatty acids were covalently attached by ester linkage directly to the polypeptide backbone.

The addition of the fatty acid moiety appears to occur at a point late in the maturation of the glycoprotein, while it is in transit to the plasma membrane (Schmidt and Schlesinger, 1980). Acylation appears to occur at a point shortly before G protein

becomes resistant to Endo- β -N-Acetylglucosaminidase H which is thought to occur in the Golgi complex (Robbins et al., 1977). Endo- β -N-Acetylglucosaminidase H removes high mannose carbohydrate chains but not processed oligosaccharides from proteins (Robbins et al., 1977; Tarantino and Maley, 1977).

One possible function of the covalently bound fatty acids is that they provide an additional means by which the glycoprotein can interact with the viral membrane. If this is the case, then one would expect the fatty acids to be located at a point on the protein which is in close proximity or in direct contact with the lipid bilayer. To test this proposition, the relative location of the fatty acids on G protein was examined.

As demonstrated in Figure 3.3.1, SDS polyacrylamide gel analysis of VSV that had been labeled with [^{3}H]-palmitate revealed that only the G protein was labeled. The fact that none of the other viral proteins were labeled to any extent indicates that the labeled palmitate was not appreciably metabolically converted into other molecules such as amino acids prior to incorporation into proteins.

The large amount of label migrating at the dye front is mostly composed of viral phospholipids. This was ascertained by examining the membrane lipids of VSV which was labeled with [^{3}H]-palmitate. Thus, viral lipids were extracted from [^{3}H]-palmitate labeled VSV by the Bligh-Dyer procedure and analyzed by thin layer chromatography. The results shown in

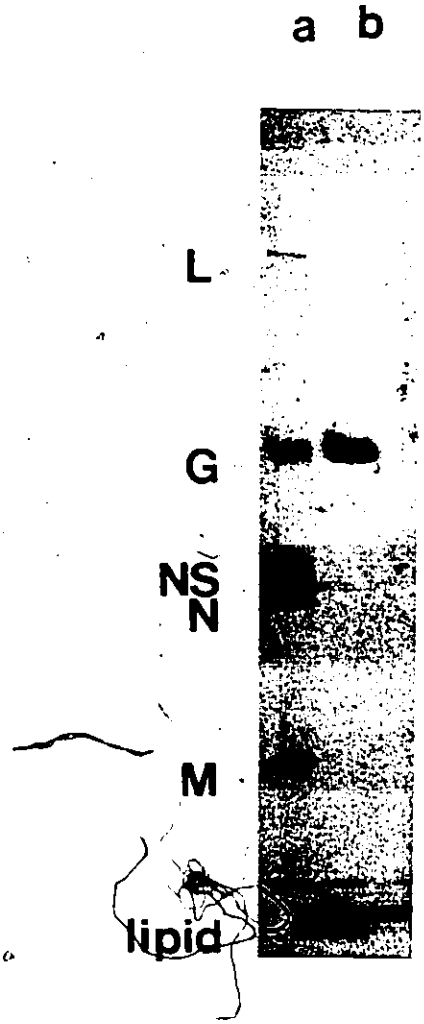


Figure 3.3.1: Polyacrylamide gel analysis of VSV labeled with [^{3}H]-palmitate.

VSV was labeled with [^{3}H]-palmitate as described in Methods and analyzed on a 10% polyacrylamide gel. Lanes a and b, represent VSV labeled with [^{35}S]-methionine and [^{3}H]-palmitate, respectively.

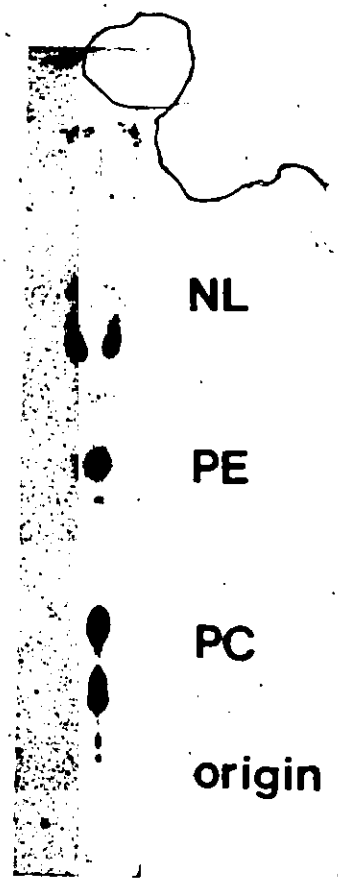
Figure 3.3.2 indicate that much of the label incorporated into the VSV lipid migrates as phosphatidylcholine and phosphatidyl ethanolamine. These are two of the major phospholipids found in the viral envelope when VSV is grown in L cells (McSharry and Wagner, 1971).

The labeling of G protein by palmitate did not seem to be due to noncovalent interactions since the label remained bound to G protein after boiling in SDS containing buffer and extraction of the gel with methanol/acetic acid or with 100% dimethylsulphoxide (Figure 3.3.1). The covalent nature of the linkage was further confirmed by treating SDS gels in which [^{3}H]-palmitate labeled VSV was fractionated with hydroxylamine prior to fluorography (Omary and Trowbridge, 1981). Thus, VSV labeled with [^{3}S]-methionine or [^{3}H]-palmitate was run in duplicate on a 10% polyacrylamide gel and one half of the gel was treated with NH_2OH prior to fluorography. As shown in Figure 3.3.3, NH_2OH treatment causes a substantial loss of [^{3}H] radioactivity from G protein (compare lanes b and d). Approximately 90% of the label is removed after a two hour treatment with hydroxylamine. Longer treatment results in complete loss of label (see Section 3.4). On the other hand, no detectable loss of radioactivity was observed when [^{35}S]-methionine labeled VSV was treated in an identical manner (Fig. 3.3.3 lanes a and b).

In order to ascertain what fraction of the total amount of [^{3}H]-palmitate incorporated into the virus was actually tightly bound to G protein, [^{3}H]-palmitate labeled VSV was solubilized

Figure 3.3.2: Phospholipid analysis of VSV labeled with [^{3}H]-palmitate.

Lipids were extracted from [^{3}H]-palmitate labeled VSV as described in Methods. An aliquot of the lipid extract was dried, suspended in $CHCl_3$/MeOH (1:1, v/v) and spotted on a silica H1B5 thin layer plate along with unlabeled phospholipid markers. The plate was developed in $CHCl_3$/MeOH/H_2O (65:25:4, v/v/v) and the radioactive spots were visualized by autoradiography. The positions of the nonradioactive markers were determined by staining the plate with I_2 vapour. NL, neutral lipids.


NL
PE
PC
origin

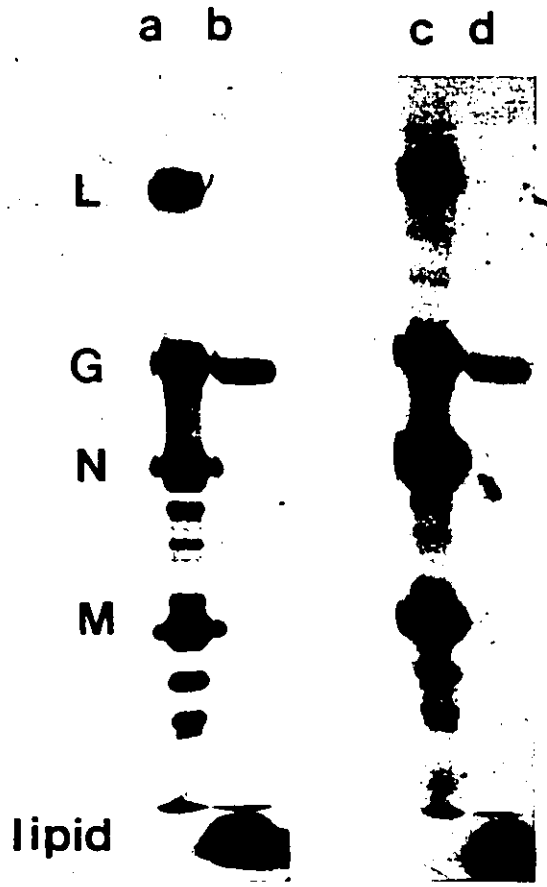


Figure 3.3.3: Effect of hydroxylamine on [^{35}S]-methionine and [^{3}H]-palmitate labeled VSV.

VSV, labeled with [^{35}S]-methionine or [^{3}H]-palmitate was run in duplicate on a 10% polyacrylamide gel. Following electrophoresis, half of the gel was directly processed by fluorography while the other half was treated with 1 N NH_2OH (pH 6.6) for 2 hours prior to fluorography. Lanes a and c, [^{35}S]-methionine labeled VSV before and after treatment with NH_2OH, respectively; lanes b and d, [^{3}H]-palmitate labeled VSV before and after treatment with NH_2OH, respectively.

with Triton X100 and the supernatant containing the G protein and the viral lipids was taken through a series os sequential organic extractions as outlined in Table II. A total of 3.8×10^6 cpm was used for the extractions. 97% of the radioactivity was accounted for after the different extractions and 98% of these counts were in organic phase. 1.8% of the radioactivity remained associated with G protein which precipitated under the extraction conditions. This radioactivity was refractory to further organic extraction and thus, in very tight association with G protein. This tightly bound radioactivity could be quantitatively removed from G protein by mild alkaline hydrolysis with 0.1 N methanolic KOH. After this treatment, greater than 90% of the radioactivity tightly bound to G protein was extractable with petroleum ether. Gel analysis of the G protein remaining in the aqueous phase after alkaline hydrolysis showed that all of the radioactivity was removed from G protein (Fig. 3.3.4 lane d).

The radioactivity remaining in the organic phase after transesterification was identified as the methyl ester of palmitate by thin layer chromatography (see Fig. 3.3.6). Thus, in agreement with the work of Schlesinger, the fatty acids found on G protein are in ester linkage to the polypeptide backbone.

If the fatty acids play a role in maintaining the proper disposition of G in the membrane or provide an additional means of protein-lipid interaction because of their lipophilic nature, then one would expect the fatty acids to be present on that portion of

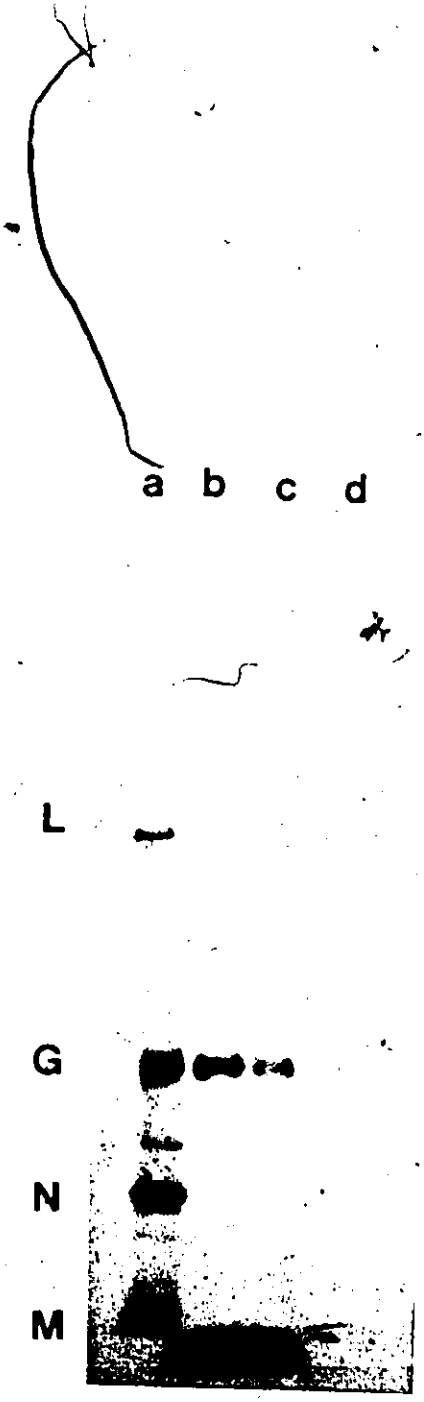
TABLE II: ORGANIC EXTRACTION OF G PROTEIN LABELED WITH [^{3}H]-PALMITATE.

ORGANIC EXTRACTION	RADIOACTIVITY IN ORGANIC LAYER (CPM)	TOTAL RADIOACTIVITY RECOVERED %
METHANOL-CHLOROFORM (2:1)	3.55×10^6	93.4
METHANOL-CHLOROFORM-WATER (2:1:0.8)	5.69×10^4	1.5
METHANOL-CHLOROFORM-WATER (2:1:0.8)	6.3×10^3	0.17
SUM OF ORGANIC EXTRACTS	3.61×10^6	95.1
RESIDUAL G PROTEIN	6.4×10^4	1.8

[^{3}H]-Palmitate labeled G protein was isolated from labeled VSV by extraction with Triton X100 as described in Methods. One ml of the detergent supernatant containing G protein and the viral lipids (3.8×10^6 cpm) was subjected to sequential organic solvent extractions as indicated in the table. 97% of the radioactivity (3.68×10^6 cpm) was recovered after the various extractions. The amount of radioactivity remaining tightly associated with G protein was determined by suspending the final protein pellet and counting an aliquot.

Figure 3.3.4: Effect of alkali on [^{3}H]-palmitate labeled G protein.

The [^{3}H]-palmitate labeled G protein remaining after the organic extractions described in Table II was suspended in 0.1 N methanolic KOH and incubated at 23°C for 30 minutes. After this period, the reaction was acidified with conc. HCl and extracted with petroleum ether. The aqueous phase was dried, suspended in electrophoresis sample buffer, and analyzed on a 10% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV; lane b, [^{3}H]-palmitate labeled VSV; lane c, material present in the Triton X100 supernatant after detergent disruption and centrifugation of [^{3}H]-palmitate labeled VSV; lane d, material remaining in the aqueous phase after transesterification and organic solvent extraction of [^{3}H]-palmitate labeled G protein.


a b c d
L
G
N
M

the molecule that is in contact with the lipid bilayer. To test this directly, VSV labeled with [^{35}S]-methionine or [^{3}H]-palmitate was digested with trypsin or thermolysin and the spikeless particles were analyzed on a polyacrylamide gel. As demonstrated in Figure 3.3.5, digestion of [^{3}H]-palmitate labeled VSV with either of these two proteases results in the complete loss of G protein and the appearance of a new labeled species migrating at a position corresponding to Fragment A, the membrane anchoring domain of G protein. No other labeled protein was observed indicating that all fatty acids present on G protein are located in the membrane interacting region.

In order to quantitate this more accurately, [^{3}H]-palmitate labeled VSV was run on a polyacrylamide gel before and after treatment with thermolysin and the regions of the gel corresponding to G protein, fragments A and B, and the lipid front were excised from the gel and the radioactivity was determined. As shown in Table III, 2.2% of the palmitate radioactivity present in VSV is present in G protein prior to enzymatic digestion. This is in good agreement with the value determined by organic extraction of detergent solubilized glycoprotein presented in Table II. After thermolysin treatment, 2.3% of the total radioactivity is found in fragment A. No radioactivity migrates at a position in the gel corresponding to fragment B. Thus, all of the fatty acids present on G protein are localized to the region of the protein that is interacting with the lipid bilayer.

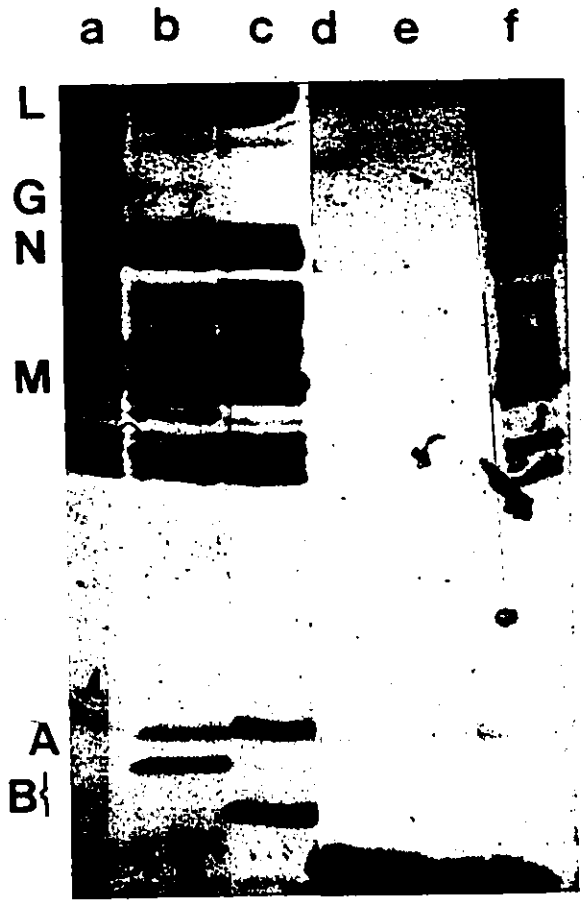


Figure 3.3.5: Proteolytic digestion of VSV labeled with [^{3}H]-palmitate.

[^{35}S]-Methionine and [^{3}H]-palmitate labeled VSV was digested with trypsin or thermolysin and analyzed on a 17.5% polyacrylamide gel. Lanes a and f, [^{35}S]-methionine labeled VSV marker; lanes b and c, [^{35}S]-methionine labeled VSV treated with trypsin and thermolysin, respectively; lanes d and e, [^{3}H]-palmitate labeled VSV treated with trypsin and thermolysin, respectively.

Table III: Quantitation of [^{3}H]-palmitate radioactivity present in Fragment A.

Gel Region	Untreated VSV cpm	% of total cpm	Thermolysin digested VSV	% of total cpm
G Protein	327	2.2	57	
Fragment A	30		305	2.3
Fragment B	35		45	
Lipid front	15,000		13,100	

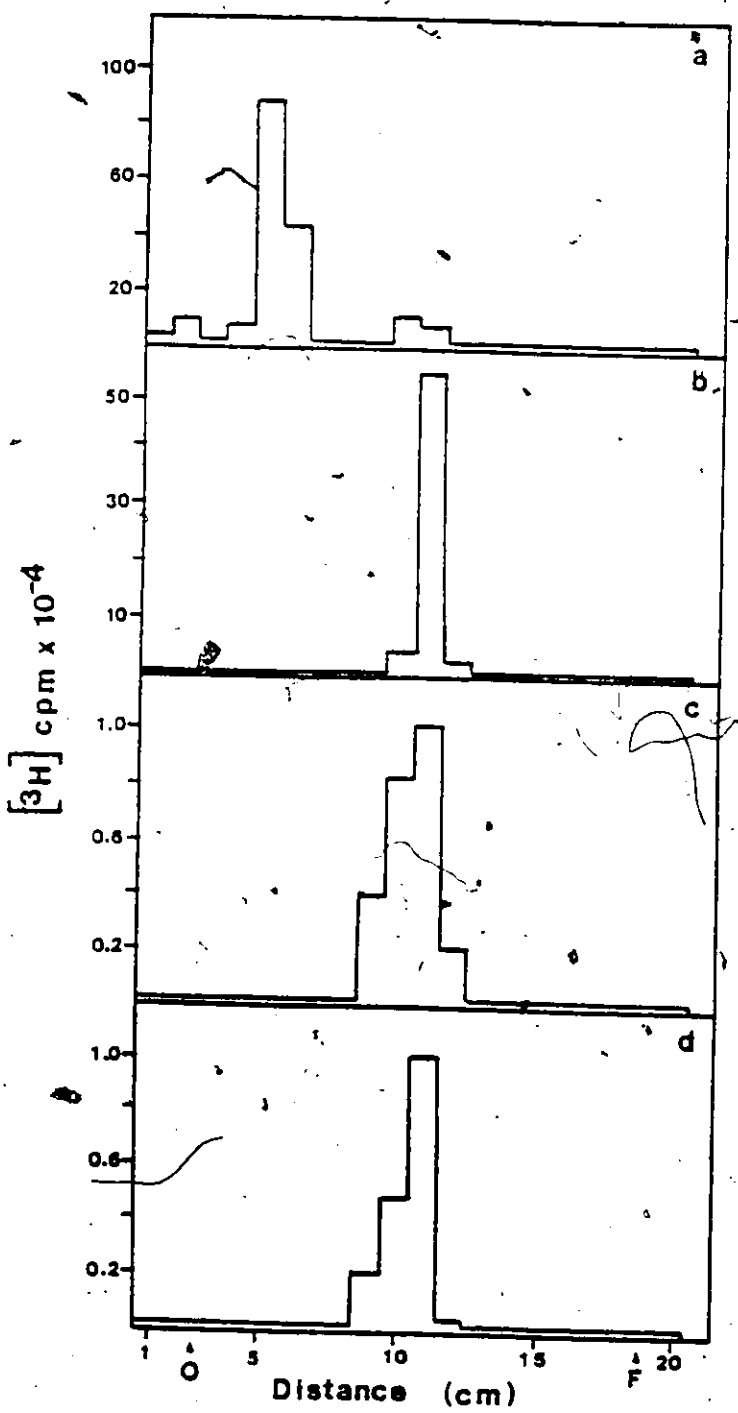
[^{3}H]-palmitate labeled VSV, treated or untreated with thermolysin was sedimented through a glycerol cushion and the pelleted virus was solubilized in electrophoresis sample buffer and run on a 17.5% polyacrylamide gel. Regions of the gel corresponding to G protein, the peptides A and B, and the gel front were excised from the dried gel, eluted with 0.05% SDS, and the radioactivity was determined.

In order to determine whether the radioactivity present in fragment A was still present as palmitate, the material derived from fragment A after transesterification and extraction with organic solvent was analyzed by thin layer chromatography. In this case, [^{3}H]-palmitate labeled fragment A was not delipidized by extraction with organic solvents due to excessive losses encountered because of the hydrophobic nature of the peptide. Rather, fragment A was delipidized by SDS gel electrophoresis. Thus, [^{3}H]-palmitate labeled Fragment A was isolated from a polyacrylamide gel as described in Section 3.2. Of the material recovered, 18,000 cpm was lyophilized and suspended directly in 0.1N methanolic KOH. After incubation at 23°C for 30 min, the reaction was acidified with concentrated HCl and extracted 4 times with petroleum ether. Approximately 16,000 cpm (89%) was recovered in the organic phase. Analysis of the radioactivity present in the organic phase by TLC revealed that it migrated as the methyl ester of palmitate (Fig. 3.3.6). Thus the fatty acids present in the membrane anchoring region of G protein are found in ester linkage to the polypeptide backbone.

The exact point of attachment of the fatty acids is obviously a point of great interest. Since the fatty acids appear to be in ester linkage directly to the polypeptide chain, potential acylation sites include the hydroxy amino acids serine, threonine and tyrosine. Cystine is also a potential candidate for a thioester linkage, however this type of linkage would be unstable since the high

Figure 3.3.6: Analysis of radioactivity released after transesterification of [^{3}H]-palmitate labeled G protein and Fragment A.

Delipidized [^{3}H]-palmitate labeled G protein (by organic extraction as described in Table II) or Fragment A (by elution from polyacrylamide gels) was incubated with 0.1 N KOH for 30 minutes at 23°C, acidified with conc. HCl, and extracted with petroleum ether. The organic phase was taken to dryness, suspended in a small amount of petroleum ether, and spotted on a silica H1B5 plate. The plate was developed with petroleum ether/ether (1:1, v/v) after which the lanes corresponding to the samples were scrapped in 1 cm divisions, eluted, and the radioactivity was determined. Panel a, [^{3}H]-palmitate; panel b, methyl ester derivative of [^{3}H]-palmitate; panel c, organic extractable material obtained from the transesterification of [^{3}H]-palmitate labeled G protein; panel d, organic extractable material obtained from the transesterification of [^{3}H]-palmitate labeled Fragment A. The palmitate methyl ester marker was prepared by the methylation of [^{3}H]-palmitate with freshly prepared diazomethane (Kates, 1972) and was kindly supplied by P. Leblanc. (o) origin; (F) solvent front.


a
b
c
d
100
60
20
50
30
10
1.0
0.6
0.2
1.0
0.6
0.2
[3H] cpm x 10-4
1
O
5
10
15
F
20
Distance (cm)

energy bond would be fairly reactive.

An examination of the COOH-terminal sequence of G protein (Fig. 3.2.3) reveals the presence of several amino acids that could serve as the acylation sites. These include serines at positions 443, 444, 447, 448, and 451, threonines at positions 479 and 486, and a tyrosine at position 485.

Indirect evidence suggests that the fatty acids are covalently attached to serine residues (Schmidt and Schlesinger, 1979; Schlesinger et al., 1981). This comes from the observation that a fatty acid containing peptide derived from a pronase digestion of [^{3}H]-palmitate labeled G protein contained a high proportion of serine residues. Acylation of serine residues, especially those present in the actual membrane anchoring segment of G could serve to mask the hydrophilic hydroxyl groups and thereby supplement the hydrophobicity of the membrane interacting region.

Attempts to determine the site of fatty acid attachment by amino acid sequence analysis were unsuccessful. In several sequencing attempts using highly purified [^{3}H]-palmitate labeled fragment A generated by trypsin digestion, radioactivity was released at distinct cycles, however, the position was not consistent in different experiments. This variability was probably not due to peptide washout since amino acid labeled sample could be readily sequenced and, in addition, the [^{3}H]-palmitate radioactivity was released at later rather than initial cycles. If peptide loss was

occurring, the radioactivity would have been detected in the early cycles.

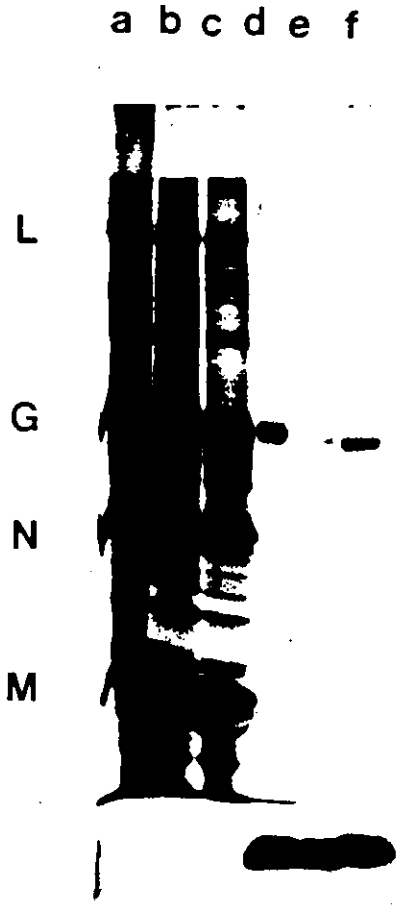
An explanation for these types of result may be that the fatty acid is being cleaved off the protein during the degradation reactions. The ester linkage should be very stable under the conditions of degradation, however, the possibility that small amounts of water may be present in the HFBA which would result in a variable amount of acid hydrolysis of the fatty acid with subsequent extraction with the butyl chloride cannot be completely discounted. This would be compounded by the elevated temperature of the sequencing reaction and the relatively long exposure times to acid.

To test this, G protein labeled with [^{3}H]-palmitate was purified and subjected to a few cycles of Edman degradation in the absence of either PITC or HFBA. In the case where there was PITC delivery but no HFBA present, 10% of the counts (300 cpm out of a total of 3,000 cpm applied to the cup) were released in the first cycle. In the case where HFBA but not PITC was present, 33% of the starting radioactivity was released in the first cycle. Thus, the HFBA appears to have at least some effect on the stability of the fatty acid. The attempt to determine the site of acylation by sequencing was not further pursued.

As stated earlier, the threonine residues present at position 479 and 486 could serve as fatty acid acylation sites. To examine this further, [^{3}H]-palmitate labeled VSV was digested with a mixture of carboxypeptidases A and B in the presence of Triton X100

Figure 3.3.7: Effect of carboxypeptidase digestion on G protein labeled with [^{3}H]-palmitate.

VSV labeled with [^{35}S]-methionine or [^{3}H]-palmitate was digested with a mixture of carboxypeptidases A and B in the presence or in the absence of detergent as described in the legend to Figure 3.1.1 and analyzed on a 10% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV; lanes b and c, [^{35}S]-methionine labeled VSV treated with carboxypeptidase in the absence and presence of Triton X100, respectively; lane d, [^{3}H]-palmitate labeled VSV; lanes e and f, [^{3}H]-palmitate labeled VSV treated with carboxypeptidase in the absence and presence of Triton X100, respectively. The autoradiograph was overexposed in order to better visualize the tritium radioactivity.


a b c d e f
L
G
N
M

as described in Section 3.1. As demonstrated in Figure 3.3.7, carboxypeptidase digestion of VSV in the presence of detergent results in a molecular weight reduction of G protein of approximately 2,500D (lanes e and f). In the case of [^{3}H]-palmitate labeled VSV, the truncated G protein still retains the fatty acid label. This indicates that, within the sensitivity of the gel system used, the fatty acids that are linked to G are not present in the extreme carboxy terminal end of the molecule. Since the two threonine residues are located 17 and 10 amino acid residues from the carboxy terminal lysine, then if the fatty acids were present on these residues, a molecular weight reduction of G protein in the range of 1,500 to 2,000D would have resulted in the loss of palmitate radioactivity.

Thus, the fatty acids must be located in either the proposed transmembrane domain, that is, starting at Ser_{447} to Leu_{466} or in the external domain comprising the site of enzyme cleavage at Asn_{433} to the Lys_{446} residue at the external membrane boundary.

Comparative tryptic maps of [^{3}H]-palmitate labeled Fragment A and amino acid labeled Fragment A were undertaken. In this case, trypsin generated Fragment A was labeled with [^{3}H] valine, leucine, glycine, and phenylalanine. Analysis of the carboxy terminal sequence of G protein indicates that 5 tryptic peptides should be generated from tryptic digestion of Fragment A labeled with these amino acids. The major ones would be Asn_{433} - Lys_{446} which contains one each of leucine, valine, phenylalanine, and

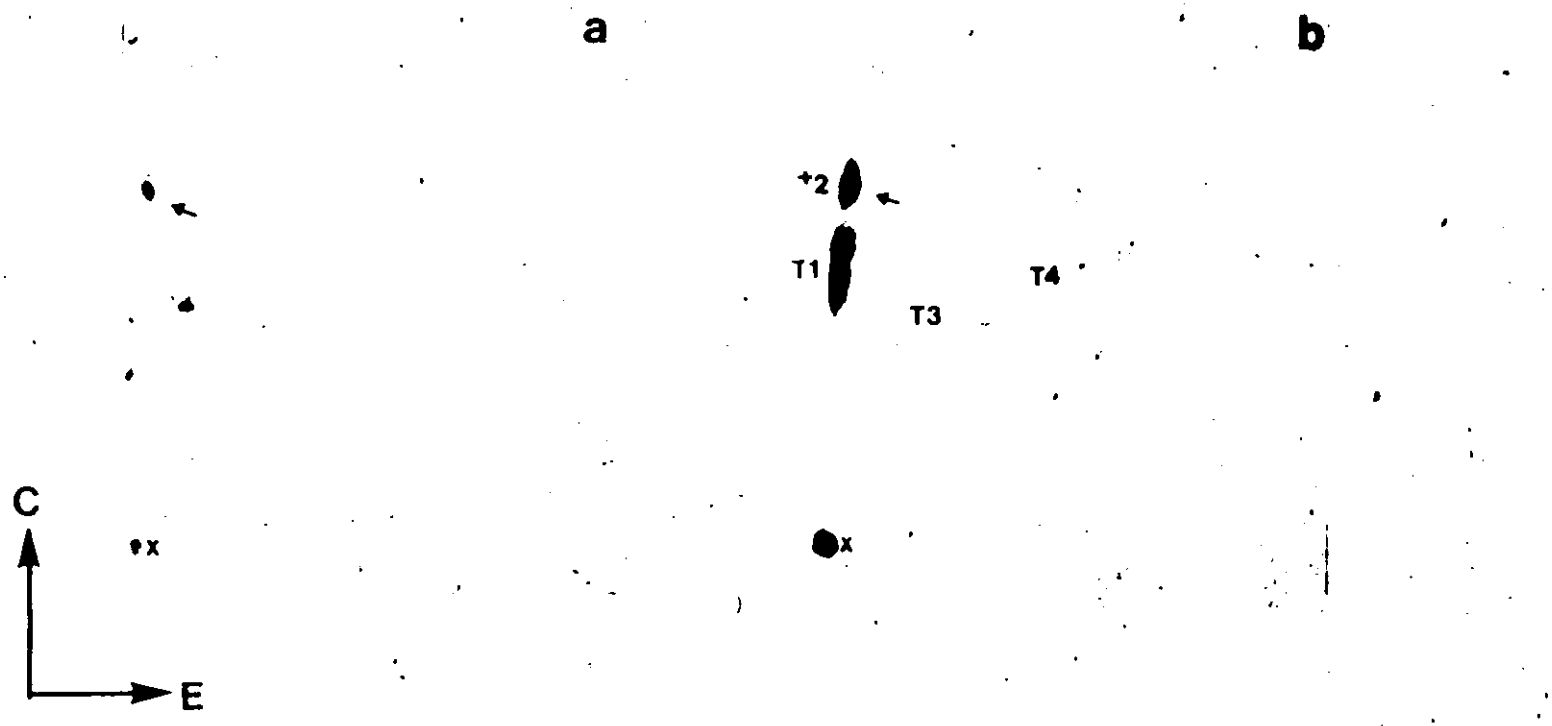
glycine ; Ser_{447} to Arg_{467} which contains four leucines, four phenylalanines, and one valine, and minor peptides comprising Val_{468} to Lys_{475}; Leu_{476} to Lys_{475}; and Leu_{493} to Lys_{495}. As demonstrated in Figure 3.3.8, tryptic digestion of Fragment A labeled with the above amino acids results in the generation of four distinct peptides as indicated in the figure. Peptide T1 because of its broad size, may contain more than one species.

Tryptic analysis of [^{3}H]-palmitate labeled Fragment A results in the generation of only one tryptic peptide. This peptide comigrates with peptide T2 from the amino acid labeled sample. Thus, the fatty acids are all contained in one tryptic peptide of G protein.

The identity of the tryptic peptide may be ascertained from a quantitation of the degree of radioactive incorporation in the tryptic peptides. The peptides with the highest degree of incorporation would be derived from the external portion and the membrane spanning portion. Both of these regions have serine residues that could act as fatty acid acceptors. When the specific activity of the isotopes used is taken into consideration (leucine, 138Ci/mmole, valine, 58 Ci/mmole, glycine, 12.3 Ci/mmole and phenylalanine, 56.2 Ci/mmole) then the ratio of radioactivity between the two peptides would be approximately 1:3 (external peptide/ transmembrane peptide). From a qualitative observation of the fingerprint shown in Figure 3.3.8, one would conclude that peptide T1 represents the membrane spanning domain. This conclusion is

Figure 3.3.8: Tryptic peptide analysis of Fragment A labeled with [^{3}H]-palmitate or with [^{3}H] amino acids.

Fragment A, labeled with [^{3}H]-palmitate or with a mixture of [^{3}H]-leucine, valine, phenylalanine, and glycine, was digested with trypsin and the tryptic peptides were analyzed on a thin layer plate as described in Methods. Panel a, Fragment A labeled with [^{3}H]-palmitate (4,000 cpm applied); panel b, Fragment A labeled with the above amino acids (18,000 cpm applied). The plate was sprayed with Enhance prior to autoradiography. The amino acid labeled peptides are numbered and the arrow indicates a peptide common to [^{3}H]-palmitate labeled and amino acid labeled Fragment A. The direction of electrophoresis (E) and chromatography (C) is indicated in the figure.


a
b
+2
T1
T3
T4
C
E

supported from the results obtained by eluting these peptides from the chromatogram and determining the radioactivity. In this case, 1,080 cpm were recovered from peptide T1 while approximately 400 cpm was recovered from peptide T2, giving a ratio of approximately 1:2.7 for T2/T1. Thus, peptide T2, with which the fatty acid containing tryptic peptide comigrates, appears to represent that domain of the glycoprotein just external to the bilayer. This would then appear to indicate that one or both of the serine residues located at positions 443 and 444 are acylated since there are no other hydroxy amino acids in this region.

This is supported by a study in which [^{3}H]-palmitate labeled G protein was cleaved at tryptophan residues (Schlesinger et al., 1981). NH_2-terminal amino acid analysis of an acyl containing peptide revealed the presence of phenylalanine at the amino terminus of this particular peptide. This would be consistent with fatty acid acylation at the serine residues mentioned above since these serine residues are flanked by tryptophan residues at positions 441 and 445. Cleavage at these tryptophan residues would generate a serine containing peptide having a phenylalanine at the amino terminus. Although this study is consistent with the fingerprint analysis presented here, it has not been confirmed.

The presence of fatty acid residues at one or both of these serines may also explain why proteolytic digestion of VSV does not cleave G protein at a point closer to the membrane surface. For instance, fatty acids at these residues may cause part of the

proposed external portion of the glycoprotein to interact more closely with the membrane surface, thereby protecting it against proteolytic degradation.

The major conclusion reported in this section is that the fatty acids covalently attached to G protein are localized in that region of the molecule that anchors it to the membrane. The fatty acids may thus augment the hydrophobic forces required for proper protein orientation and stabilization.

SECTION 3.4: INCORPORATION OF A PHOTOREACTIVE FATTY ACID INTO VSV

Studies concerning the topological organization, the distribution, and the nature of the hydrophobic interactions of membrane proteins and phospholipids provide valuable information regarding the stability, complexity, and function of biological membranes. In addition, in the simple model systems of enveloped viruses, knowledge of the spatial arrangements of the constituent membrane proteins and the nature of the lipid-protein, and protein-protein interactions provides a useful means of understanding and examining the processes involved in membrane assembly and viral morphogenesis.

In the case of VSV, numerous studies employing a variety of techniques have led to considerable advances in our understanding of the structural organization of the virus. Thus, the use of hydrophobic probes, chemical crosslinkers, and reversible bifunctional reagents have been utilized in examining interactions occurring in the membrane (Dubovi and Wagner, 1977; Mudd and Swanson, 1978; Zakowski and Wagner, 1980). These and other studies have, for the most part, demonstrated that G protein may exist in the membrane in oligomeric structures and that the M protein exists on the inner surface of the membrane in close proximity to the membrane surface as well as to other structural components of the virus.

These studies have provided important information concerning

the structure of the viral membrane, however, in each case exogenous probes were used. In addition, some of the methods require harsh conditions with the consequence that membrane perturbing events may possibly give rise to results not truly indicative of the physiological condition of the membrane (Patzer et al., 1979).

In recent years, chemical approaches utilizing fatty acids and phospholipids containing photoreactive nitrene or carbene precursors have been used as an effective means of examining hydrophobic interactions between phospholipids and proteins (Greenberg et al., 1976; Gupta et al., 1977, 1979a; Brunner et al., 1980; Quay et al., 1981). The use of these photolabile compounds has provided a novel and elegant approach for examining the precise spatial arrangement of proteins within the bilayer.

The demonstration that G protein contains covalently linked fatty acid residues in the membrane spanning domain makes it an ideal candidate for substitution with one of these photoreactive fatty acid derivatives. Dr. Gerber and his group in the Department of Biochemistry at McMaster is involved in the synthesis of various radioactive fatty acid analogues which contain the photosensitive diazirinophenoxy (DAP) group. In this respect, a collaborative effort was undertaken in which the tritiated synthetic photoreactive fatty acid ω-[9-^{3}H] diazirinophenoxy nonanoate ([^{3}H] DAP-nonanoate) was used to biosynthetically label VSV.

The use of the compound to label VSV would have two implications: (i) Since the phosopholipids of the virus would be

substituted with the photoreactive derivative, irradiation would make it possible to identify those proteins that are in intimate contact with the hydrophobic core of the membrane. (ii) If the fatty acid derivative could be used to acylate G protein in place of the normal fatty acids, a photoaffinity derivative of G protein would be produced. This would provide a novel method for examining nearest neighbour organization in the membrane through irradiation dependent crosslinking to proximate proteins. Thus, by labeling VSV with [^{3}H] DAP-nonanoate, it should be possible to examine phospholipid-protein as well as protein-protein interactions in a non perturbed system in which the probe is intrinsically incorporated during viral maturation and assembly.

Biosynthetic Utilization of [^{3}H] DAP-nonanoate by L Cells

Diazirine containing fatty acids and phospholipids have been utilized mainly in reconstituted artificial membrane systems and only recently have their biosynthetic potential been demonstrated in prokaryotic cells (Quay et al., 1981). There have been no reports that these compounds can be recognized and utilized for phospholipid synthesis by eukaryotic cells. This is, of course, an essential prerequisite towards the ultimate goal of incorporating the fatty acid derivative into G protein and the viral envelope. Thus, initial experiments were carried out with uninfected L cells to determine if [^{3}H] DAP-nonanoate could be incorporated into these cells and, more importantly, to ascertain whether the photoreactive group is

retained and remains capable of undergoing irradiation dependent activation.

The chemical structure of [^{3}H] DAP-nonanoate is shown in Figure 3.4.1 and the synthesis has been described (Leblanc et al., 1982). Irradiation results in the generation of carbene intermediates with the consequent formation of intramolecular crosslinks, which have been shown to generally occur by insertion into neighbouring C-H bonds (Gupta et al., 1979a, 1979b). This particular structure has advantages for biologically meaningful studies since the phenyl ring to which the diazirine is coupled allows for activation at 360 nm, a wavelength at which protein and membrane damage is minimal.

For the biosynthetic studies reported here, [^{3}H] DAP-nonanoate was usually solubilized in ethanol to a specific activity of 2 Ci/mmole unless noted otherwise. Immediately prior to use, the material was saponified with an equimolar amount of NaOH and added to the media such that the final alcohol concentration did not exceed 0.5 - 1%.

In initial experiments in which incorporation into L cells was monitored by counting aliquots of a cell suspension after various times of exposure to the radioactive medium, it was found that the inclusion of high concentration of serum in the medium resulted in much lower levels of incorporation. This was presumably due to the presence of large amounts of endogenous fatty acids in the serum. Subsequent experiments were carried out in the absence of serum

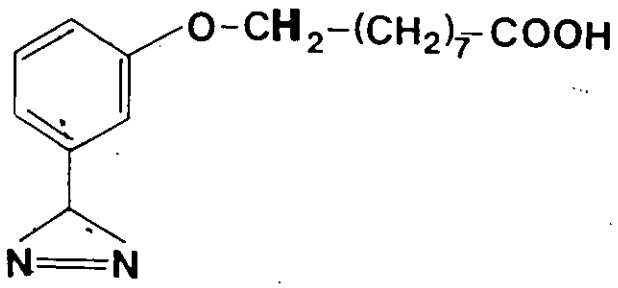


Figure 3.4.1: Structure of [^{3}H] Diazirinophenoxy nonanoate. The bold face letter indicates the position of the tritium label.

for relatively short labeling periods (2-4 hrs).

The incorporation of [^{3}H] DAP-nonanoate into the phosphatidylcholine and phosphatidylethanolamine fractions of L cells is shown in Figure 3.4.2a and 3.4.2b, respectively. As can be seen, incorporation of [^{3}H] DAP-nonanoate increased with time up to 4 hours and was comparable with the incorporation of [^{3}H]-palmitate. Most of the label present in the phospholipid fraction was found in the PC and PE fractions. This is in agreement with the phospholipid composition of the L cell membrane (Weinstein et al., 1969).

Irradiation Dependent Crosslinking of L Cell Phospholipids and Proteins

To demonstrate whether the photoreactive group is retained and remains capable of irradiation dependent excitation, L cells were labeled with [^{3}H] DAP-nonanoate for 4 hours and photolyzed for various lengths of time as described in the Methods section. After photolysis, the lipids were extracted and analyzed by TLC. As shown in Figure 3.4.3, photolysis resulted in extensive phospholipid crosslinking. Photolysis dependent crosslinking was shown to be essentially complete at the shortest time interval used (5 sec, Fig. 3.4.3b) and the products formed were found to be unaffected by further photolysis (Fig. 3.4.3 c-d).

SDS polyacrylamide gel analysis of proteins present in [^{3}H] DAP-nonanoate labeled L cells before and after photolysis is shown

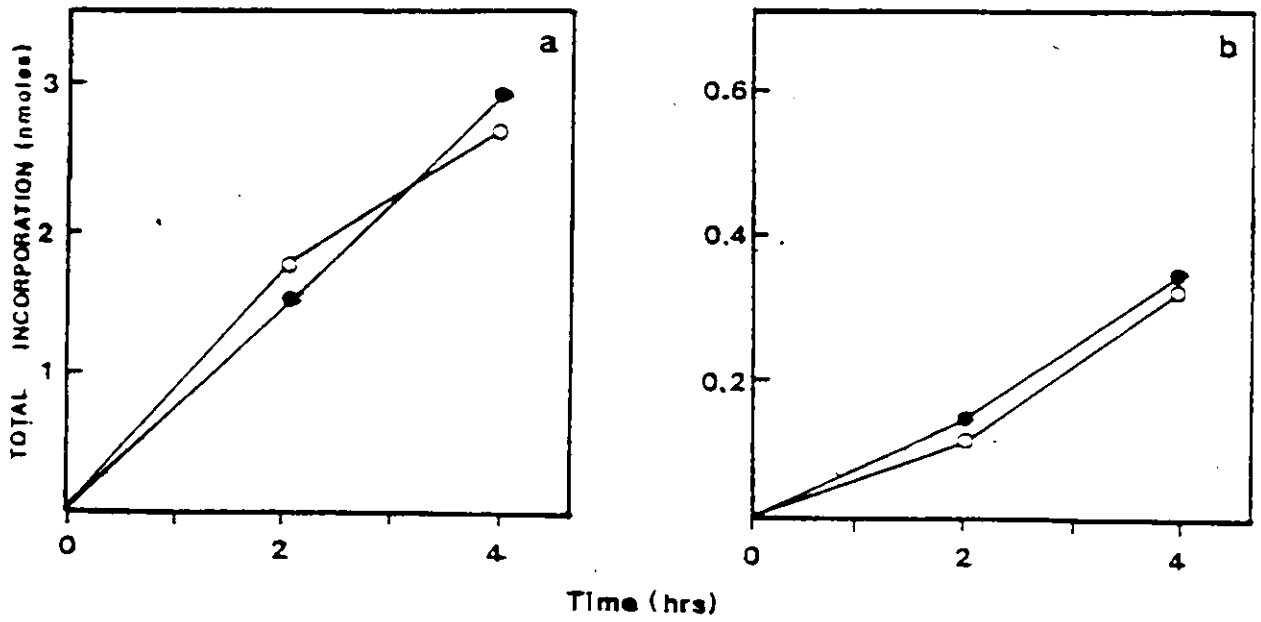


Figure 3.4.2: Incorporation of [^{3}H] DAP-nonanoate into L cell phospholipids.

L cells were labeled with [^{3}H] DAP-nonanoate (2 Ci/mmole) as described in Methods. At the times indicated, the cells (2.5×10^6) were harvested, washed 3 times with PBS, and the phospholipids were extracted. 10% of each extract was analyzed on a silica thin layer plate ($CHCl_3$/MeOH/H_2O, 65:25:4, v/v/v). The radioactive spots corresponding to each phospholipid were visualized by fluorography, recovered, and the radioactivity was determined by liquid scintillation counting. The figure shows the extent of incorporation into phosphatidylcholine (panel a) and phosphatidylethanolamine (panel b). A parallel analysis was carried out with L cells labeled with [^{3}H]-palmitate (17.2 Ci/mmole). (•) [^{3}H] DAP-nonanoate; (o) [^{3}H]-palmitate.

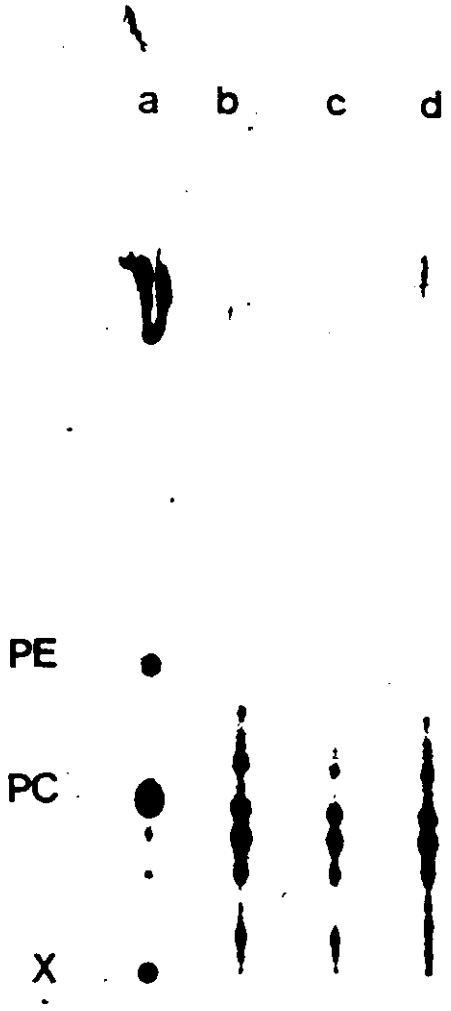


Figure 3.4.3: Phospholipid analysis of photolyzed L cells.

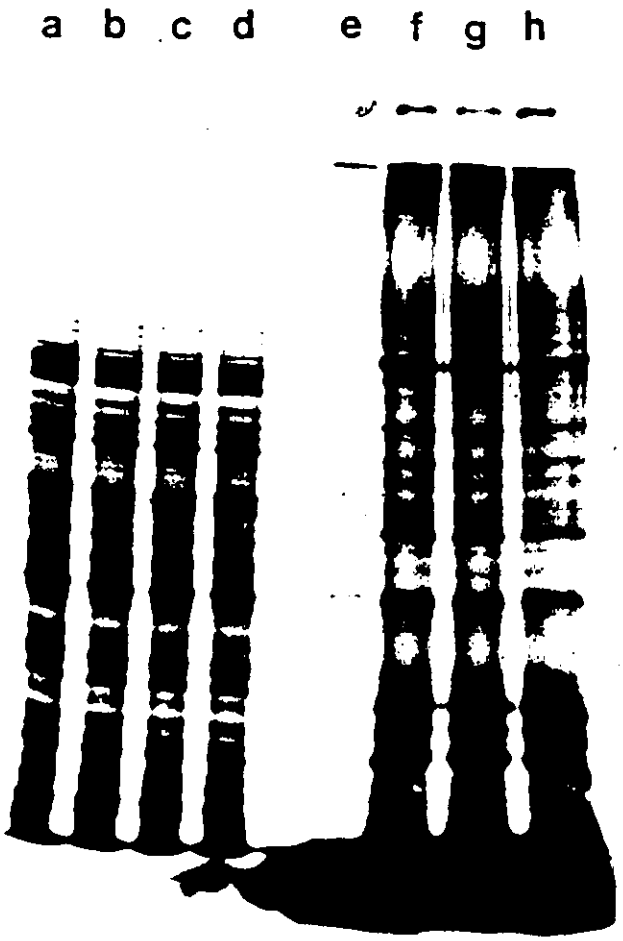
L cells were labeled for 4 hours with [^{3}H] DAP-nonanoate (160 mCi/mmole) after which time the cells were harvested, washed with PBS, and photolyzed for various lengths of time. Following irradiation, the cells were recovered and the phospholipids were extracted and analyzed by TLC as described in Figure 3.4.2. Lanes a, b, c, and d, represent the phospholipids extracted from L cells which were photolyzed for 0, 5, 30, and 60 seconds, respectively. Approximately 100,000 cpm was applied to each lane. (x) origin.

in Figure 3.4.4. A number of proteins were shown to be labeled prior to irradiation (Fig. 3.4.4 lane e). This labeling is not due to metabolic conversion of the [^{3}H] DAP-nonanoate to amino acids since the protein pattern is distinct from the pattern observed with an [^{35}S]-methionine labeled L cell sample analyzed in parallel (Fig. 3.4.3 lane a). A similar if not identical pattern would be expected if the labeling was due to amino acid conversion. These results are in agreement with those recently reported with CEF cells metabolically labeled with [^{3}H]-palmitate (Schlesinger et al., 1980) and supports the finding that a number of normal cellular proteins contain covalently linked fatty acids.

The effect of photolysis on [^{3}H] DAP-nonanoate labeled L cells is shown in Figure 3.4.4 lanes f-h. As can be seen, some of the proteins acylated prior to photolysis are further labeled after irradiation indicating that they are integral membrane proteins in close association with the hydrophobic core of the lipid bilayer. Not all of the proteins labeled prior to photolysis undergo irradiation dependent increases in radioactivity and others are only observed after photolysis. This would argue against the possibility that the products observed prior to photolysis represent products of partial photolysis by ambient light. More likely, the bands that are observed prior to irradiation which do not undergo irradiation dependent phospholipid-protein crosslinking represent soluble proteins which covalently interact with fatty acids since the total cell lysate was analyzed.

Figure 3.4.4: Photolysis of L cells labeled with [^{3}H] DAP-nonanoate.

L cells were labeled for 4 hours with [^{35}S]-methionine or [^{3}H] DAP-nonanoate. The washed cells (2 x 10^6) were suspended in 4 ml of PBS and 1 ml aliquots were photolyzed as described in Methods. The cells were recovered by centrifugation, washed with PBS, and incubated for 10 minutes at 37°C in a buffer containing 10 mM Tris-HCl (pH 7.6), 2 mM $CaCl_2$, 1 mM PMSF, phospholipase A_2 (100 ug/ml) and deoxyribonuclease (50 ug/ml). The reaction was terminated with the addition of electrophoresis sample buffer and the samples were analyzed on a 10% polyacrylamide gel. Lanes a, b, c, and d, [^{35}S]-methionine labeled L cells photolyzed for 0, 2, 4, and 10 seconds, respectively; lanes e, f, g, and h, [^{3}H] DAP-nonanoate labeled L cells photolyzed for 0, 2, 4, and 10 seconds, respectively.


a b c d e f g h

The possibility that the irradiation dependent crosslinking that is observed is nonspecific could also be ruled out since the photolysis pattern involves only a discrete reproducible subset of proteins that are different from the major proteins observed in [^{35}S]-methionine labeled samples. In addition, photolysis of control cells labeled with [^{35}S]-methionine under identical conditions showed no change in protein pattern, indicating the absence of any photolytic protein damage or photoinduced polymerization of proteins (Fig. 3.4.4 lanes a-d).

Since it has been shown that similar fatty acids do not label the headgroups of phospholipids (Gupta et al., 1979a) and that carbene precursors once placed in the membrane cannot be reduced by agents such as glutathione (Bayley and Knowles, 1978) and in fact would be rapidly scavanged by water if they reached the membrane surface, it is concluded that the proteins labeled upon photolysis are integral membrane proteins.

Thus, [^{3}H] DAP-nonanoate is readily recognized by mammalian cells and incorporated biosynthetically into phospholipids and proteins. The fatty acid derivative so incorporated undergoes irradiation dependent phospholipid-phospholipid and phospholipid-protein crosslinking and can thus be used for the identification of integral membrane proteins.

Biosynthetic Incorporation of [^{3}H] DAP-nonanoate into VSV

Having established the usefulness of the photoreactive fatty acid in the L cell system, the ability to incorporate [^{3}H].

DAP-nonanoate into VSV was examined. Incorporation of [^{3}H] DAP-nonanoate into VSV was roughly linear with increasing amounts of radioactivity and paralleled the incorporation observed with [^{3}H] palmitate (Fig. 3.4.5). The concentration of serum in the medium was found to have an effect on the level of incorporation. Highest incorporation was achieved when 2% serum was included in the medium. A high concentration of serum (10%) reduced total incorporation into virus as did the complete absence of serum. This was also the case when incorporation of [^{3}H]-palmitate into VSV was examined in a similar fashion. The inclusion of 10% serum reduced total incorporation into VSV by approximately 40% at the highest concentration of label used. This was presumably due to a large pool of endogenous fatty acids contributed by the serum. The reduction observed in the absence of serum is probably due to decreased levels of virus production.

In order to examine if [^{3}H] DAP-nonanoate had any effect on infectious virus production, plaque assays were performed on virus grown in the presence of different concentrations of the fatty acid. As demonstrated in Table IV, the presence of [^{3}H] DAP-nonanoate had little effect on total infectious virus produced. The highest concentration used (2.5×10^{-4}M) reduced infectious virus yield by approximately 35%. However, the [^{3}H] DAP-nonanoate used in the experiment had a specific activity of only 160 mCi/mmole. Subsequent preparations of material having a specific activity of 2 Ci/mmole reduced the actual amount of compound that was necessary for maximum

Figure 3.4.5: Effect of serum on the incorporation of [^{3}H] DAP-nonanoate into VSV.

VSV was grown in L cells in the presence of different concentrations of serum and various concentrations of [^{3}H] DAP-nonanoate or [^{3}H]-palmitate as indicated. Total incorporation was determined by purifying the virus in each case and determining the radioactivity present in the viral pellet. Panel a, [^{3}H] DAP-nonanoate (160 mCi/mmole) incorporation; Panel b, [^{3}H]-palmitate (12 Ci/mmole) incorporation. (●) no serum; (▲) 2% serum; (△) 10% serum.

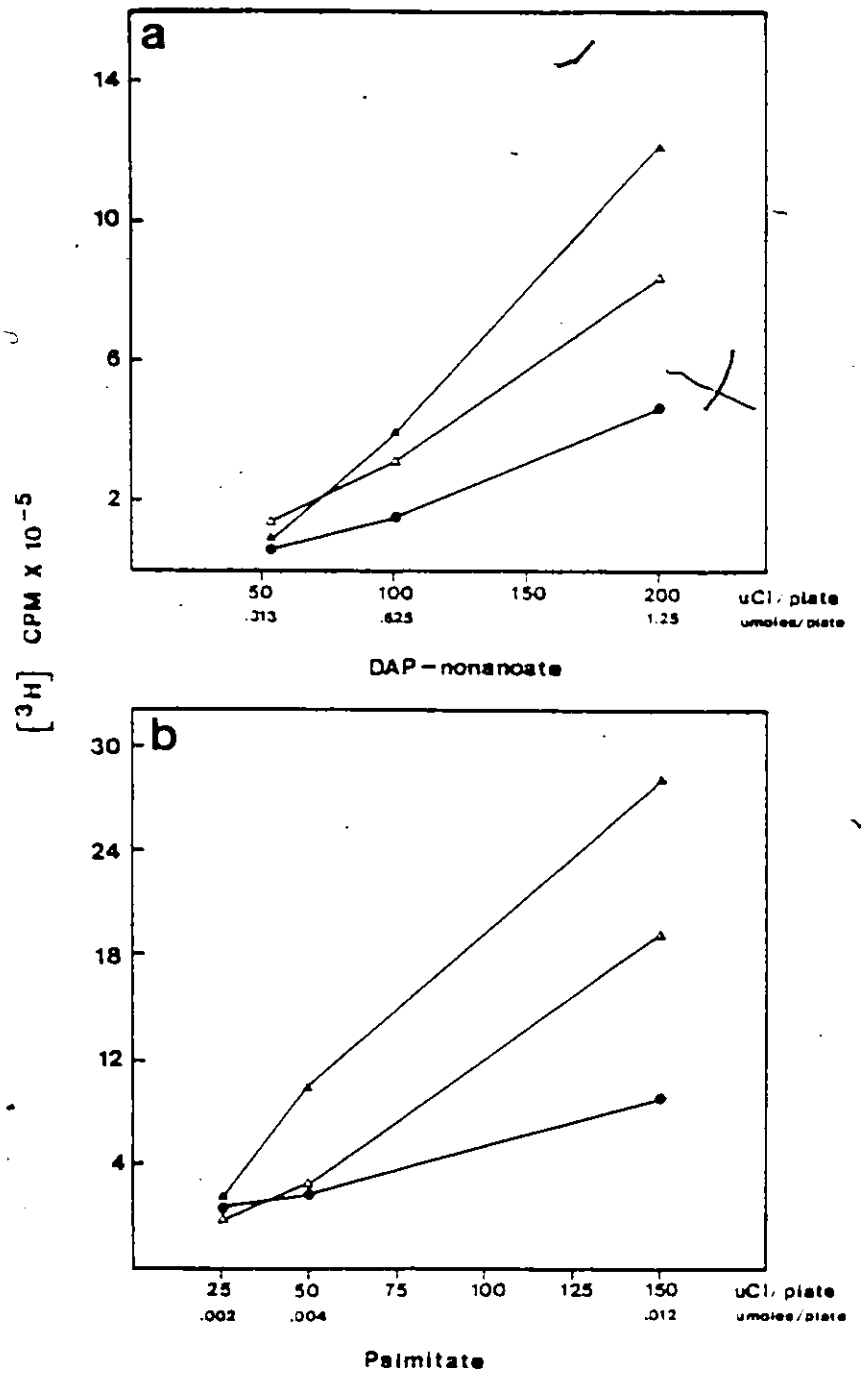

a
14
10
6
2
50
.313
100
.625
150
200
1.25
uCi/ plate
umoles/plate
DAP – nonanoate
[3H] CPM X 10-5
b
30
24
18
12
4
25
.002
50
.004
75
100
125
150
.012
uCi/ plate
umoles/plate
Palmitate

TABLE IV: EFFECT OF DAP-NONANOATE ON INFECTIOUS VIRUS PRODUCTION.

ADDITION	YIELD OF VIRUS PFU/ML x 10^{-8}
NONE	10.1
6.25×10^{-5} M DAP-NONANOATE	8.0
1.25×10^{-4} M DAP-NONANOATE	7.1
2.5×10^{-4} M DAP-NONANOATE	6.5

Plaque assays were performed on L cells using VSV which was grown in the presence of 10% NBCS and different concentrations of DAP-nonanoate as indicated in the table.

incorporation. Consequently, all viral labeling experiments were carried out in the presence of 2% NBCS and 40 uCi/ml of [^{3}H] DAP-nonanoate (2 Ci/mmole).

Phospholipid analysis of VSV labeled with [^{3}H] DAP-nonanoate showed that the major phospholipid classes were biosynthetically labeled with the fatty acid derivative (Fig. 3.4.6). Most of the label was found in the PC and PE fractions, in agreement with the phospholipid composition of VSV grown in L cells (Patzer et al., 1979).

Thus, the [^{3}H] DAP-nonanoate modified L cell plasma membrane is incorporated into budding virions during viral assembly.

Incorporation of [^{3}H] DAP-nonanoate into G Protein

When [^{3}H] DAP-nonanoate labeled VSV was analyzed on an SDS polyacrylamide gel, the only protein labeled was G protein (Fig. 3.4.7 lane b). The fact that none of the other viral proteins were detectably labeled precludes the possibility that the fatty acid was metabolically converted to amino acids prior to incorporation. Thus, it is possible to biosynthetically incorporate a fatty acid analogue in addition to or in place of the normal fatty acids that are found on G protein.

[^{3}H] DAP-nonanoate had no effect on the synthesis of the various viral proteins as determined by gel electrophoresis of a virus sample doubly labeled with [^{3}H] DAP-nonanoate and [^{35}S]-methionine (Fig. 3.4.7, lane c).

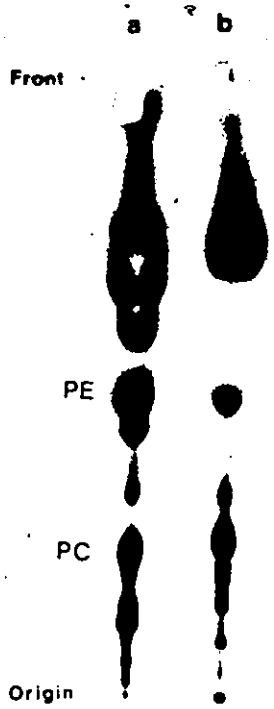


Figure 3.4.6: Phospholipid analysis of VSV labeled with [^{3}H] DAP-nonanoate.

Phospholipids were extracted from VSV labeled with [^{3}H] DAP-nonanoate or [^{3}H]-palmitate and analyzed by TLC as described in Methods. Lane a, phospholipids extracted from [^{3}H]-palmitate labeled VSV (150,000 cpm applied); lane b, phospholipids extracted from [^{3}H] DAP-nonanoate labeled VSV (100,000 cpm applied).

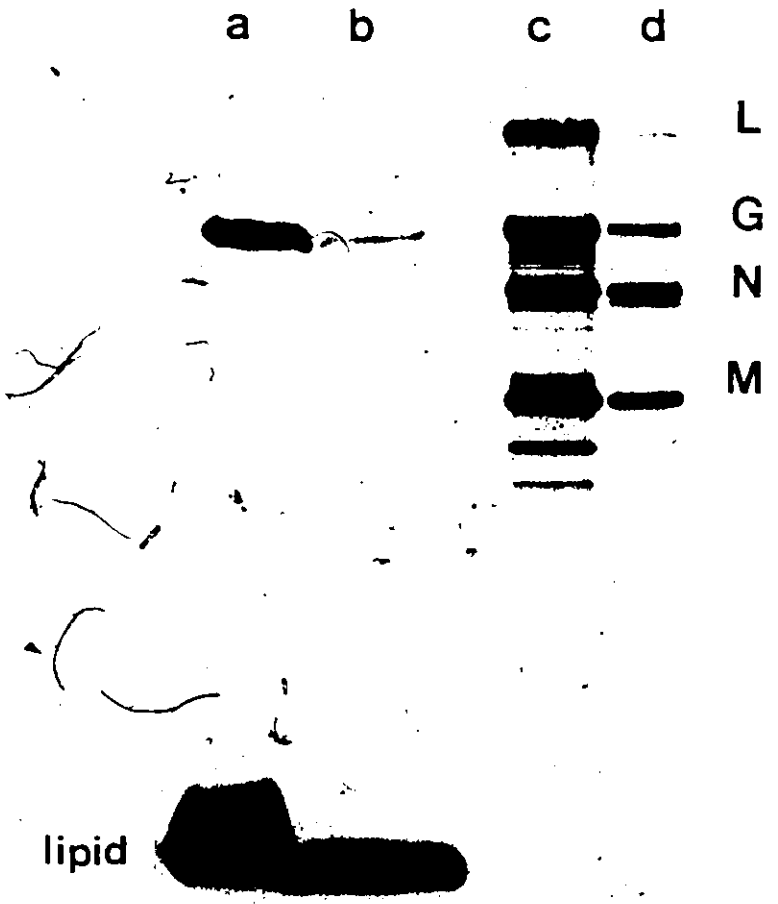


Figure 3.4.7: Incorporation of [^{3}H] DAP-nonanoate into VSV G protein.

VSV was labeled with [^{3}H] DAP-nonanoate, [^{3}H]-palmitate, or a mixture of [^{3}H] DAP-nonanoate and [^{35}S]-methionine. For dual labeled virus, infected L cells were prelabeled with [^{3}H] DAP-nonanoate and the medium was replaced four hours post infection with fresh medium containing 10 uCi/ml [^{35}S]-methionine and 50 uCi/ml [^{3}H] DAP-nonanoate. Lane a, [^{3}H]-palmitate labeled VSV; lane b, [^{3}H] DAP-nonanoate labeled VSV; lane c, VSV labeled with [^{3}H] DAP-nonanoate and [^{35}S]-methionine; lane d, [^{35}S]-methionine labeled VSV.

Quantitation of the [^{3}H] radioactivity present in G protein as a percentage of total incorporation into virus showed that approximately 1% of the total amount of radioactivity incorporated into VSV was present in G protein. The quantitation was performed by running [^{3}H] DAP-nonanoate labeled virus on a 10% polyacrylamide gel and determining the radioactivity present in gel slices (see Fig 3.4.12). In this case, of the approximately 110,000 cpm recovered from the individual gel slices, 1,000 cpm (0.9%) was recovered in the region of the gel corresponding to G protein. In comparison, approximately 2% of the [^{3}H]-palmitate incorporated into VSV was present in G protein.

Demonstration that [^{3}H] DAP-nonanoate is Covalently Attached To G Protein

The fact that the label on G protein was retained after polyacrylamide gel electrophoresis and subsequent fluorographic processing indicates that the label is very tightly complexed to G protein. In order to demonstrate whether the [^{3}H] DAP-nonanoate is covalently attached to G protein, VSV labeled with [^{35}S]-methionine or [^{3}H] fatty acids was separated on a polyacrylamide gel and treated with hydroxylamine. As shown in Figrue 3.4.8, NH_2OH treatment results in loss of label from both [^{3}H]-palmitate and [^{3}H] DAP-nonanoate labeled G protein. No detectable loss of radioactivity was observed when [^{35}S]-methionine labeled samples were treated in a similar manner (Fig. 3.4.8, lanes a and f).

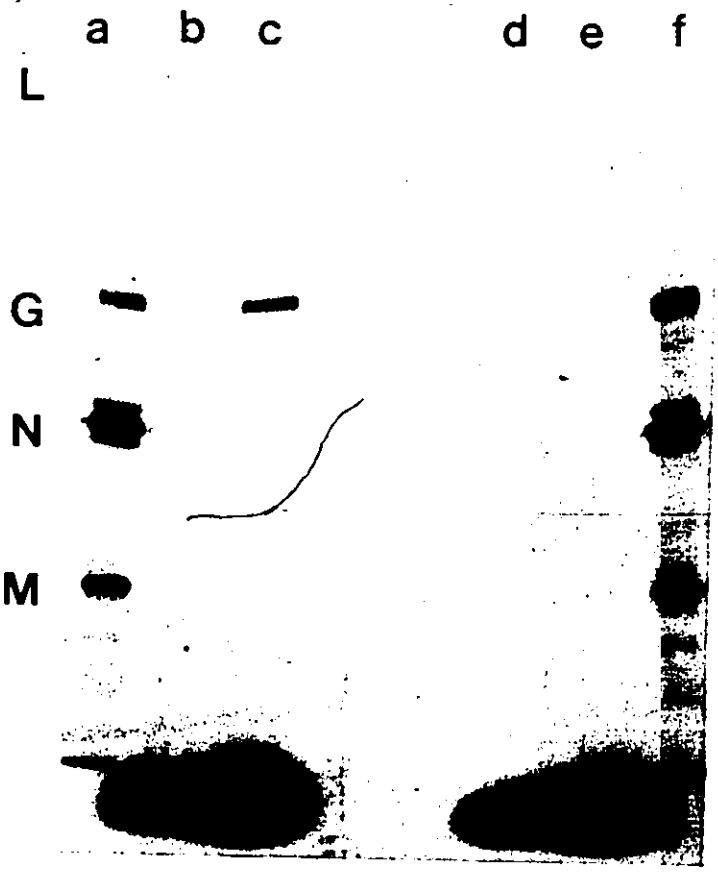


Figure 3.4.8: Effect of hydroxylamine on [^{3}H] DAP-nonanoate labeled VSV.

[^{35}S]-methionine, [^{3}H]-palmitate, and [^{3}H] DAP-nonanoate labeled VSV were run in duplicate on a 10% polyacrylamide gel. Following electrophoresis, half of the gel was directly processed by fluorography while the other half was treated with 1 M NH_2OH (pH 6.6 with NaOH) prior to fluorography. Lanes a, b, and c, untreated [^{35}S]-methionine, [^{3}H] DAP-nonanoate, and [^{3}H]-palmitate labeled VSV, respectively; lanes d, e, and f, NH_2OH treated [^{3}H] DAP-nonanoate, [^{3}H]-palmitate, and [^{35}S]-methionine labeled VSV, respectively.

The covalent nature of the linkage between [^{3}H] DAP-nonanoate and G protein was further confirmed by isolating G protein from labeled virus by detergent solubilization and delipidizing the G protein by exhaustive organic solvent extraction (Table V). This treatment resulted in the removal of all but 1.1% of the radioactivity found in the G protein supernatant. This value is in excellent agreement with the amount of radioactivity found to be present on G protein by SDS polyacrylamide gel electrophoresis.

The bound radioactivity was quantitatively released from G protein by transesterification with methanolic KOH. Of the 10,200 cpm remaining bound to G protein following organic solvent extractions, 98% became extractable with petroleum ether following transesterification.

The released fatty acid methyl ester derivatives were analyzed by HPLC (Fig. 3.4.9). As demonstrated in this figure, most of the radioactivity released from G protein migrated as the DAP-nonanoate methyl ester. However, significant amounts of longer chain length fatty acids were also present. Thirty percent of the applied radioactivity was recovered in the DAP-nonanoate methyl ester peak, while 16, 10, and 17% were recovered in the peaks corresponding to the methyl ester derivatives of DAP-undecanoate, tridecanoate, and pentadecanoate, respectively. Thus, it appears that a fraction of the [^{3}H] DAP-nonanoate is metabolically converted to longer chain length species prior to incorporation into G protein. This is in agreement with studies involving [^{3}H]-palmitate labeled VSV in

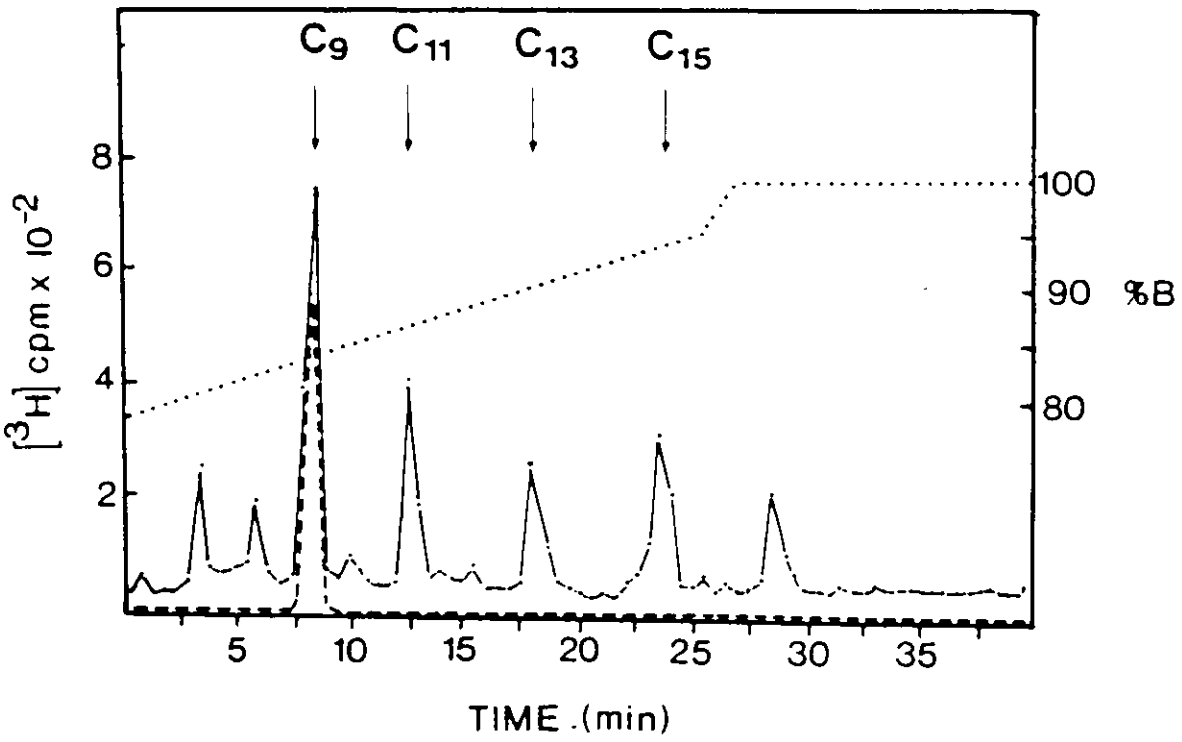
TABLE V: ORGANIC EXTRACTION OF G PROTEIN LABELED WITH [^{3}H] DAP-NONANOATE.

ORGANIC EXTRACTION	RADIOACTIVITY IN ORGANIC LAYER (CPM)	TOTAL RADIOACTIVITY RECOVERED %
METHANOL-CHLOROFORM (2:1)	8.84×10^5	93
METHANOL-CHLOROFORM-WATER (2:1:0.8)	1.96×10^4	2.06
METHANOL-CHLOROFORM-WATER (2:1:0.8)	8.0×10^3	0.84
METHANOL-CHLOROFORM-WATER (2:1:0.8)	1.5×10^3	0.16
SUM OF ORGANIC EXTRACTS	9.13×10^5	96.1
RESIDUAL G PROTEIN	1.02×10^4	1.1

[^{3}H] DAP-nonanoate labeled G protein was isolated from labeled VSV by extraction with Triton X100. One ml of the detergent supernatant (950,000 cpm) was subjected to organic solvent extractions as indicated in the table. 97.2% of the radioactivity (923,310 cpm) was accounted for after the various extractions.

Figure 3.4.9: HPLC analysis of radioactivity associated with [^{3}H] DAP-nonanoate labeled G protein.

[^{3}H] DAP-nonanoate labeled G protein was purified as described in Table V. The fatty acid residues remaining attached to G protein were transesterified by treatment with 0.1 N methanolic KOH for 20 minutes at 23°C. The resulting fatty acid methyl esters were recovered by extraction with petroleum ether, dried, suspended in methanol, and analyzed by reverse phase HPLC as described in Methods. The flow rate was 2 ml per minute using the gradient shown (A, water; B, methanol) and 0.5 minute fractions were collected and the radioactivity was determined. The solid line represents the fatty acid methyl ester derivatives released from [^{3}H] DAP-nonanoate labeled G protein by transesterification. The dashed line represents the starting methyl ester derivative of [^{3}H] DAP-nonanoate which was obtained from a separate analysis. The arrows represent the elution positions of radioactive markers corresponding to the methyl ester derivatives of DAP-nonanoate, undecanoate, tridecanoate, and pentadecanoate, in order of increased retention time, respectively. The radioactive markers were provided by P. Leblanc. Approximately 6,000 cpm of the organic solvent extractable material obtained from the transesterification of G protein was applied to the column.


C9
C11
C13
C15
8
6
4
2
[3H] cpm x 10-2
100
90
80
%B
5
10
15
20
25
30
35
TIME (min)

which small amounts of stearate and oleate in addition to palmitate were found on G protein (Schmidt and Schlesinger, 1979). The peak eluting at 28.5 min may represent a fatty acid of chain length longer than 15 carbons while the two small peaks eluting at 3 and 5.5 min have not been identified.

Localization of [^{3}H] DAP-nonanoate to the Membrane Interacting Domain of G

As demonstrated in Section 3.3, all the fatty acids covelently attached to G protein are present in the carboxy terminal membrane interacting domain. In order to determine if this is the case with the photoreactive fatty acid, [^{3}H] DAP-nonanoate labeled virus was digested with thermolysin to remove the external portions of G and the spikeless virions were isolated and analyzed on a polyacrylamide gel. The results shown in Figure 3.4.10 demonstrate that proteolytic digestion of [^{3}H] DAP-nonanoate labeled VSV results in the disappearance of G protein and the appearance of a new peptide corresponding to the membrane interacting domain of G. When expressed as a percentage of radioactivity recovered from the gel, it was found that all the [^{3}H] DAP-nonanoate label present in G is localized in the carboxy terminal domain.

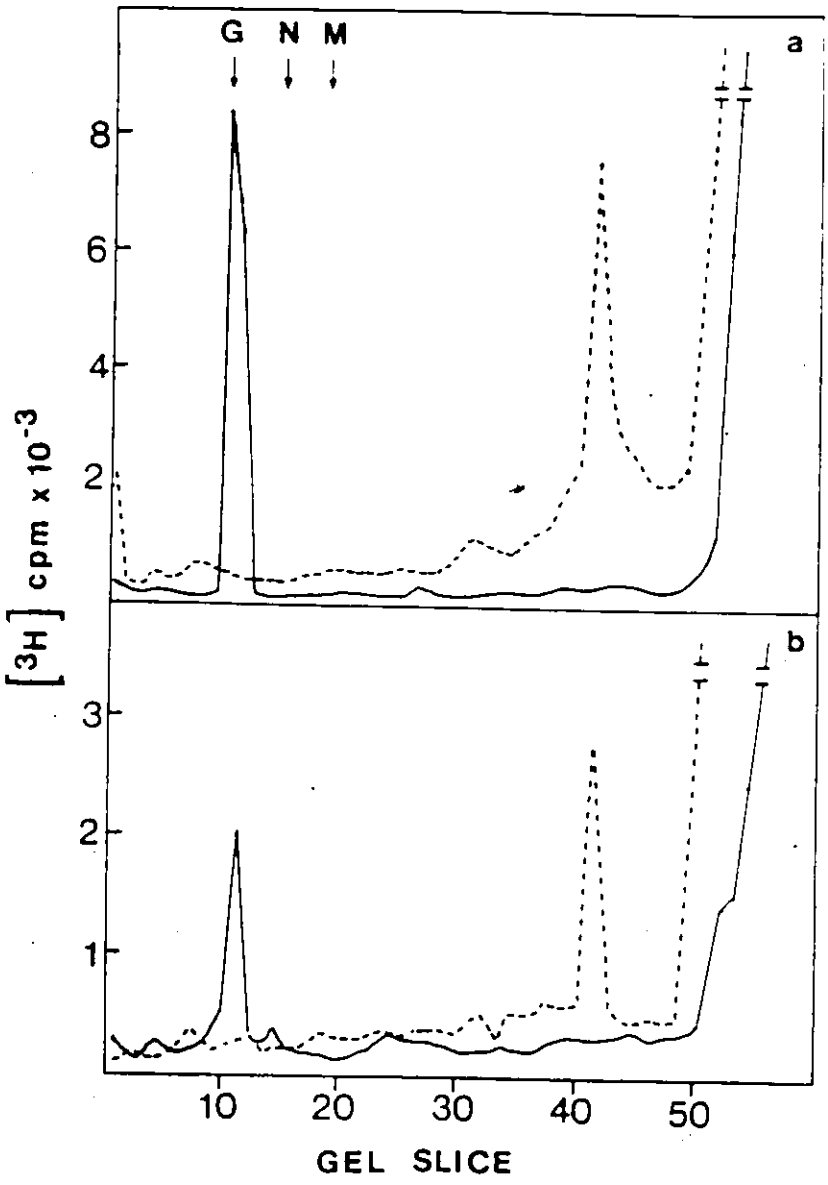
These results taken together indicate that [^{3}H] DAP-nonanoate behaves in an identical manner to palmitate that is normally acylated to G protein.

Photolysis of [^{3}H] DAP-nonanoate Labeled VSV

In order to examine the effect of irradiation on VSV labeled

Figure 3.4.10: Proteolytic digestion of VSV labeled with [^{3}H] DAP-nonanoate.

Labeled VSV was digested with thermolysin and the spikeless virions were purified and analyzed on a 17.5% polyacrylamide gel as described in Methods. Gel lanes, corresponding to the different samples, were cut and sectioned into 3 mm slices. Radioactivity was eluted from each slice and counted. Panel a, [^{3}H]-palmitate labeled VSV before (____) and after (---) treatment with thermolysin, respectively; panel b, [^{3}H] DAP-nonanoate labeled VSV before (____) and after (---) treatment with thermolysin, respectively. The position of the viral proteins, obtained from an [^{35}S]-methionine labeled VSV sample run on a separate lane, are indicated at the top of the figure.


G
N
M
a
b
8
6
4
2
3
2
1
[3H] cpm x 10-3
10
20
30
40
50
GEL SLICE

with the photoreactive fatty acid, VSV was photolyzed and analyzed on a 10% polyacrylamide gel. As shown in Figure 3.4.11, the most apparent result of photolysis of [^{3}H] DAP-nonanoate labeled VSV is a large increase in the labeling of the G protein. In order to quantitate this increase, labeled VSV was run on a polyacrylamide gel before and after photolysis, and the radioactivity was determined after fractionating the gel (Figure 3.4.12). In the case of non-photolyzed VSV, of the 110,000 cpm recovered from the gel, approximately 1000 cpm (0.9%) was recovered in the region of the gel corresponding to G protein. In contrast, of the 179,000 cpm recovered from the photolyzed sample, 16,400 cpm or 9.1% was present in the region corresponding to G protein. Thus, photolysis results in a 10 fold increase in the labeling of G protein. This increase is due to the formation of photolytic phospholipid-protein adducts resulting from the crosslinking of boundary lipids to G protein. If this is the case, then the increased label should be confined to the carboxy terminal membrane interacting domain of the glycoprotein, since this is the region that is in direct contact with the lipid bilayer. In order to test this, [^{3}H] DAP-nonanoate labeled VSV was photolyzed, digested with thermolysin, and analyzed on a polyacrylamide gel. As shown in Figure 3.4.13, proteolysis of photolyzed VSV demonstrates that the irradiation dependent increase in radioactivity in G protein is localized almost exclusively to the hydrophobic membrane anchoring domain.

Thus, the irradidation dependent phospholipid-protein

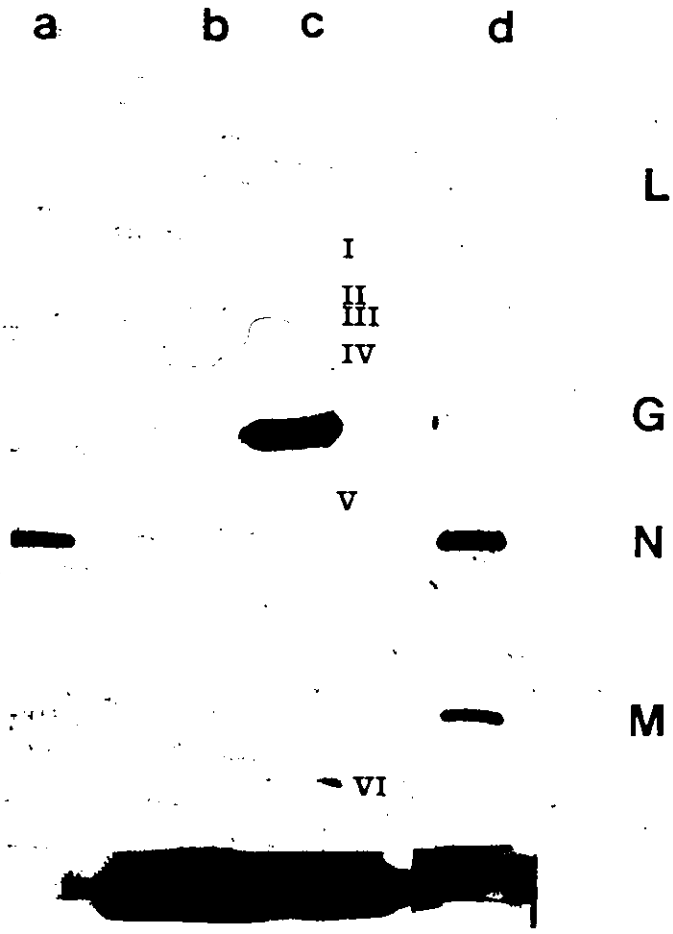
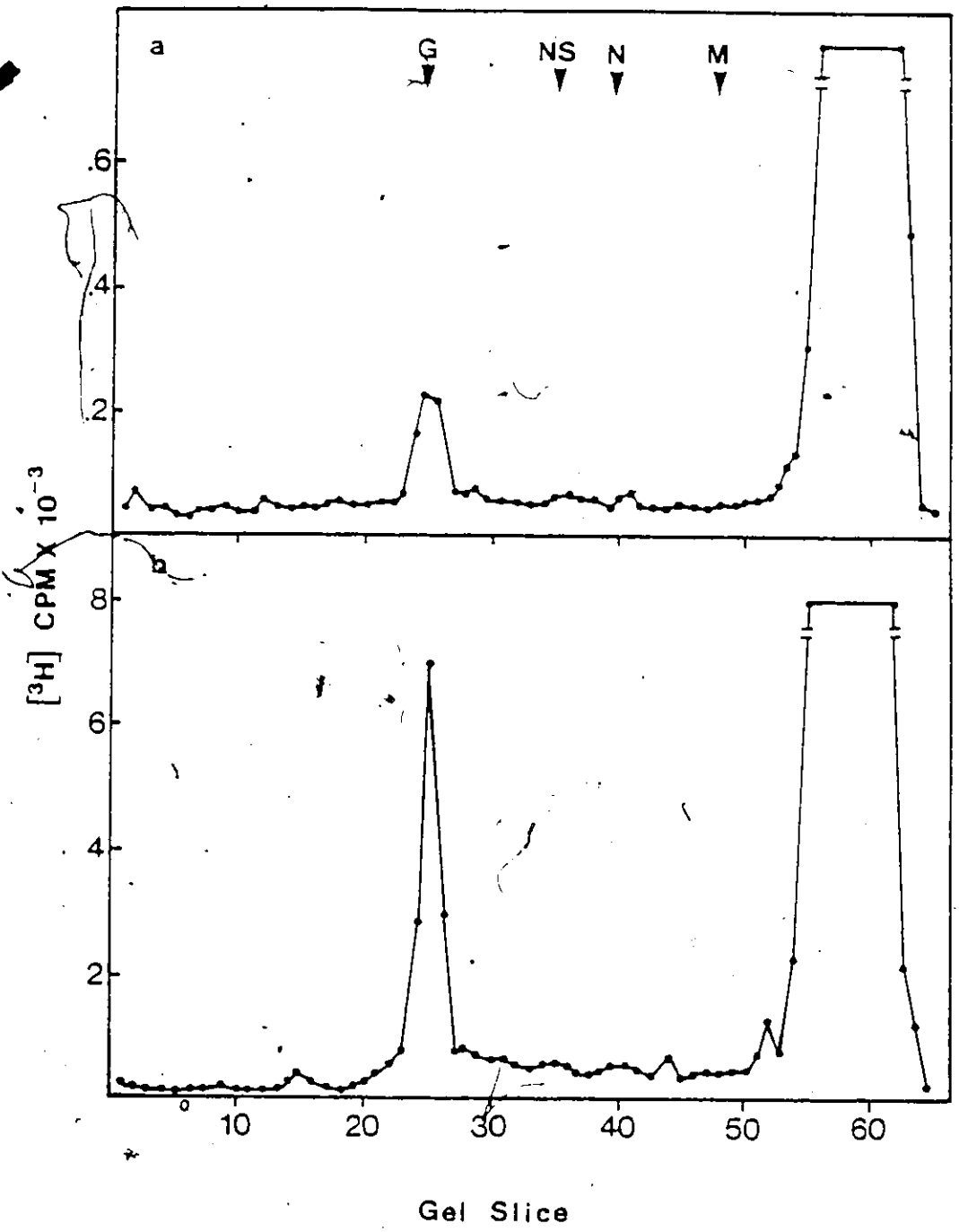


Figure 3.4.11: Photolysis of [^{3}H] DAP-nonanoate labeled VSV.

[^{3}H] DAP-nonanoate labeled VSV was diluted in buffer containing 10 mM Tris-HCl (pH 7.5), 100 mM NaCl and photolyzed for 30 seconds as described in Methods. The virus was recovered and analyzed on a 10% polyacrylamide gel. Lanes a, and d, [^{35}S]-methionine labeled VSV marker; lanes b and c, [^{3}H] DAP-nonanoate labeled VSV before and after photolysis, respectively. Bands I, II, III, IV, and VI, represent irradiation dependant products and have apparent molecular weights of 140,000D, 107,000D, 100,000D, 90,000D, 55,000D, and 24,000D, respectively.

Figure 3.4.12: Quantitation of the irradiation dependent increase in the labeling of G protein.

[^{3}H] DAP-nonanoate labeled VSV was split into two equal portions. One of the aliquots was photolyzed prior to analysis on a 10% polyacrylamide gel. Gel lanes corresponding to the samples were cut and sectioned into 2 mm slices. The radioactivity was eluted from each gel slice and counted. Lanes a and b, [^{3}H] DAP-nonanoate labeled VSV before and after photolysis, respectively. The positions of the viral proteins are indicated at the top of the figure.


a
G
NS
N
M
.6
.4
.2
b
8
6
4
2
0
10
20
30
40
50
60
[³H] CPM X 10⁻³
Gel Slice

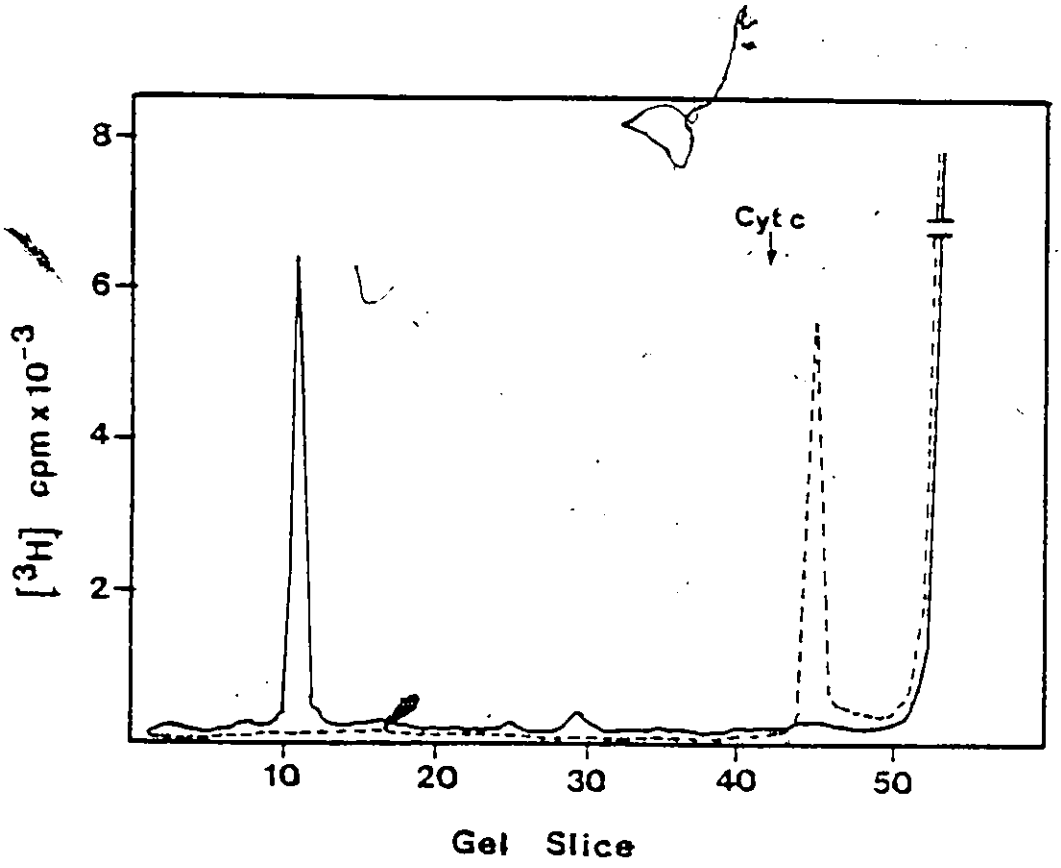


Figure 3.4.13: Proteolytic digestion of photolyzed [^{3}H] DAP-nonanoate labeled VSV.

[^{3}H] DAP-nonanoate labeled VSV was photolyzed, digested with thermolysin, and analyzed on a 17.5% polyacrylamide gel. Gel lanes were cut and sectioned into 3 mm slices and the radioactivity was eluted and counted. (——) before treatment with thermolysin; (---) after treatment with thermolysin. The position of cytochrome c, which was used as a molecular weight marker, is indicated. Of the total radioactivity recovered, 6.5% was present in the G protein peak in the photolyzed sample prior to enzymatic digestion while approximately 8% of the total recovered radioactivity was present in the membrane anchoring fragment after digestion.

crosslinking is confined to the hydrophobic core of the membrane and accurately identifies G protein as an integral membrane protein. This is readily apparent from the fact that none of the other viral proteins, especially the M protein, were detectably labeled following photolysis. This is particularly important in the case of M protein since M has been identified as a membrane protein by various criteria. The fact that the M protein is not labeled to any extent following irradiation indicates that it does not penetrate the lipid bilayer to a significant extent. Thus, photolysis of [^{3}H] DAP-nonanoate labeled VSV identifies only those membrane proteins that are in intimate contact with the hydrophobic core of the lipid bilayer.

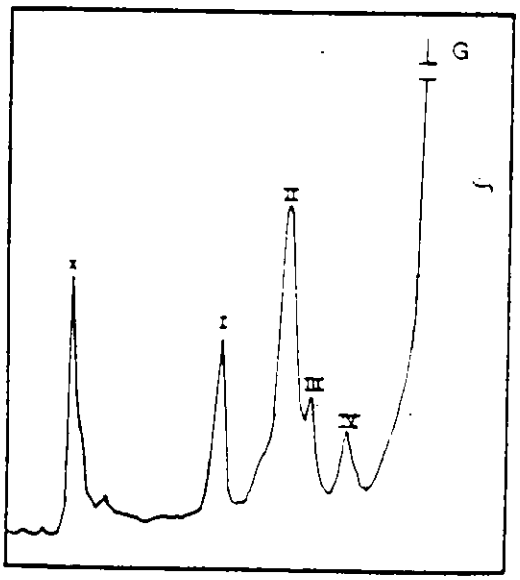
The biosynthetic incorporation of [^{3}H] DAP-nonanoate into the G protein affords a nearest neighbour probe of this integral membrane protein since the carbene generated upon photolysis will react with all groups within its vicinity.

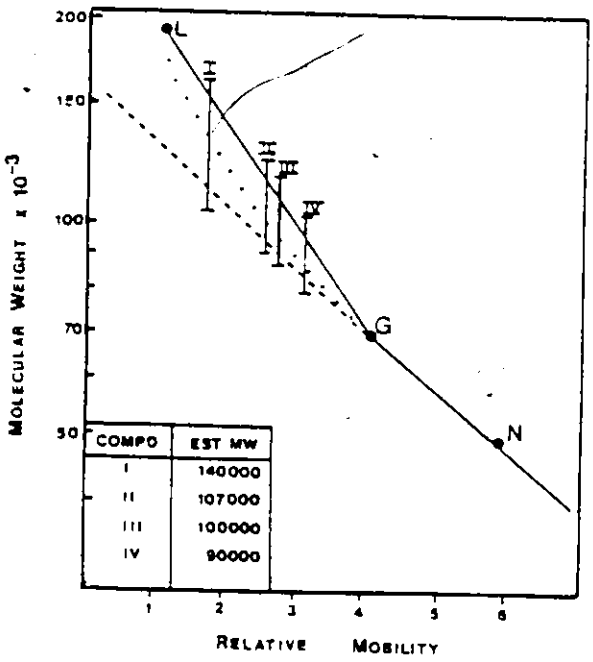
In addition to the irradiation dependent increase in labeling of G protein, photolysis of [^{3}H] DAP-nonanoate labeled VSV results in the appearance of new protein species of molecular weight higher than G as well as some species that exhibit a lower molecular weight (Fig. 3.4.11).

The photoinduced higher molecular weight species are designated I, II, III, and IV and are shown more clearly in Figure 3.4.14 which represents a densitometric scan of this region of the gel shown in Figure 3.4.11, lane c. The compounds I, II, III, and IV

Figure 3.4.14: Densitometric scan of the high molecular weight proteins present in photolyzed [^{3}H] DAP-nonanoate labeled VSV.

The high molecular weight region of the autoradiogram shown in Figure 3.4.11 was scanned with a Joyce-Loebl microdensitometer. The major irradiation dependant peaks and their estimated apparant molecular weights are indicated in the bottom panel. (x) indicates a deflection arising from the interface between the stacking and separating gel.


G


200
150
100
70
50
L
G
N
MOLECULAR WEIGHT x 10⁻³
RELATIVE MOBILITY
1
2
3
4
5
6
COMPO
EST MW
I
140000
II
107000
III
100000
IV
90000

were shown to have molecular weights of 140,000, 107,000, 100,000, and 90,000 D, respectively, based on their electrophoretic mobilities. The appearance of these protein species is entirely dependent on the presence of the photoreactive fatty acid since photolysis of [^{3}H]-palmitate labeled VSV did not result in an increase in labeling of G protein or the appearance of new protein species (Fig. 3.4.15, lanes b and c). Photolysis of [^{35}S]-methionine labeled VSV resulted in a stable protein pattern except for the appearance of a minor species below L protein which had an apparent molecular weight of 175,000 D (Fig. 3.4.15, lanes d and e). The origin of this species is not known but its appearance was variable depending on the length of photolysis. It is not due to photopolymerization of G protein since the species is not produced upon photolysis of [^{3}H] palmitate labeled virus. It may be the result of a variable degradation of L protein since given the large size of L protein (190,000 D) it would be the most susceptable to photolytic damage.

Identification of Photoinduced Protein Species

The photoinduced high molecular weight species produced by irradiation of [^{3}H] DAP-nonanoate labeled VSV could be the result of two things. They may be due to the presence of minor host derived protein species present in the viral membrane that become labeled after photolysis by phospholipid-protein crosslinking. For instance, Hecht and Summers have demonstrated the existence of H2 histocompatability antigens of mouse cells in highly purified preparations of VSV (Hecht and Summers, 1972, 1976). This may

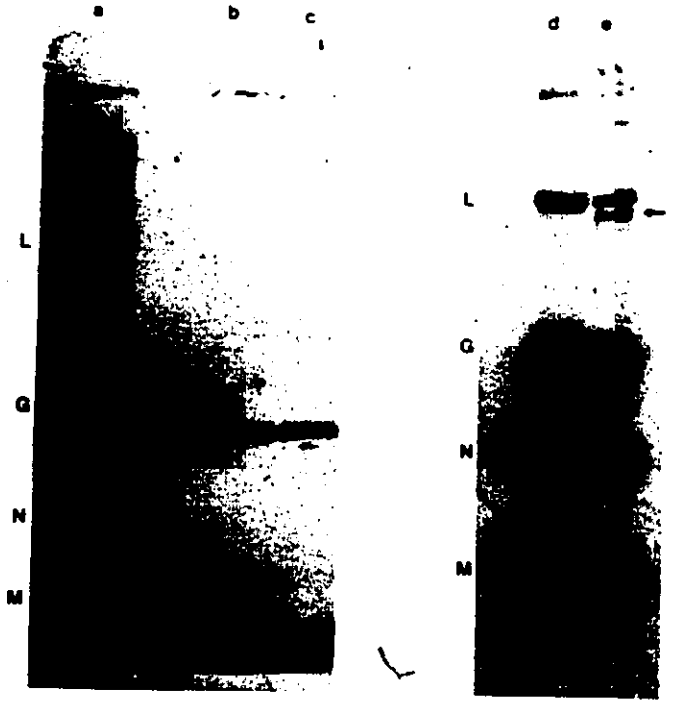


Figure 3.4.15: Photolysis of [^{3}H]-palmitate and [^{35}S]-methionine labeled VSV.

Labeled virus was photolyzed as described in Methods and analyzed on a 10% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV marker; lanes b and c, [^{3}H]-palmitate labeled VSV before and after photolysis, respectively; lanes d and e, [^{35}S]-methionine labeled VSV before and after photolysis, respectively. The arrow in lane e indicates a new species resulting from the irradiation which has an apparent molecular weight of approximately 175,000D and is only present in the [^{35}S]-methionine labeled sample.

explain the appearance of protein species below G protein following photolysis (labeled V in FIg. 3.4.11). More recently, Lodish and Porter have used surface labeling techniques to detect minor species of host derived proteins in the membrane of purified VSV (Lodish and Porter, 1980). The major species detected had molecular weights of 110,000 and 20,000 D, respectively. This is consistent with the species labeled VI in Figure 3.4.11 observed after photolysis of VSV. This protein has a molecular weight of 24000 D based on its electrophoretic mobility. If this is the case, then this protein is likely to be transmembranal given the large irradiation dependent increase in labeling that it exhibits.

The second possibility is that the high molecular weight species represent products of protein-protein crosslinking mediated by the photoreactive fatty acid covalently attached to G protein. Based on apparent molecular weights, the compound designated I is consistent with a G-G homodimer while compounds II and III have molecular weights more consistent with a G-M heterodimer.

One method of examining whether in fact these species are due to protein-protein crosslinking mediated by G protein is to immunoprecipitate photolyzed VSV with specific antibodies raised against the various viral proteins.

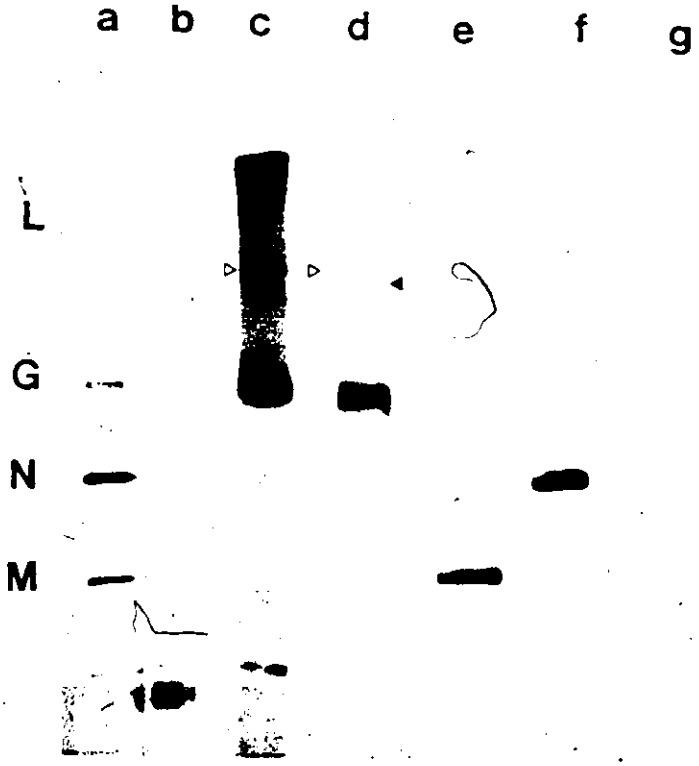
Initial attempts were unsuccessful due to the small amount of label present in the photoinduced species. Recently, an ultrasensitive immunoblotting technique has been developed to detect

proteins present in trace amounts (Towbin et al., 1979). The technique involves fractionating proteins of interest on a polyacrylamide gel and then transferring the proteins onto a nitrocellulose membrane. The immobilized proteins are then reacted with a specific antibody and the antibody-protein complex is then detected by reaction with radioiodinated Protein A. This technique allows for the detection of as little as 100 pg of protein.

This procedure was used to characterize the species produced following photolysis of VSV. Thus, [^{3}H] DAP-nonanoate labeled VSV was photolyzed, run on a polyacrylamide gel, transferred onto nitrocellulose, and immunoblotted with specific viral antibodies followed by reaction with [^{125}I] Protein A. The results are shown in Figure 3.4.16. In the case of antibody directed against G protein, two higher molecular weight species, in addition to G protein, were detected in the photolyzed sample (Fig. 3.4.16, lane d). These two species have molecular weights of 160,000 and 140,000 D, respectively. The 160,000 D species was present in the nonphotolyzed sample (lane c), thus the 140,000 D species is the only bonafide photoinduced compound reactive with anti-G antibody. This species corresponds in molecular weight to compound I of Figure 3.4.11. No other proteins were detected with anti-G antibody. Thus, the other photoinduced species do not have a G protein component. The fact that the 24,000 D species was not detected by the radioiodination suggests that it does not represent a breakdown product of the G protein.

Figure 3.4.16: Immunoblotting of [^{3}H] DAP-nonanoate labeled VSV.

[^{3}H] DAP-nonanoate labeled VSV was photolyzed, split into 4 equal portions, and run in separate wells on a 10% polyacrylamide gel along with an equivalent amount of non-photolyzed sample and an [^{35}S]-methionine labeled VSV marker. Following electrophoresis, the proteins were electrophoretically transferred onto a nitrocellulose membrane. The membrane was cut into lanes corresponding to the different samples and reacted with various VSV specific antibodies. This was followed by reaction with [^{125}I] protein A as described in Methods. Shown here is an autoradiogram of the immunoblots as visualized with the aid of an intensifying screen. Lane a, [^{35}S]-methionine labeled VSV marker; lane b, non photolyzed [^{3}H] DAP-nonanoate labeled VSV; lanes d, e, f, and g, photolyzed [^{3}H] DAP-nonanoate labeled VSV reacted with specific antibodies directed against VSV G, N, M proteins, and normal serum, respectively. Lanes a and b were exposed for longer times in order to visualize the [^{3}H] and [^{35}S] radioactivity. Lane c was exposed for 30 minutes while lanes d-g were exposed for 15 minutes. The open arrowhead in lanes c and d represents a G protein antibody reacting species which is present in the virus prior to photolysis. The filled in arrowhead represents an irradiation specific species present only in the photolyzed sample and reactive only with anti-G antibody. This species has an apparant molecular weight of approximately 140,000D.


a b c d e f g
L
G
N
M

Only the N and M proteins were detected when the photolyzed sample was reacted with antibody directed against N or M protein, respectively (Fig. 3.4.16, lanes e and f). Therefore, none of the new protein species produced by photolysis contain an N or an M component.

This reinforces the assignment of the 140,000 D photoinduced species as a bonifide G-G homodimer and rules out the possibility that compounds II and III represent G-M heterodimers. Thus, at least a portion of the G protein molecules exist in the membrane in close enough proximity to each other to form dimers upon photolysis. This dimer formation is likely the result of stable protein-protein interaction rather than due to collisional interactions. Studies examining the rigidity of the viral membrane and the fluidity of G protein in the lipid environment have invariably demonstrated that the viral membrane is considerably more rigid than the plasma membrane (Aitstiel and Landsberger, 1981; Landsberger and Aitstiel, 1980; Landsberger and Compans, 1976). As a consequence, the lateral mobility of G protein is much reduced.

The immunoblotting experiment also indicates that the remaining photoinduced species (compounds II, III, IV, V, and VI) represent minor species of host proteins present in the viral preparation since they were not reactive with any of the specific antibodies tested.

The results put forth in this section have demonstrated the usefulness of biosynthetically utilizing photoreactive fatty acids

as a general means of examining membrane topology and, in particular, as a unique way of identifying integral membrane proteins. This was readily apparent in the case of VSV where the nature of the membrane interaction of the G and M proteins were clearly demonstrated. In this case, photolysis of [^{3}H] DAP-nonanoate labeled VSV resulted in extensive phospholipid crosslinking to G protein, an integral membrane protein, however, no crosslinking occurred with the M protein. In addition, the demonstration that a photoreactive fatty acid could be covalently attached to G protein and used as a photoaffinity label of this protein has provided insight on the nature of some of the protein-protein interactions occurring in the viral envelope.

SECTION 3.5: INCORPORATION OF THE MATRIX PROTEIN M INTO PHOSPHOLIPID VESICLES AND ITS ASSOCIATION WITH THE NUCLEOCAPSID PROTEIN N

The non-glycosylated M protein of VSV is believed to be localized on the inner surface of the viral membrane (Cartwright et al., 1969; Wagner et al., 1970) and to play an important role in viral maturation through its interaction with the glycoprotein modified lipid bilayer and the RNP core (Shnitzer and Lodish, 1979).

Fluorescent photobleaching studies have shown that M protein modulates the surface mobility of G protein (Reidler et al., 1981; Johnson et al., 1981) while other studies have demonstrated that M protein can interact with the nucleocapsid (Mancarella and Lenard, 1981; Wilson and Lenard, 1981). M protein appears to play a functional role in maintaining the nucleocapsid in a condensed form. In the absence of M protein, the nucleocapsid was shown to exist in an extended form while in the presence of M, the nucleocapsids were found to more closely resemble the native RNP (Newcomb and Brown, 1981; Newcomb et al., 1981).

The close association of the M protein with the viral or cellular lipid has been demonstrated with the use of hydrophobic probes (Pepinsky and Vogt, 1979; Zakowski and Wager, 1980) and cell fractionation studies (Morrison and McQuain, 1978). However, no direct evidence exists that M protein interacts with the viral membrane. The studies with the photoreactive probe described in the previous section indicate that M protein does not penetrate the

envelope to an appreciable extent, if at all. Since the M protein is protected from proteolytic digestion because of its internal location, it is not possible to carry out the type of experiments that were done with G protein. Therefore, we decided to examine the affinity of M protein for membranes by attempting to reconstitute the purified protein into artificial phospholipid vesicles.

These artificial membrane systems seem to reflect the true nature of protein-lipid interaction, at least in the case of viral glycoproteins. For instance, studies with the glycoproteins of VSV (Petri and Wagner, 1979, 1980; Miller et al., 1980) Semliki Forest Virus (Helenius et al., 1977) and influenza virus (Huang et al., 1979; Hsu et al., 1979) have demonstrated that the glycoproteins insert into articifical vesicles in a manner analogous to that found in the normal viral envelope. In addition, reconstituted G protein lipid vesicles were able to inhibit VSV infection (Miller et al., 1980; Miller and Lenard, 1980) and liposomes containing glycoproteins of influenza virus were shown to undergo fusion with cell membranes (Huang et al., 1980).

Association of M Protein with Artificial Phospholipid Vesicles

Recently, the nature of the lipid-protein interaction of the M protein of influenza virus has been examined by incorporating the purified protein into lipid vesicles (Gregoriades, 1980; Gregoriades and Frangione, 1981). A similar approach was used to determine if M protein of VSV has an affinity for lipid vesicles.

The VSV M protein is difficult to study because of

aggregation and solubility problems. The M protein used in this study was obtained directly from polyacrylamide gels or purified from VSV by detergent extraction and gel permeation chromatography as described in Methods. In both cases, the protein was judged to be at least 90% pure by gel electrophoresis. Figure 3.5.1 shows the purified M protein that was obtained by elution from polyacrylamide gels. The protein was extensively washed with ethanol-ether and acetone to remove SDS and finally suspended in buffer containing 1% deoxycholate. The protein was insoluble in aqueous buffers in the absence of detergent.

Nondenatured M protein was obtained by extracting VSV with octyl glucoside in the presence of 0.25M NaCl. The supernatant was fractionated on an Ultrogel AcA-54 column and the M protein collected (Fig. 3.5.2). Octyl glucoside was used for the detergent disruption of VSV instead of Triton X100 since it is much easier to remove by dialysis (Furth, 1980).

To investigate the ability of purified M protein to interact with lipid vesicles, the protein was mixed with phosphatidylcholine in the presence of 1% DOC and vesicles were formed by removing the detergent by dialysis. The association of protein with lipid was determined by floating the mixture through a sucrose gradient. During equilibrium centrifugation, intact lipid vesicles will band in the sucrose gradient according to their densities. If proteins are tightly associated with the lipid, they will float upwards, otherwise they will remain at the bottom of the tube (Klenk et al., 1974).

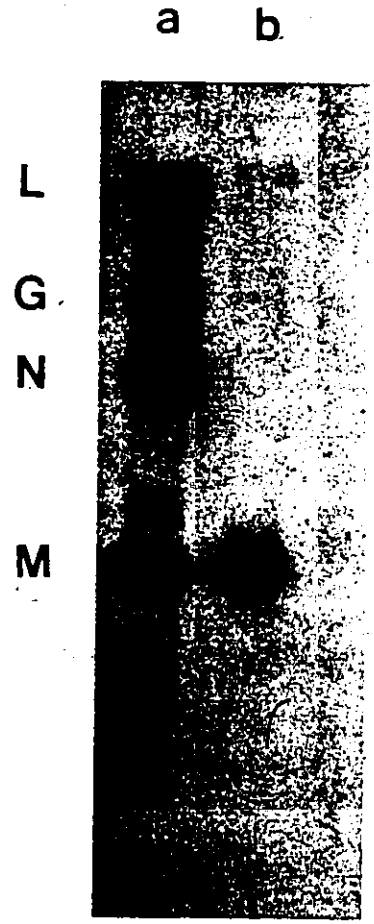


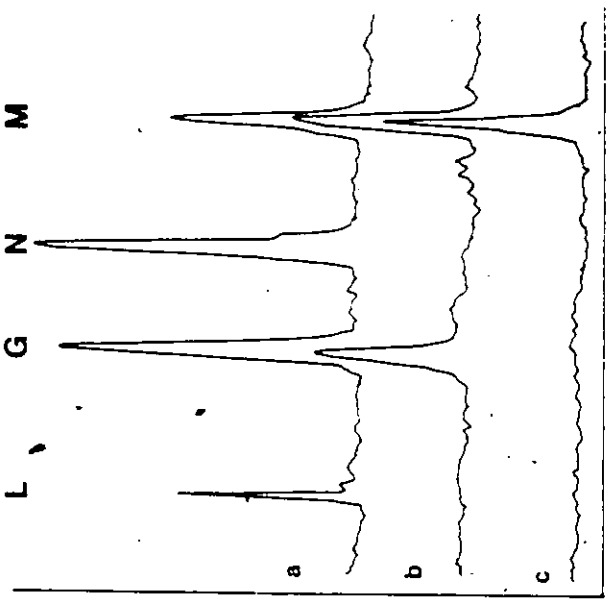
Figure 3.5.1: Purification of M protein by polyacrylamide gel electrophoresis.

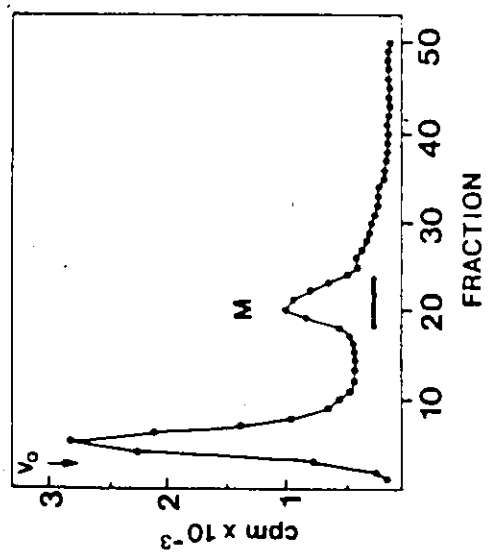
VSV specific proteins were separated on a 10% polyacrylamide gel and the region of the gel was excised and the protein eluted and purified as described in Methods. The purity of the protein was determined by running an aliquot on a polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV; lane b, purified M protein labeled with [^{35}S]-methionine.

Figure 3.5.2: Purification of M protein by detergent extraction.

Right: [^{35}S]-Methionine labeled VSV was extracted with octyl-D-glucoside and fractionated on an Ultrogel AcA-54 column as described in Methods. The peak elutin between fractions 18-23 was pooled, dialyzed, and analyzed on a 10% polyacrylamide gel. V_o indicates the void volume as determined with Blue Dextran.

Left: The figure shows a densitometric scan of the autoradiogram of a 10% polyacrylamide gel of the purified M protein. Lane a, [^{35}S]-methionine labeled VSV marker; lane b, supernatant recovered after high salt detergent extraction and centrifugation of VSV; lane c, purified M protein.


L
G
N
M
a
b
c


cpm x 10⁻³
3
2
1
V₀
M
10
20
30
40
50
FRACTION

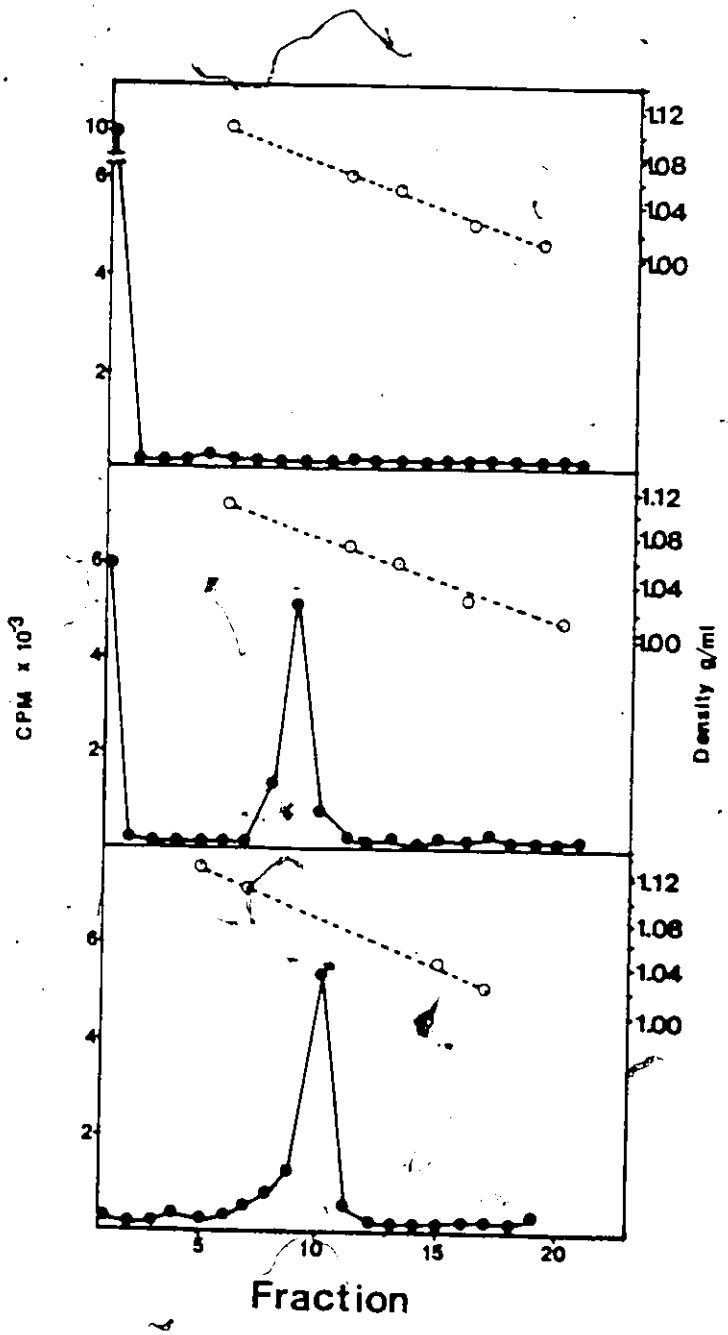
After centrifugation, the gradients were fractionated from the bottom and aliquots were assayed for radioactivity. As seen in Figure 3.5.3, M protein extracted from polyacrylamide gels and reconstituted in the absence of PC remained at the bottom of the tube following centrifugation (panel a). However, when M was mixed with PC prior to detergent dialysis, it floated upwards due to its association with the lipid vesicles. The peak of radioactivity containing [^{35}S]-methionine labeled M protein and PC vesicles had a density of 1.09 g/ml and approximately 60% of the input radioactivity was incorporated into the vesicles with the remainder having pelleted to the bottom of the tube (panel b). Nondenatured M protein purified by the detergent extraction procedure also exhibitied a strong affinity for lipid vesicles (Fig. 3.5.4). In this case, approximately 45% of the M protein floated upwards with lipid vesicles.

It was also demonstrated that M protein could spontaneously interact with preformed vesicles. This was accomplished by mixing detergent free M protein (detergent removed by dialysis) with preformed PC vesicles, vortexing the mixture, and floating it through a sucrose gradient. As shown in Figure 3.5.3, panel c, approximately 90% of the M protein was reconstituted when preformed vesicles were used. In this case, the peak containing M protein and phospholipid had a density of 1.085 g/ml.

Thus, M protein, irrespective of method of isolation, has a strong affinity for phospholipid vesicles whether the reconstitution

Figure 3.5.3: Incorporation of polyacrylamide gel purified M protein into artificial lipid vesicles.

[^{35}S]-methionine labeled M protein, purified directly from polyacrylamide gels, was added to dried egg phosphatidyl choline in the presence of 1% Na deoxycholate and reconstituted by detergent dialysis as described in Methods. After flotation centrifugation through a 5-30% sucrose density gradient, fractions were collected from the bottom of the tube and aliquots were assayed for radioactivity. Panel a, reconstitution of M protein in the absence of lipid; panel b, reconstitution of M protein in the presence of lipid (protein/lipid ratio, 1:10, w/w); panel c, reconstitution of detergent free M protein with preformed phospholipid vesicles (protein/lipid ratio, 1:10, w/w). The density of the fractions collected was determined by refractometry.


CPM x 10^{-3}
Density g/ml
Fraction
1.12
1.08
1.04
1.00
10
6
4
2
5
10
15
20

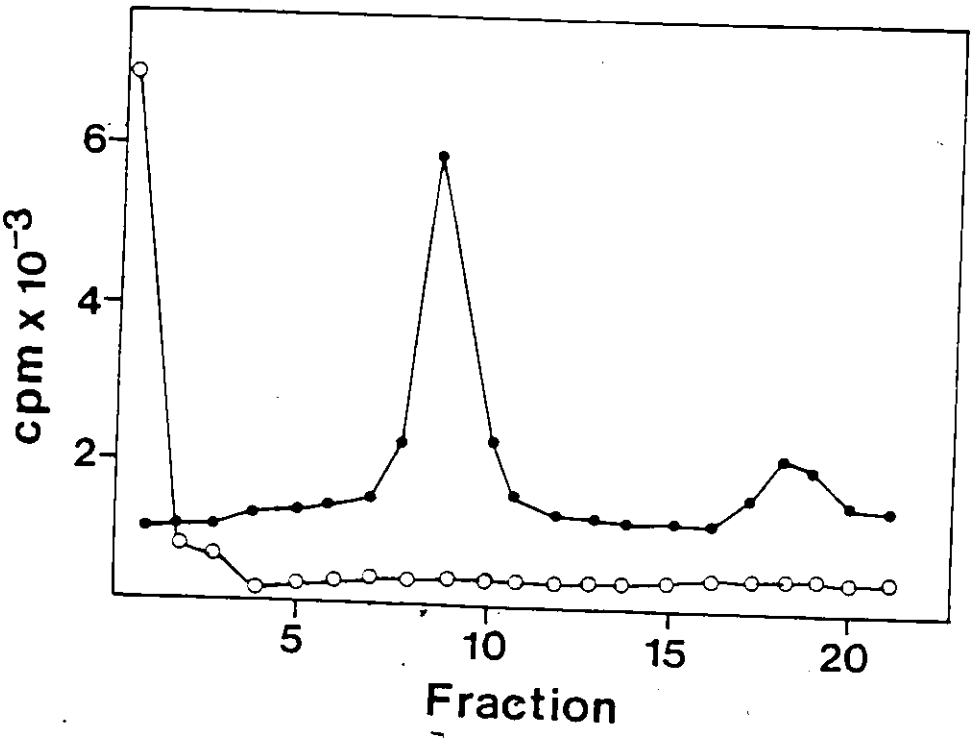


Figure 3.5.4: Incorporation of detergent extracted M protein into artificial lipid vesicles.

[^{35}S]-Methionine labelled M protein, purified as described in Figure 3.5.2, was reconstituted with phosphatidylcholine (protein/lipid ratio, 1:10, w/w) by detergent dialysis. The figure shows the profile obtained from the sucrose density gradient following centrifugation. (o) M protein reconstituted in the absence of lipid; (•) M protein reconstituted in the presence of lipid.

was done by the detergent dialysis procedure or with preformed vesicles. This is apparent from the fact that the M protein floated upwards through a sucrose gradient against a centrifugal force in excess of 100,000 g.

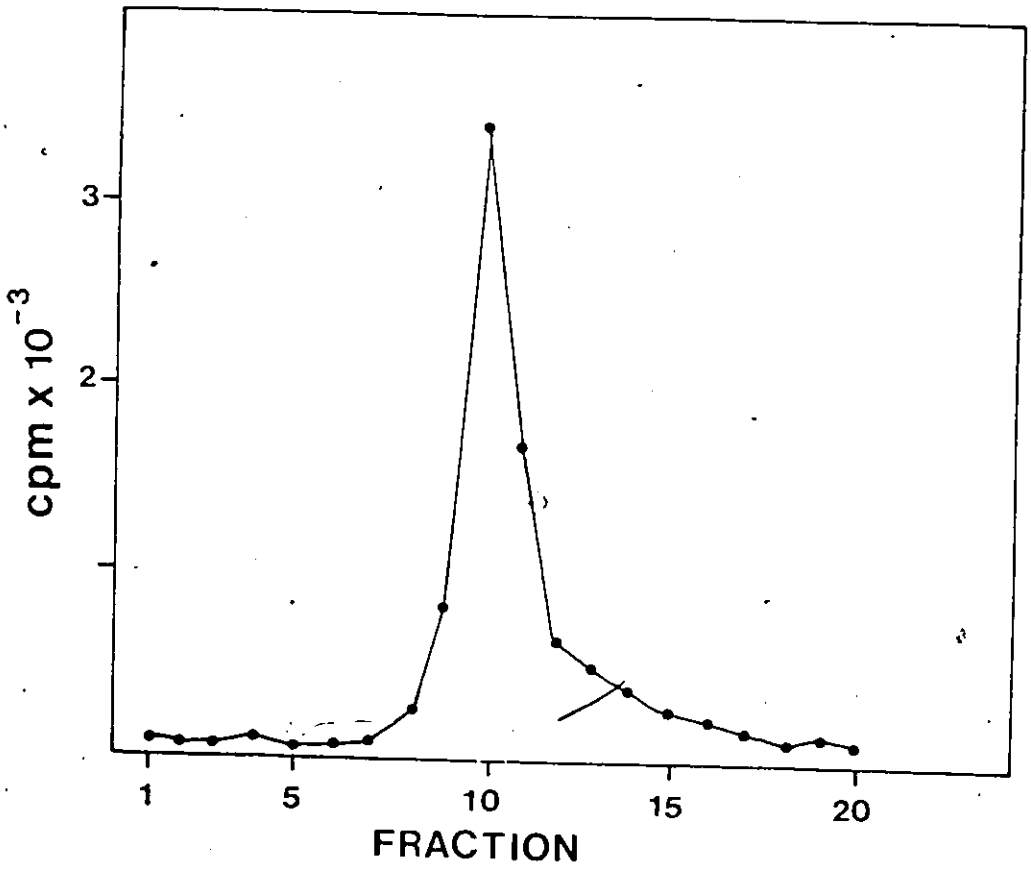
The M protein-lipid interaction observed is not due to nonspecific electrostatic association since M protein remained associated with the vesicles even after reflotation on a sucrose gradient containing 0.5 M NaCl (Fig. 3.5.5). The fact that salt does not disrupt the association indicates that the interaction is at least partly hydrophobic in nature.

In order to demonstrate that the M protein reconstituted by the detergent dialysis procedure was not due to physical entrapment of the protein during vesicle formation, reconstituted vesicles were isolated, treated with thermolysin, and reisolated on flotation gradients. The protease treated vesicles were pelleted and analyzed on a polyacrylamide gel. Analysis of the vesicles by gel electrophoresis showed that all the M protein was susceptible to proteolytic attack (Fig. 3.5.6, lane c). Thus, the M protein is present on the external surface of the lipid vesicles and the observed interaction is therefore not due to physical entrapment during vesicle formation.

The incorporation of M protein into lipid vesicles was further studied by reconstituting M protein which was labeled with [^{3}H] glycine, leucine, valine, and phenylalanine with lipid vesicles. The reconstituted vesicles were then digested with either

Figure 3.5.5: Effect of salt on the association of M protein with lipid vesicles.

[^{35}S]-Methionine labeled M protein was reconstituted with phosphatidylcholine vesicles by detergent dialysis as described in Methods. Following centrifugation, the peak fractions from the gradient were pooled and the vesicles were recovered by centrifugation. The pellet was suspended in 0.05 M Tris-HCl (pH 7.6), 0.005 M KCl, 0.5 M NaCl, made 40% with respect to sucrose and re-analyzed by flotation through a 5-30% sucrose density gradient.


cpm x 10^{-3}
3
2
1
5
10
15
20
FRACTION

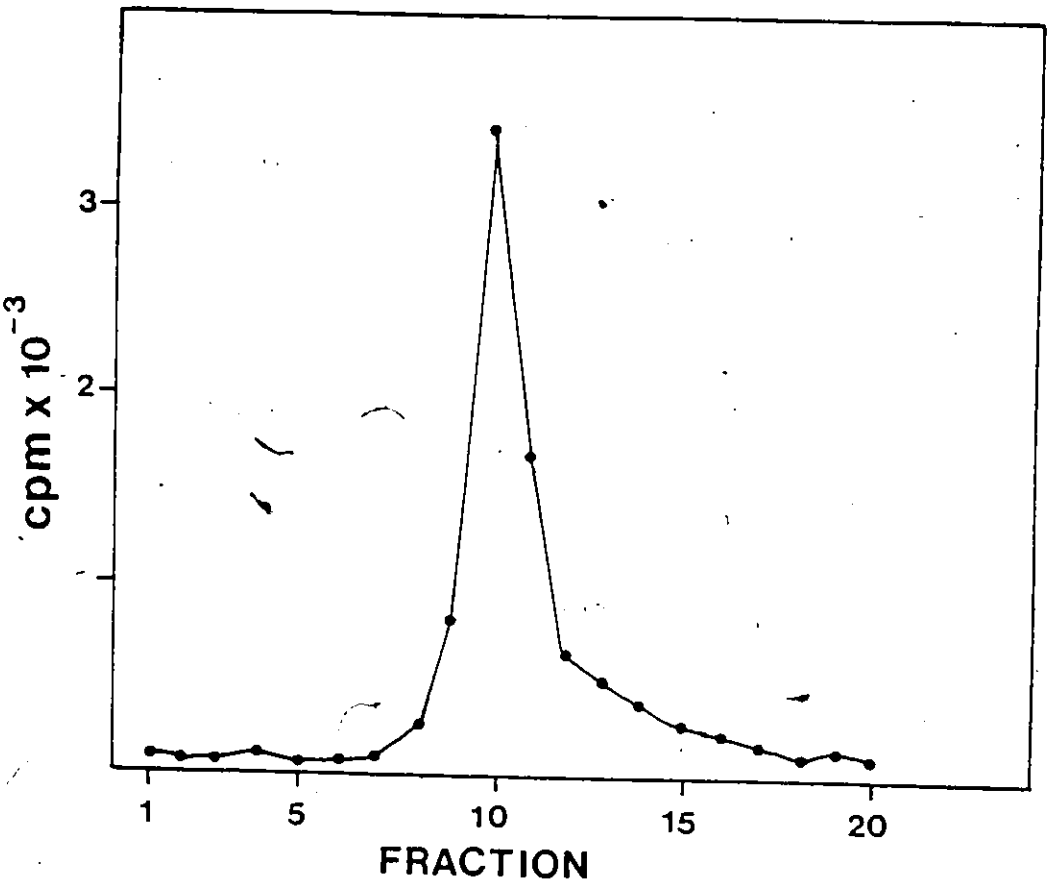


Figure 3.5.5: Effect of salt on the association of M protein with lipid vesicles.

[^{35}S]-methionine labelled M protein was reconstituted with phosphatidylcholine by detergent dialysis as described in Methods. The peak fractions were pooled, and the vesicles were recovered by centrifugation. The pellet was suspended in 0.05 M Tris-HCl (pH 7.6), 0.005 M KCl , 0.5 M NaCl, made 40% with respect to sucrose and reanalyzed by flotation through a 5-30% sucrose density gradient.

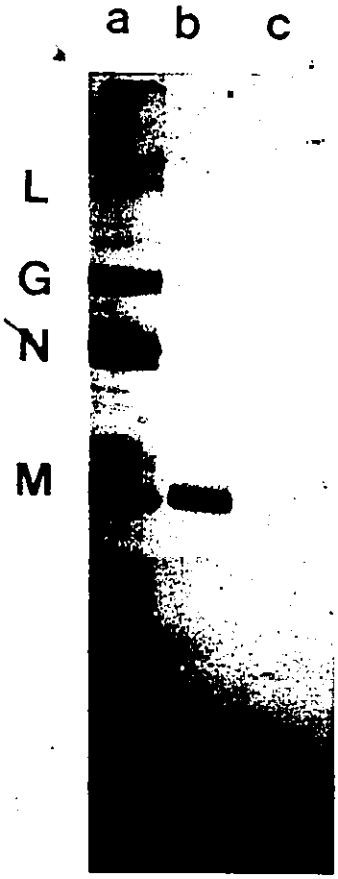


Figure 3.5.6: Polyacrylamide gel analysis of M protein incorporated into lipid vesicles.

The peak fraction in Figure 3.5.2, panel b, was collected and the vesicles pelleted by centrifugation. The pellet was suspended and one half of the sample was incubated with thermolysin (protein/enzyme ratio, 3:1, w/w) as described in Methods. Following incubation, the vesicles were made 40% with respect to sucrose and reisolated by flotation through a sucrose gradient. The vesicles were recovered and analyzed on a 15% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV; lane b and c, M protein containing vesicles before and after treatment with thermolysin, respectively.

thermolysin or trypsin and analyzed on flotation gradients. The results presented in Figure 3.5.7, a and b, show that approximately 10% of the radioactivity associated with the vesicles floated upwards to a density of 1.09 g/ml while the rest of the radioactivity was removed by the protease treatment and remained at the bottom of the tube. Since it was shown previously that none of the M protein remains intact following proteolytic digestion (see Fig. 3.5.6, lane c), the evidence suggested that a part of the M protein was protected from proteolytic digestion by being associated with the lipid vesicles. Analysis of the material present in the vesicles before and after digestion with protease showed that all the M protein was sensitive to proteolytic treatment, however, small peptides remained associated with the vesicles (fig. 3.5.8). These peptides, having an apparent molecular weight in the range of 4 - 5000 D presumably represent that region of M protein that is directly interacting with the lipid bilayer and thus inaccessable to proteolytic digestion. The fact that the M protein was completely digested by the added proteases suggested that the M protein associated with the vesicles had an external orientation and thus the material floating upwards after proteolytic digestion did not represent undegraded M protein or M protein entrapped within the vesicles.

In order to investigate whether tryptic fragments of M protein are able to associate with lipid vesicles, detergent free M protein labeled with [^{3}H] [illegible]ine, leucine, valine, and phenylalanine was digested with trypsin. The tryptic fragments were

Figure 3.5.7: Proteolytic digestion of lipid vesicles containing M protein.

Purified M protein, labeled with [^{3}H]-glycine, valine, leucine, and phenylalanine, was reconstituted into phospholipid vesicles as described. The vesicles were treated with trypsin or thermolysin (protein/enzyme ratio, 3:1, v/v) and analyzed by flotation through a sucrose gradient. Panels a and b show the flotation profile obtained when vesicles containing M protein were treated with trypsin and thermolysin, respectively. Panel c shows the profile obtained with M protein that was treated with trypsin prior to incorporation into preformed phospholipid vesicles.

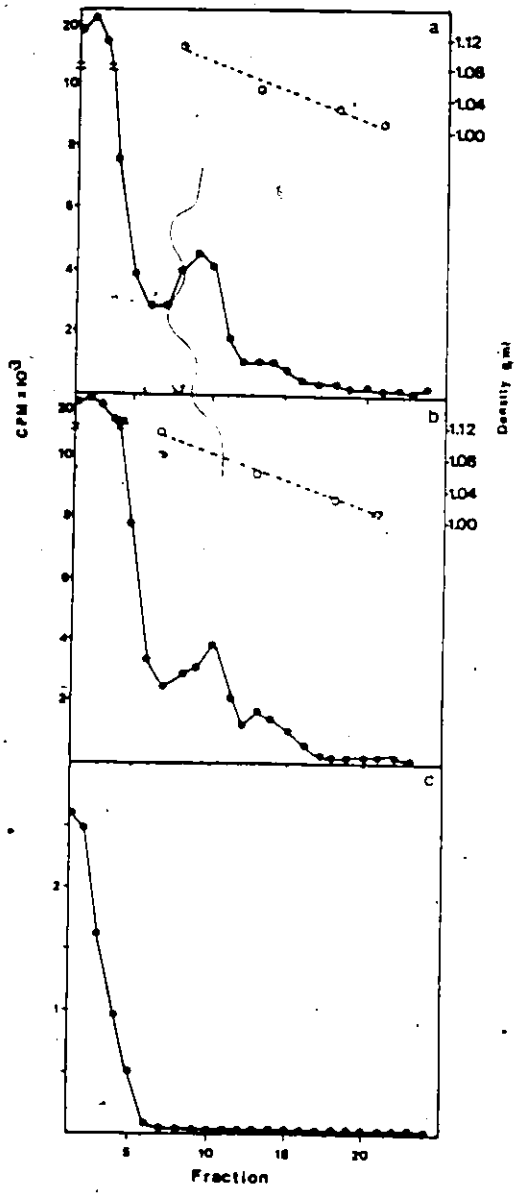
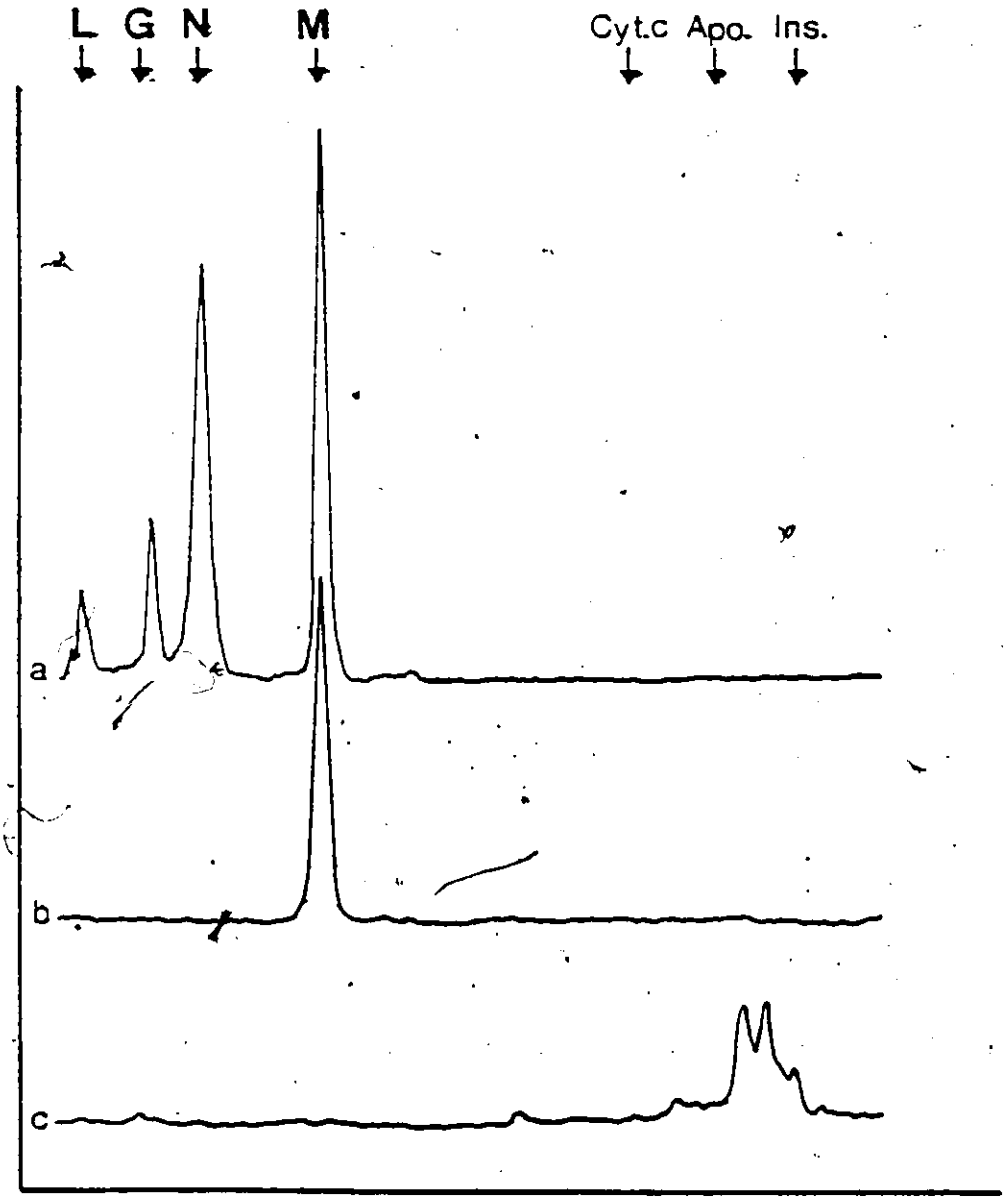

a
b
c
1.12
1.08
1.04
1.00
CPM x 10^3
Density g/ml
Fraction
5
10
15
20

Figure 3.5.8: Polyacrylamide gel analysis of M protein containing lipid vesicles after treatment with trypsin.

The peak fraction shown in Figure 3.5.7, panel a, was collected and the vesicles were pelleted by centrifugation. The pellet was suspended in electrophoresis sample buffer and analyzed on a 20% polyacrylamide gel. The figure shows a densitometric tracing of the autoradiogram. Lane a, [^{35}S]-methionine labeled VSV; lanes b and c, lipid vesicles containing M protein before and after treatment with trypsin, respectively. Molecular weight markers, which were visualized by staining the gel with Coomassie Blue, included cytochrome c (12,600D), aprotinin (6,500D), and insulin (3,500D).


L
G
N
M
Cyt.c
Apo.
Ins.
a
b
c

added to preformed phospholipid vesicles, vortexed, and analyzed by flotation through a sucrose gradient. As shown in Fig. 3.5.7, panel c, all the radioactivity remained at the bottom of the tube. Thus, tryptic fragments of M protein did not associate with lipid vesicles to an observable extent.

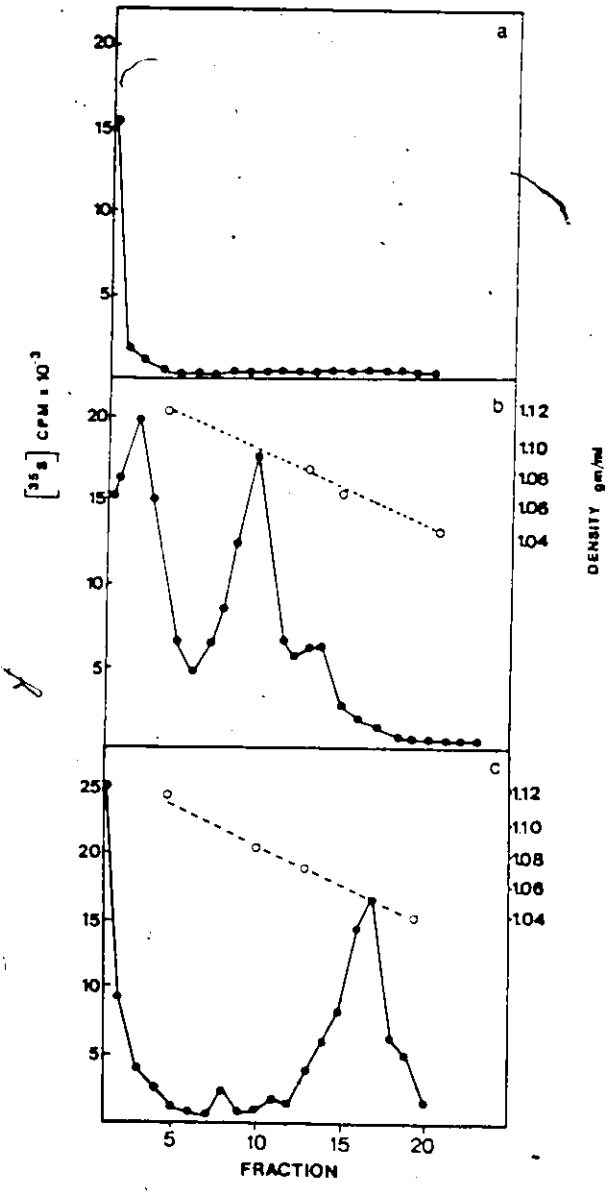
Association of N Protein with Lipid Vesicles Containing M Protein

The ability to reconstitute the M protein on the external surface of phospholipid vesicles provides a model system with which one can directly investigate protein-protein interactions.

Much of the available evidence suggests that the association of the viral nucleocapsid with the envelope occurs through interaction between N and M proteins (Mancarella and Lenard, 1981; Mudd and Swanson, 1978; Shnitzer and Lodish, 1979; Wilson and Lenard, 1981; Knipe et al., 1977; Wagner, 1975). To directly test this idea, lipid vesicles were reconstituted in the presence of either N protein alone, or a mixture of N and M proteins. The results shown in Figure 3.5.9, panel a, demonstrate that N protein has no affinity for lipid vesicles and thus remained at the bottom of the gradient. When, however, an equal mixture of N and M proteins, both labeled with [^{35}S] methionine, was reconstituted with phosphatidylcholine vesicles, a radioactive peak was observed in the sucrose gradient at a density of 1.10 g/ml (Fig. 3.5.9, panel b). When a mixture of G and N protein was used for the reconstitution, a peak of radioactivity was observed at a density of 1.05 g/ml (Fig. 3.5.9, panel c).

Figure 3.5.9: Association of N protein with lipid vesicles.

[^{35}S]-Methionine labeled M, N, and G protein, were purified from polyacrylamide gels and equivalent amounts of N and M protein, or N and G protein, were mixed with phospholipid (protein/lipid ratio, 1:5, w/w) and reconstituted by detergent dialysis. Panel a, flotation profile of N protein reconstituted with lipid vesicles; panel b, reconstitution of a mixture of N and M protein with lipid vesicles; panel c, reconstitution of a mixture of N and G protein with lipid vesicles.


a
b
c
20
15
10
5
25
1.12
1.10
1.08
1.06
1.04
[35S] CPM x 10-3
DENSITY gm/ml
5
10
15
20
FRACTION

To establish the identity of the proteins associated with the vesicles at these densities, the vesicles were sedimented and analyzed by gel electrophoresis. The results show that the peak obtained from the reconstitution of a mixture of N and M proteins contained both N and M proteins (Fig. 3.5.10, lane b). The ratio of N to M was approximately 1:3 as determined from a densitometric tracing of the autoradiogram. This ratio was arrived at after correcting for the number of methionines present in the respective proteins as determined from cDNA sequences (Rose and Gallione, 1981; Gallione et al., 1981). Similar results were obtained when N protein and detergent extracted nondenatured M protein was used for the reconstitution (Fig. 3.5.10, lanes f and g). In this case, the N to M ratio was approximately 1.1:1.

In the control experiment where an equal amount of G and N protein was used for the reconstitution, the only species present in the peak that floated upwards was G protein (Figure 3.5.10, lanes d and e). The density of the peak (1.05 g/ml) agrees with the value obtained previously by Petri and Wagner (1979; 1980) when purified G protein was incorporated into artificial PC vesicles.

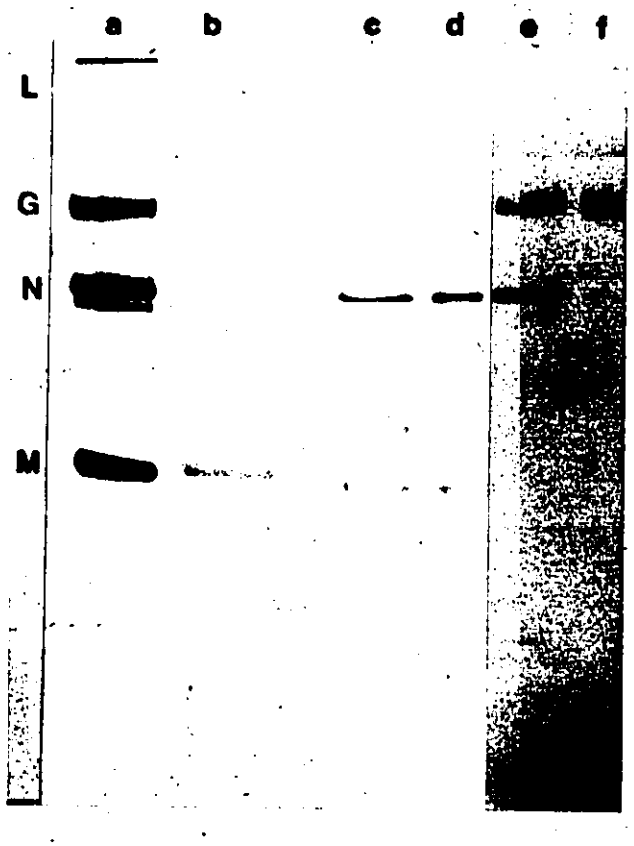
Thus, it appears that N protein can associate with lipid vesicles only in the presence of M protein.

Reconstitution of N Protein and M Protein Synthesized in vitro

The N protein used in the above reconstitution studies was isolated from polyacrylamide gels and, thus, presumably in a

Figure 3.5.10: Polyacrylamide gel analysis of proteins associated with lipid vesicles.

The peak fractions from Figure 3.5.9, panels a and b, were recovered by centrifugation and analyzed on a polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV; lane b, lipid vesicles recovered from the reconstitution of a mixture of N and M proteins; lane e, mixture of G and N proteins prior to reconstitution; lane f, vesicles recovered from the reconstitution of a mixture of G and N proteins. Lanes c and d represent a separate experiment in which N protein and M protein which was purified by detergent extraction as described in Figure 3.5.2 was used for reconstitution. Lane c, mixture of N and M proteins prior to reconstitution; lane d, vesicles recovered from a reconstitution of a mixture of N protein and detergent extracted M protein.


a
b
c
d
e
f
L
G
N
M

denatured state. The results observed may thus not truly reflect the type of interaction that may be occurring in vivo.

The isolation of N protein from virions or from infected cells in a soluble nondenatured state is difficult because of its rapid association with nucleocapsid structures soon after its synthesis (Hsu et al., 1979). One possible way of overcoming these problems is to perform reconstitutions with proteins synthesized in vitro using soluble cell free protein synthesizing extracts. The proteins synthesized in vitro would have distinct advantages since they would be soluble in a membrane free environment because they are made in very low amounts. In addition, the fact that viral replication and transcription is not occurring precludes the possibility of forming RNP complexes.

Reconstitution experiments were, therefore, attempted with proteins synthesized in vitro. The source of viral mRNA was an S4 extract from VSV infected L cells (Toneguzzo and Ghosh, 1976). Protein synthesis was carried out using a rabbit reticulocyte cell free lysate (Pelham and Jackson, 1975). Figure 3.5.12, lanes b and c, show the proteins synthesized in the reticulocyte lysate in the presence or absence of viral specific mRNA.

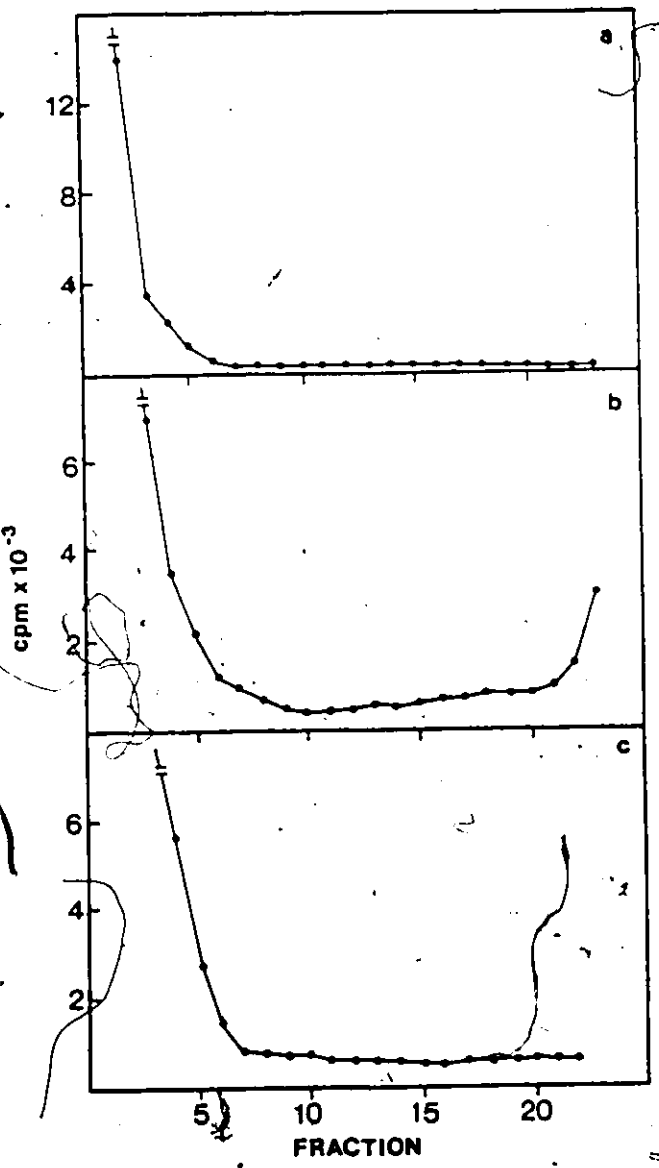
For the reconstitution experiments, viral proteins were synthesized in the presence of [^{35}S]-methionine. After a one hour incubation, unincorporated radioactivity was removed by dialysis. The reaction mixture was then mixed with 100 ug of preformed PC

vesicles, vortexed for five minutes, and made 40% with respect to sucrose. The mixture was then floated through a discontinuous sucrose gradient. In this case, the gradient consisted of only a 35% sucrose solution overlaid on the 40% sucrose solution. Thus, after equilibrium centrifugation, the lipid vesicles and any associated material should float to the top of the centrifuge tube. This type of gradient, rather than the linear gradient, was used since it was not possible to determine the protein to lipid ratio of the starting material.

Figure 3.5.11, panel b, shows the profile of radioactivity obtained from the gradients that contained in vitro synthesized proteins and preformed phospholipid vesicles. Most of the [^{35}S]-methionine labeled proteins synthesized in the reticulocyte lysate remained at the bottom of the tube, however, approximately 10% of the input radioactivity floated to the top of the tube. In the case where the reaction contained vesicles but was not programmed with exogenous viral mRNA, all the radioactivity remained at the bottom of the tube (Fig. 3.5.11, panel b). When the reaction was programmed with viral mRNA but no phospholipid vesicles were added, the radioactivity also remained at the bottom of the tube. Thus, the peak observed at the top of the gradient in panel b of Figure 3.5.11 must represent viral specific proteins made in the cell free extract that are associated with the lipid vesicles. When the phospholipid containing fractions were recovered by centrifugation and analyzed by gel electrophoresis, it was found that both N and M proteins were

Figure 3.5.11: Association of *in vitro* synthesized proteins with lipid vesicles.

VSV specific proteins were synthesized *in vitro* in a nuclease treated reticulocyte extract as described in Methods. After a 90 minute incubation, the reaction was dialyzed to remove unincorporated label and mixed with preformed phosphatidylcholine vesicles. The mixture was vorxeted, made 40% with respect to sucrose, and floated through a layer of 35% sucrose. Panel a, flotation pattern of a reaction mixture which was not programmed with exogenous VSV mRNA; panel b, flotation pattern of a reaction mixture programmed with VSV mRNA; panel c, flotation pattern of VSV mRNA programmed reaction mixture reconstituted in the absence of lipid vesicles. The amount of starting radioactivity (approximately 250,000 cpm) was the same for each reconstitution.


a
b
c
12
8
4
6
4
2
6
4
2
cpm x 10⁻³
5
10
15
20
FRACTION

associated with the vesicles (Fig. 3.5.12, lane f). The ratio of N to M was found to be approximately 1:2.3 as determined by a densitometric scan of the autoradiogram. This is in contrast to a ratio of N/M of 3:1 found in the proteins synthesized in vitro prior to reconstitition (see Fig. 3.5.12). The association of in vitro synthesized M and N with the lipid vesicles was not due to protein entrapment since preformed vesicles were used for the reconstitutions.

At least some of the N protein was very tightly associated since the interaction could not be disrupted by the inclusion of 0.5 M NaCl to the 35% sucrose solution prior to flotation (Fig. 3.5.12, lane g). The inclusion of salt in the gradient resulted in a change in the ratio of N to M from 1:2.3 to 1:4.5 (Table VI). Thus, approximately half of the N molecules that are present with M in the lipid vesicles can be removed by high salt treatment.

No proteins were present in the membrane fraction of the gradient when reconstitution was performed with reaction mixtures in the absence of added viral mRNA (Fig. 3.6.12, lane e). Thus, both N and M proteins synthesized in vitro and, therefore, in a soluble nondenatured state, can associate with lipid vesicles in vitro.

The previous experiments using proteins purified directly from virions indicated that the presence of N protein with the lipid vesicles was entirely dependent on the presence of M protein. This cannot be said of the reconstitutions with proteins synthesized in vitro since both N and M proteins are present at the same time.

Figure 3.5.12: Characterization of *in vitro* synthesized proteins associated with lipid vesicles.

The lipid vesicles reconstituted from the reaction mixtures containing *in vitro* synthesized proteins shown in Figure 3.5.11 were recovered by centrifugation and analyzed on a 10% polyacrylamide gel. Lanes a and d, [^{35}S]-methionine labeled VSV; lanes b and c, proteins synthesized *in vitro* in the reticulocyte lysate in the absence and presence of VSV mRNA, respectively; lane e, lipid vesicles recovered from the reaction mixture in the absence of VSV mRNA; lane f, lipid vesicles recovered from the reaction mixture in the presence of VSV mRNA; lane g, same as in lane f but the vesicles were pelleted through a buffer solution containing 0.5 M NaCl prior to polyacrylamide gel electrophoresis. Each reconstitution was done with an equivalent amount of radioactivity, and after reconstitution, an equivalent amount of vesicles were analyzed on the gel in each case.

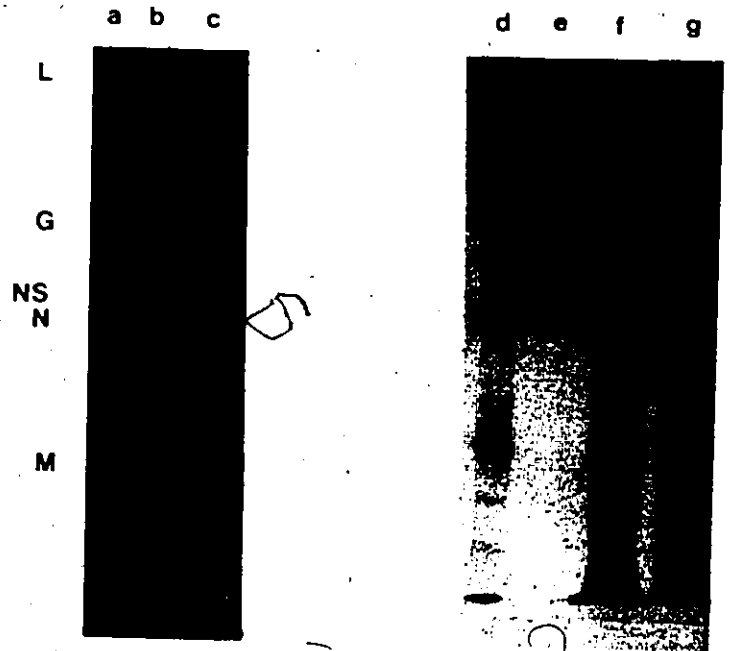

a b c
d e f g
L
G
NS
N
M

TABLE VI: RELATIVE PROPORTION OF N AND M PROTEINS ASSOCIATING WITH LIPID VESICLES

	N/M RATIO
Reaction mixture	3:1
Associated with lipid vesicles after reconstitution	1:2.3
Associated with lipid vesicles after treatment with 0.5 M NaCl	1:4.5

The table represents a quantitation of the in vitro synthesized proteins found associated with the lipid vesicles after reconstitution as determined from a densitometric scan of the autoradiogram shown in Figure 3.5.12. The values obtained were corrected for the number of predicted methionine residues present in N and M protein as deduced from the nucleotide sequence of the respective mRNAs (Rose and Gallione, 1981; Gallione et al., 1981).

In order to ascertain whether the association of N protein with the lipid vesicles is dependent upon the presence of M protein, it is necessary to examine reconstitution characteristics with separated proteins. This can be done by purifying the respective viral mRNA species and translating them separately in vitro. Alternatively, one can specifically suppress the translation of a specific mRNA species by hybridization arrest with a DNA complementary to that mRNA molecule (Preston and McGeogh, 1981). The DNA-RNA duplex formed would thus prevent translation of that particular message. This technique has been widely used as a method of mapping gene products to their respective messages.

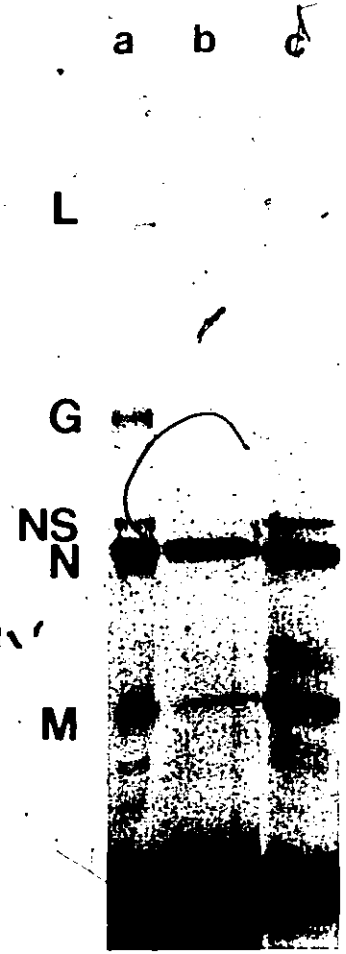
The latter approach was taken in order to prevent the synthesis of M protein in the reticulocyte lysate system. The hybridization probe used was the plasmid pM32 which was kindly provided by J. Rose. pM32 is a derivative of the cloning vector pBR322 and contains a cDNA copy of the COOH terminal coding sequences of the M protein mRNA (Rose and Iverson, 1979). Hybridization of RNA from an extract of VSV infected cells with this plasmid followed by in vitro translation should selectively inhibit the synthesis of M protein.

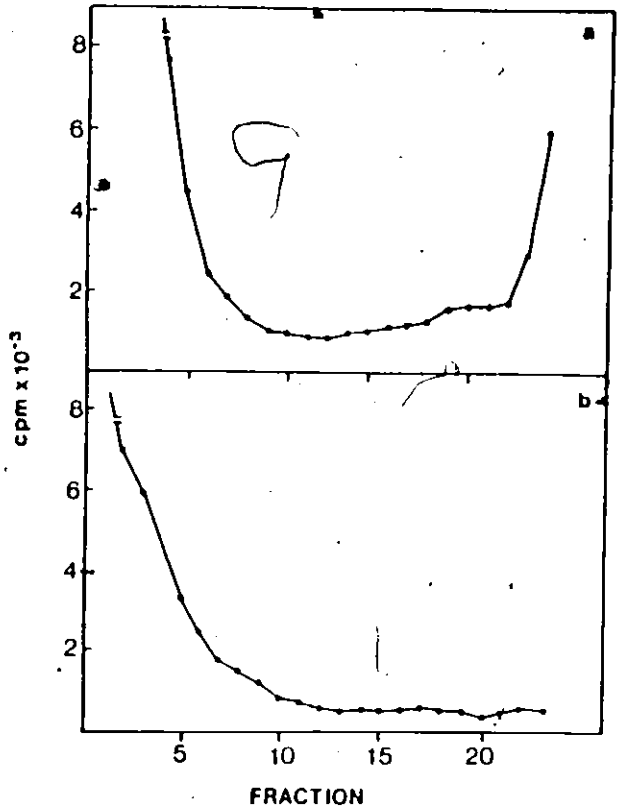
As can be seen in Figure 3.5.13, hybridization arrest with pM32 results in a substantial decrease in the synthesis of M protein. Prior to arrest, the N/M ratio in the reaction mixture was 1.2:1 as determined by densitometry, while after hybridization arrest the N/M ratio was found to be 9:1. Therefore, hybridization arrest resulted in a nearly 8 fold decrease in the synthesis of M protein.

Figure 3.5.13: Association of *in vitro* synthesized proteins with lipid vesicles in the absence of M protein.

Right: The translation of M protein in the *in vitro* protein synthesizing system was prevented by the hybridization of the M protein specific mRNA with plasmid pM32 prior to translation as described in Methods. Lane a, [^{35}S]-methionine labeled VSV marker; lane b, proteins synthesized *in vitro* with VSV mRNA in which the translation of M mRNA was arrested; lane c, proteins synthesized *in vitro* with VSV mRNA.

Left; The reaction mixtures shown in lanes a and b above were reconstituted with preformed phospholipid vesicles and floated upwards through a 35% sucrose solution. Panel a, flotation pattern of complete reaction mixture; panel b, flotation pattern of the reaction mixture in which the translation of M mRNA was arrested.


a
b
c
L
G
NS
N
M


a
b
8
6
4
2
cpm x 10⁻³
5
10
15
20
FRACTION

The hybridization arrested reaction mixture was reconstituted with preformed vesicles and analyzed on flotation gradients (Fig. 3.5.13). As seen in this figure, a definite peak of radioactivity was observed at the top of the gradient in the case where the untreated reaction mixture was used. In the case where the hybridization arrested reaction mixture was used for reconstitution, no significant amount of radioactivity was present in the lipid fraction.

Analysis of the proteins present in the lipid fractions by gel electrophoresis showed that both N and M proteins floated upwards in the complete reaction mixture, however, no labeled proteins were present in the lipid fraction when the hybridization arrested reaction was used (Fig. 3.5.14, lanes c and d).

Thus, it appears that the association of the in vitro synthesized N protein with artificial lipid vesicles is dependent upon the presence of M protein. This was further confirmed by the addition of unlabeled detergent free M protein (purified by the detergent extraction procedure) to the M arrested reaction mixture after translation. When this mixture was used for the reconstitution and the lipid vesicles analyzed on a gel, it was found that labeled N protein was now present in the lipid fraction (Fig. 3.5.14, lane e).

Association of M Protein and RNP

N protein, both newly synthesized and present in the virus, is normally found in tight association with the RNP complex. It was therefore decided to examine whether the RNP itself had any affinity

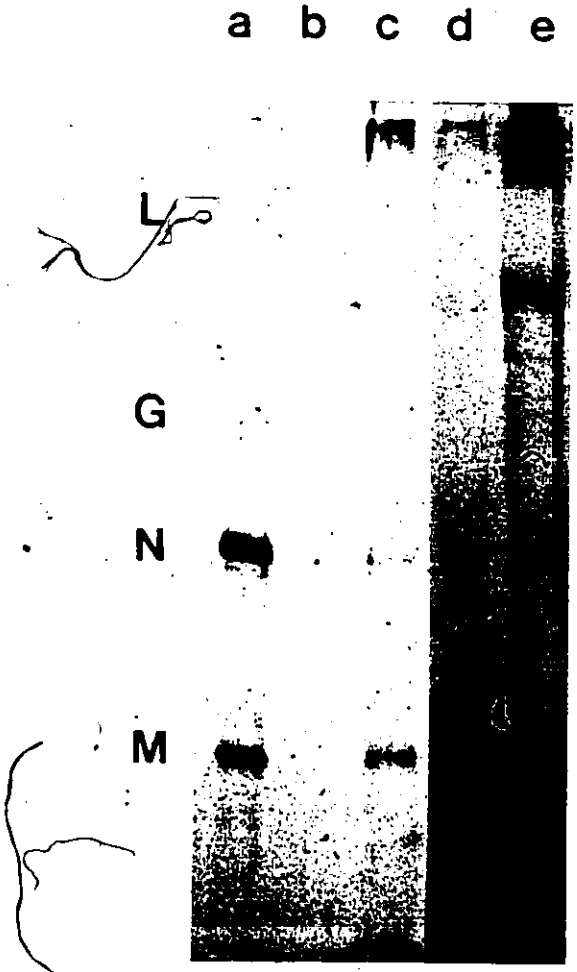


Figure 3.5.14: Characterization of *in vitro* synthesized proteins associated with lipid vesicles in the presence and in the absence of M protein.

The lipid vesicles reconstituted from the reaction mixtures containing *in vitro* synthesized proteins as shown in Figure 3.5.13 were recovered by centrifugation and analyzed on a 10% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV marker; lane b, vesioles recovered from a reconstitution of a reaction mixture in which no viral mRNA was added; lane c, vesicles recovered from a reaction mixture in which viral mRNA was added; lane d, same as lane c, but the translation of M mRNA was prevented by hybridization arrest; lane e, same as lane d, but 10 ug of unlabeled, detergent purified, M protein was added to the reaction mixture after translation but before reconstitution.

for purified M protein.

The RNP complex was purified from [^{35}S]-methionine labeled virus by detergent disruption followed by centrifugation on a glycerol gradient (Carrol and Wagner, 1978). As seen in Figure 3.5.15, the purified RNP complex still retained a small amount of M protein. This is in agreement with the observation that a small amount of M protein was consistently associated with in vitro synthesized RNP particles even after centrifugation through a 30% CsCl solution (Ghosh and Ghosh, 1982).

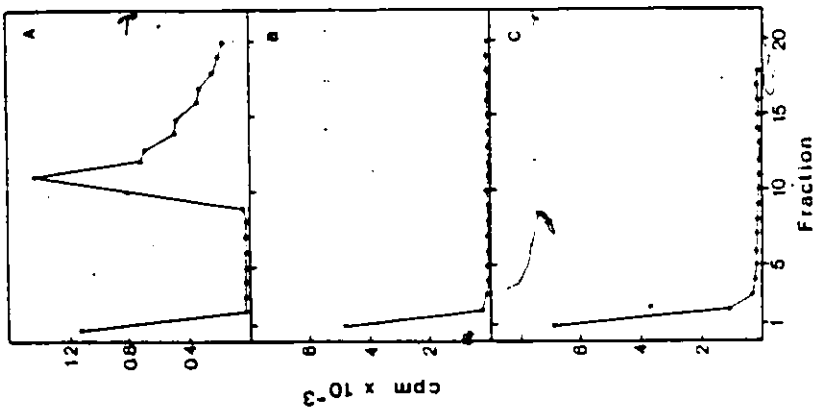
To test whether purified M protein could associate with the RNP, reconstituted phospholipid vesicles containing M protein were purified, mixed with the RNP and reanalyzed by flotation through a linear sucrose gradient. If the M is associating with the RNP it should remain at the bottom of the tube since the large size of the RNP would prevent flotation. As shown in Figure 3.5.15, RNP reconstituted with preformed phospholipid vesicles in the absence of M protein remains at the bottom of the tube. However, when reconstituted vesicles containing M protein were mixed with the RNP, all the radioactivity remained at the bottom of the tube also. This suggested that the vesicles containing M protein interacted with the RNP but could no longer float upwards due to the size of the associated nucleocapsid.

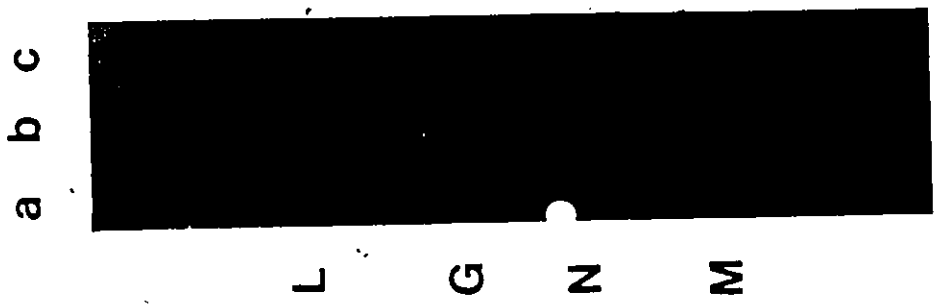
The data presented in this section demonstrates that the M protein of VSV, irrespective of method of isolation or synthesized in vitro, has a strong affinity for artificial membranes. Presumably

Figure 3.5.15: Association of M protein containing lipid vesicles with purified RNP particles.

Right: [^{35}S]-Methionine labeled VSV was disrupted with octyl-D-glucoside and RNP particles were purified on a glycerol gradient as described in Methods. Shown here is an autoradiogram of the purified RNP as determined by electrophoresis on a 10% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV marker; lane b, proteins remaining in the supernatant after detergent disruption of VSV and centrifugation; lane c, purified RNP.

Left: Lipid vesicles containing [^{35}S]-methionine labeled M protein were isolated from a sucrose gradient and mixed with an equivalent amount (in terms of radioactivity) of purified [^{35}S]-methionine labeled RNP. The mixture was made 40% with respect to sucrose and floated upwards through a 5-30% sucrose density gradient. Panel a, M protein reconstituted with lipid vesicles (protein/lipid ratio, 1:10, w/w); panel b, RNP reconstituted with preformed lipid vesicles; panel c, RNP reconstituted with lipid vesicles containing M protein.


A
B
C
12
08
04
8
4
2
8
4
2
1
5
10
15
20
Fraction
cpm x 10^-3


a b c
L G N M

this property is also manifested in the intact virion. The fact that M protein remains bound to liposomes even after treatment with 0.5 M NaCl suggests that the interaction is hydrophobic in nature. This is supported by the fact that small molecular weight peptides remain associated with the vesicles following proteolytic digestion. These results are similar to those reported with the influenza M protein (Gregoriades, 1980), however, in the case of VSV, M protein does not contain any substantial stretch of hydrophobic or nonpolar amino acids that would facilitate this type of interaction (Rose and Gallione, 1981). Perhaps the tertiary conformation of the protein plays a role in lipid association. This is supported from the observation that tryptic peptides of M protein did not associate with preformed vesicles, which would be expected to occur if long stretches of M protein had an affinity for lipid vesicles.

The interacion of M protein with artififical lipid vesicles has recently been reported by another group (Zakowski et al., 1981). The results regarding M protein are substantially similar to those reported here, however, they could achieve efficient incorporation of M protein only with lipid vesicles containing negatively charged phospholipids such as phosphatidylserine, phosphatidylinositol, or phosphatidic acid and not in vesicles composed solely of phosphatidycholine. In contrast, the data in this section demonstrates that M protein can be efficiently incorporated into neutral PC vesicles by the method of Gregoriades (1980). The non requirement of negatively charged phospholipids in this case as well

as in the case of the influenza M protein (Gregoriades, 1980) could be due to the fact that we used the charged detergent deoxycholate. In contrast, Zakowski et al., used the neutral detergent octyl glucoside for their reconstitition studies. In addition, detergent dialysis in the case of deoxycholate was carried out against buffer containing low concentrations of salt and a monovalent cation whereas Zakowski et al., dialyzed the detergent out against Tris buffer only. It has been demonstrated that the association of M protein with cytoplasmic membrane fractions is dependent upon the presence of cations (Morrison and McQuain, 1977).

In contrast to the studies reported here, Zakowski et al., did not observe the presence of low molecular weight peptides after proteolysis of M protein containing vesicles. This may reflect differences in protein-lipid interactions due to the differences in the phospholipid composition of the membranes.

The important observation reported in this section is that the nucleocapsid protein N can associate with lipid vesicles only in the presence of M protein. N protein itself has no affinity for membranes as demonstrated by the observation that N protein alone or mixed with G protein does not associate with liposomes. Its association with lipid vesicles in the presence of M protein must, therefore, be mediated by protein-protein interactions between N protein and M protein.

It is, therefore, concluded that M protein of VSV has two

distinct functional domains; one that can recognize and interact with the lipid bilayer, and one that can recognize and interact with the nucleocapsid protein N.

SECTION 3.6: CONSTRUCTION OF A CHIMERIC GENE BETWEEN THE VSV G AND M PROTEINS

It is clear that the synthesis of membrane proteins, their insertion into membranes, and their cellular sorting involves a complex series of co-translational and post-translational events which must be specified in part by information present in the polypeptide. It has been postulated that the information necessary for the specialized synthesis, transport, and modifications of secretory and membrane proteins is contained in discrete regions or domains of the polypeptide which are decoded by specialized mechanisms in the cell (Sabatini et al., 1982). These domains, or "topogenic sequences" (Blobel, 1980) which can be a transient or permanent feature of the polypeptide, include the signal sequence which initiates the translocation of proteins into and across specific membranes, the stop transfer sequence which, in the case of transmembrane proteins, serves to interrupt the signal sequence initiated translocation and insures a proper asymmetric orientation of the polypeptide, and sorting sequences which serve to target the newly synthesized polypeptide along selected subcellular pathways to its site of function. The VSV G protein must contain all of these sequences, or signals, encoded in its polypeptide backbone.

Studies with bacterial systems have helped elucidate the molecular basis for the specificity and function of some of these

topogenic sequences. Genetic studies have allowed the isolation of mutants containing alterations in the signal sequence region of a number of exported proteins of E. coli which have provided direct evidence for the role of this region of the polypeptide in the translocation process (Lin et al., 1978; Bassford and Beckwith, 1979; Bedouelle et al., 1980).

Until recently, the complexity of eukaryotic systems have made studies of this type virtually impossible to carry out. During the past few years, however, new advances in genetic engineering have made it possible to clone genes in large amounts. The parallel advancements made in the expression of these cloned genes in eukaryotic systems and the ease with which they can be manipulated has provided vast new opportunities to study the functional roles that various topogenic sequences may play in the biogenesis of membrane proteins.

This section describes the construction and cloning of a chimeric gene which contains the signal sequence coding region of G protein fused to the bulk of the coding sequence of M protein and initial attempts at expressing this hybrid gene in cultured mammalian cells.

There are several reasons for constructing and expressing a hybrid gene of this type. The fusion of the signal sequence coding region of the G protein to a protein that does not normally transverse membrane bilayers would indicate whether or not the signal sequence is sufficient to initiate and maintain vectorial

translocation across the membrane or whether additional sequences distal to the amino terminal signal sequence are required. M protein is chosen as the recipient of this signal sequence because is normally resides on the internal membrane surface of the infected cell and the completed virion. Thus, the maturation of M protein does not require the passage of the polypeptide across hydrophobic membrane bilayers.

In an elegant series of experiments in which the signal sequence of an E. coli outer membrane protein was fused to the gene coding for the cytoplasmic enzyme β -galactosidase, it was determined that the signal sequence was not sufficiant to lead β -galactosidase out of the cytoplasm (Moreno et al., 1980). In addition, using mutants of β-lactamase, a normally exported E. coli protein, it was determined that sequences near the COOH terminal end of the molecule were necessary for the efficient secretion of this molecule (Koshland and Botstein, 1980).

Another important reason for choosing M protein as the signal sequence acceptor is that the predicted amino acid sequence of M (Rose and Gallione, 1981) reveals the presence of a sequence of the type Asn-X- $\frac{\text{SER}}{\text{THR}}$. This is a concensus sequence that normally serves as a glycosylation site for glycoproteins (Sharon and Lis, 1982). In the M protein, this target site is not glycosylated since M protein is not inserted into the RER or Golgi complex where the glycosylating enzymes are localized. However, if M is extruded into the lumen of the RER by virtue of having the signal sequence of G attached to it,

then it may be possible to glycoslyate this site.

Thus, by the construation of this hybrid gene, one may be able to determine if the signal sequence itself is sufficient to translocate a normally cytoplasmic protein across the membrane barrier, and if so, whether the sequence Asn-X- $\frac{\text{SER}}{\text{THR}}$ on its own is sufficient to trigger glycosylation. If additional information other than this concensus sequence is required for glycosylation, then one would not expect the hybrid protein to be glycosylated even though it may be inserted into membranes via the signal sequence of G.

Results

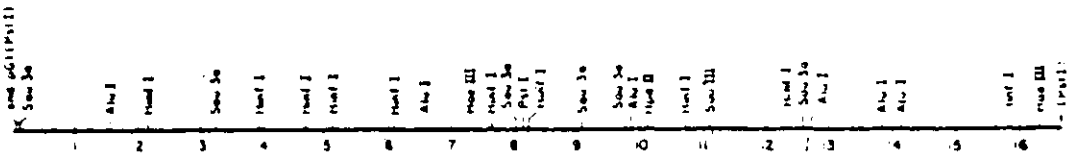
Cloning Strategy:

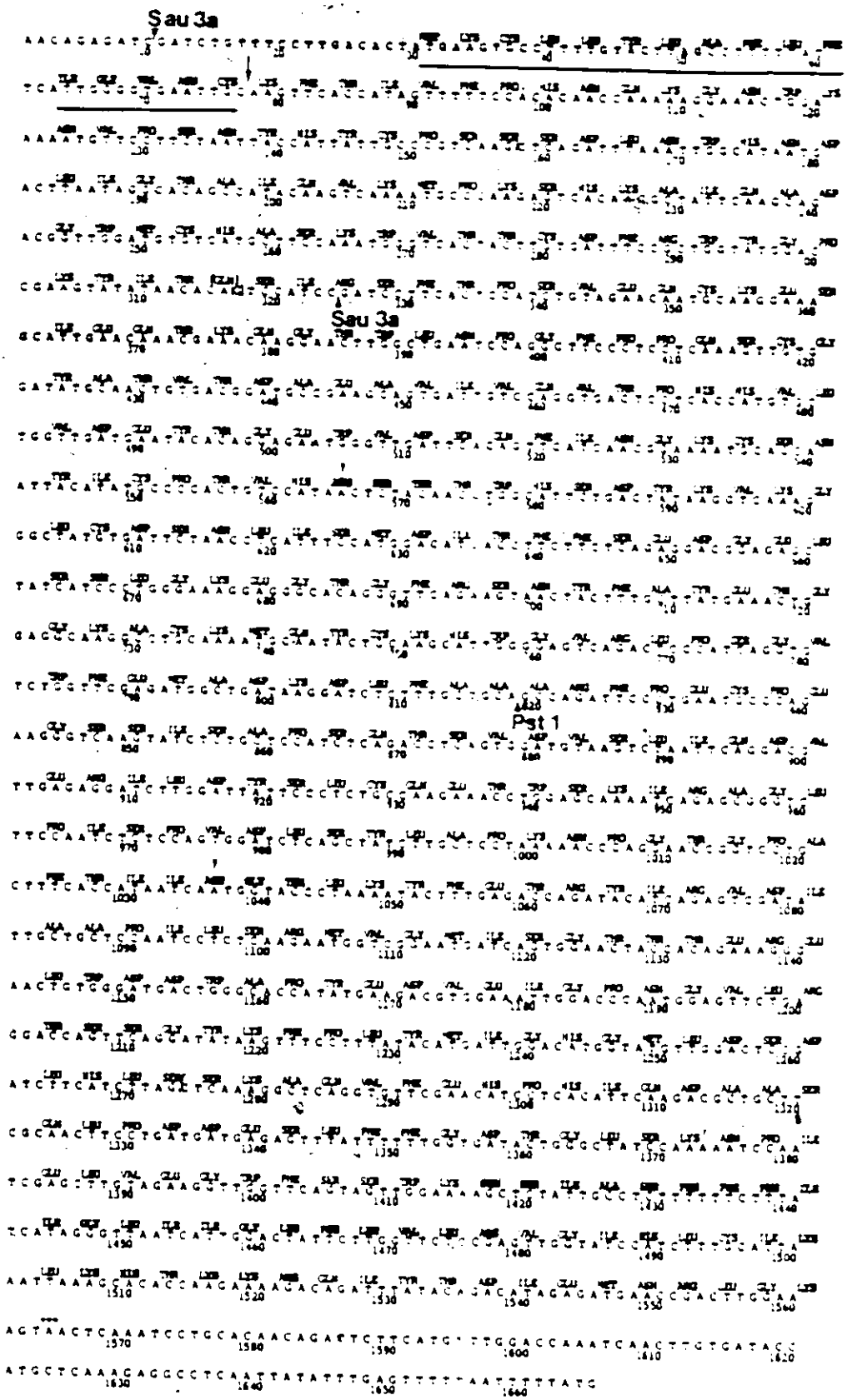
The work described in this section would not have been possible without the generosity of Dr. J. Rose in providing the recombinant plasmids pG1 and pM309 which contain cDNA copies of the complete coding sequences of the mRNA specific for G and M protein, respectively (Rose and Gallione, 1981). The nucleotide sequences and the predicted amino acid sequences for pG1 and pM309 are shown in Figure 3.6.1 and 3.6.2, respectively.

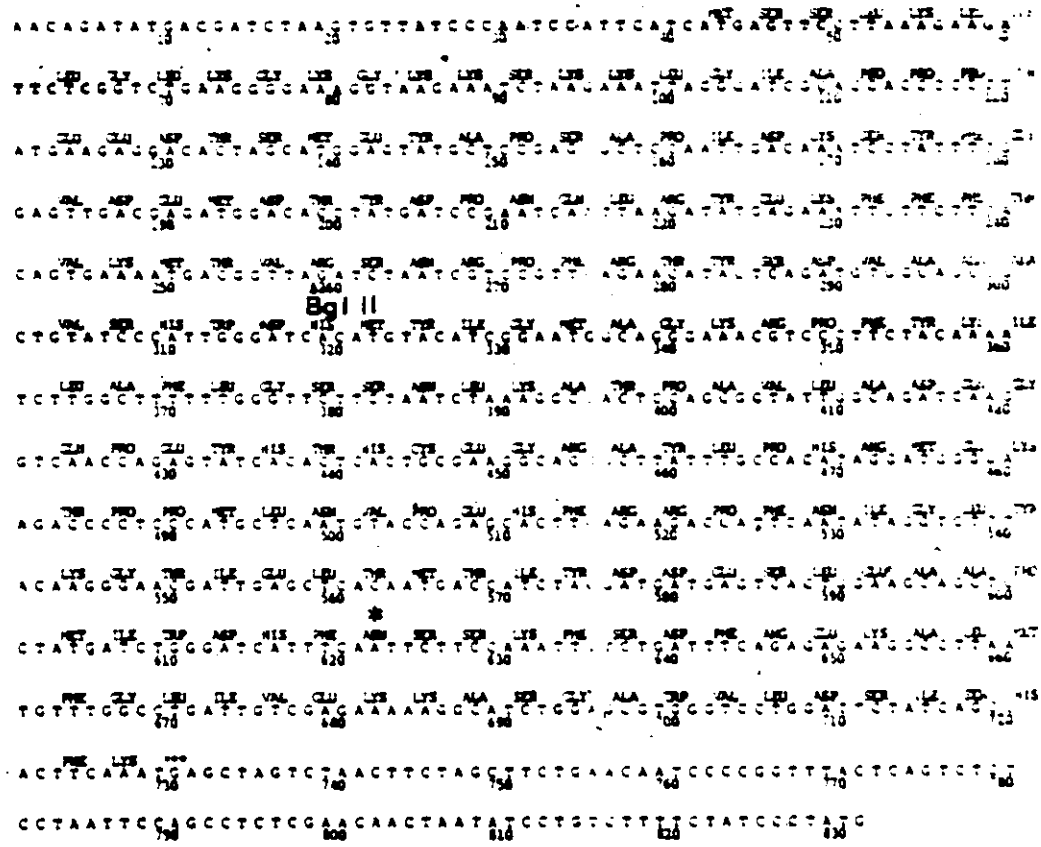
In choosing a strategy for the construction of the hybrid gene, a number of considerations had to be taken into account. It was desired to minimize the amount of G specific sequences downstream from the signal sequence coding region and, at the same time, maximize the coding sequence of the M gene. The overriding consideration, however, was to maintain the correct reading frame at the G/M junction to insure that the correct codons are read for

Figure 3.6.1: Nucleotide sequence of the VSV G mRNA and the predicted protein sequence.

This figure was taken from Rose and Gallione (1981). The signal sequence coding region is underlined. The relevant restriction enzyme sites used in the cloning strategy outlined in figure 3.6.4 are indicated by the bold face lettering. The restriction map of the pG1 insert, also taken from Rose and Gallione (1981) is shown at the bottom.




Sau 3a
Sau 3a
Pst 1



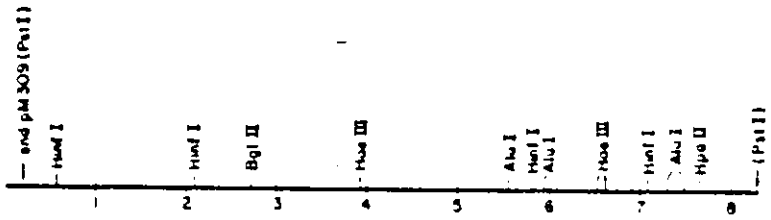


Figure 3.6.2: Nucleotide sequence and predicted protein sequence of the pM309 insert DNA.

This figure was taken from Rose and Gallione (1981). The Bgl II site used in the cloning strategy outline in Figure 3.6.4 is indicated. The bottom figure is the restriction map of the insert. The potential glycosylation site Asn-Ser-Ser is indicated by the asterik.

for the M sequence.

The G sequence (Fig. 3.6.1) reveals a number of restriction sites near the 5' end which could be used to excise the signal sequence coding region. These include the Alu I site at nucleotide position 157 from the 5' end as well as the Hinf I site at position 218. pM309 (Fig. 3.6.2) contains Hinf I sites at positions 59 and 209 on the map. However, none of these combinations would produce an in phase chimeric gene after fusion, even after repair synthesis of protruding ends followed by blunt end ligation. As well, pM309 contains several other Hinf I site which would make isolation of the M coding region difficult.

The coding strategy that was arrived at is based on the fact that the restriction enzyme Sau3A recognizes the sequence $^{\downarrow}$GATC (the sequence is written 5'-3' and the arrow indicates the point of cleavage at the palindromic site) while Bgl II recognizes the sequence A$^{\downarrow}$GATCT. Thus, these enzymes produce cohesive termini which can be easily ligated to each other. pG1 contains Sau3A sites at positions 10 and 324. Thus, this restriction fragment would contain the entire signal sequence coding region as well as the 5' untranslated region required for ribosome binding (Rose, 1980). This fragment would also contain sequences coding for the first 82 amino acids of the mature G protein. pM309 contains a unique Bgl II site located at position 259. Ligation of the Sau3A pG1 fragment with Bgl II cut pM309 would produce an in phase hybrid gene containing the G

specified sequences mentioned above as well as sequences coding for the 157 carboxy terminal amino acids and the entire 3' untranslated region of the M gene.

Since the ultimate goal is to attempt to express this hybrid molecule in a eukaryotic system, a vector suitable for expression of cloned genes in clutured mammalion cells was used as the cloning vehicle. This vector, designated pCVSVe, was kindly provided by Dr. R. Kaufman (Kaufman and Sharp, 1982). It is a derivative of pAdD26-1 which contains sequences coding for dihydrofolate reductase and was used to express this protein in cultured CHO cells (Kaufman and Sharp, 1982). The structure of the plasmid is shown in Figure 3.6.3. This vector includes a number of features that enhance expression of cloned genes. These include the major late transcription promoter of adenovirus as well as the SV40 72 base pair repeated sequence (enhancer sequence). The SV40 72 base repeat has been shown to greatly increase the level of transcription of cloned genes (Banerji et al., 1981). The vector contains SV40 origin and early region sequences as well as the polyadenylation site. Thus, cloned genes can be expressed in COS cells. COS cells are CV-1 (monkey cells) cells transformed with a replication origin defective SV40 DNA molecule (Mellon et al., 1981). These cells, which are permissive for SV40 replication, do not contain a complete copy of SV40 DNA but produce large quantities of large T antigen. Thus, they will support replication of introduced plasmids which contain SV40 sequences that can compliment the defective genome. Therefore, this

MAP OF pCVSVe

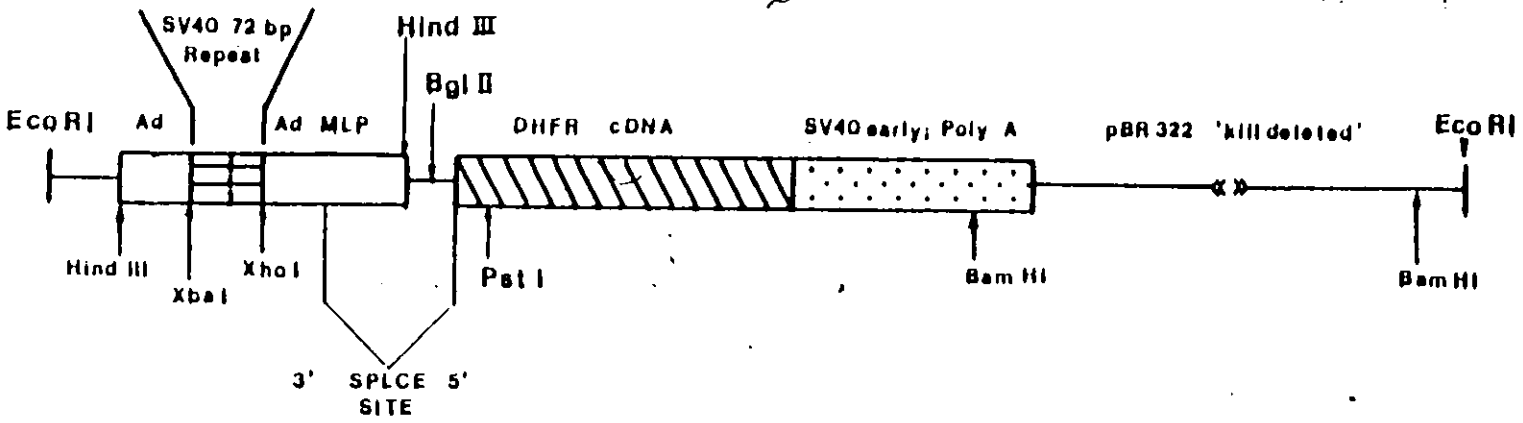


Figure 3.6.3: Structure of pCVSVe

The information shown in this figure was provided by R. Kaufman (MIT). The physical map indicates the positions of the unique Bgl II and Pst I cleavage sites just downstream from the adenovirus major late promoter (Ad MLP) which are useful for cloning DNA fragments into this vector. The SV40 specific sequence contains the replication origin and the early polyadenylic acid addition signal. The pBR322 sequence contains the gene coding for tetracycline resistance and has had the poison sequence deleted.

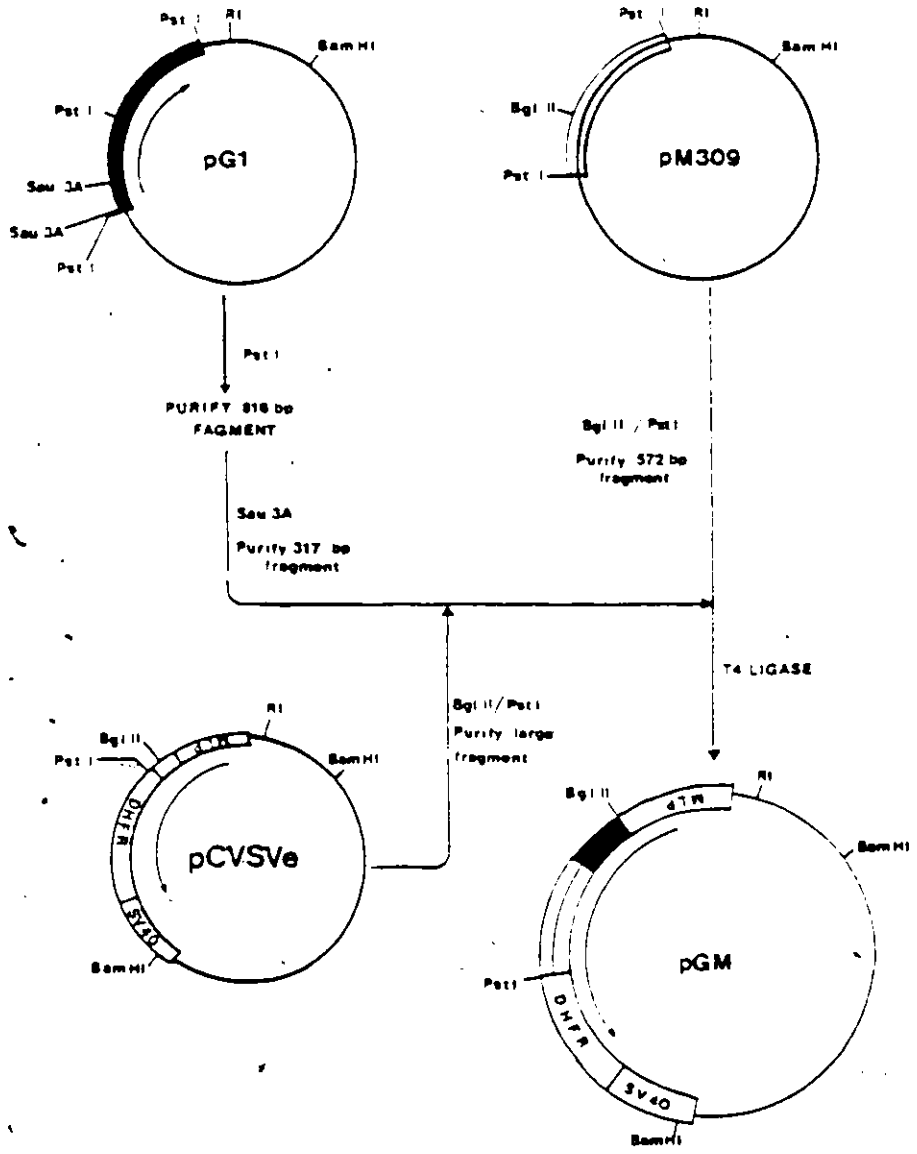
system provides a means of amplifying as well as expressing cloned genes in mammalian cells. pCVSVe can be grown in large quantities in bacterial cells since it contains pBR322 sequences necessary for replication in E. coli. As well, the pBR322 sequences known to inhibit SV40 replication in monkey cells have been deleted (Lusky and Botchan, 1981).

As shown in Figure 3.6.3, pCVSVe contains unique Bgl II and Pst I cleavage sites suitable for insertion of foreign DNA. The fusion of the Sau3A G1 fragment with the Bgl II/Pst I cut pM309 fragment would generate a product having a Bgl II cohesive end at the 5' end and a Pst I cohesive end at the 3' end. Thus, cloning the hybrid gene so constructed into Bgl II/Pst I cut pCVSVe would place the gene in the correct orientation for transcription just downstream from the regulatory sequences of the vector. This cloning strategy is summarized in Figure 3.6.4.

Figure 3.6.5 shows the agarose gel electrophoretic pattern of purified pG1 and pM309 digested with Pst I. The cloning of the cDNA of G mRNA into the Pst I site of pBR322 was done by GC tailing (Rose and Gallione, 1981). In addition to these two Pst I sites, pG1 contains an additional Pst I site located at position 818. Thus, Pst I digestion of pG1 produces two fragments in addition to the pBR322 vector, an 849 nucleotide long fragment derived from the 3' end of the insert and an 816 nucleotide long fragment derived from the 5' end (not including the GC tails). Pst I digestion of pM309 produces a 831 bp (base pair) fragment representing the entire cDNA insert (Fig. 3.3.5 f).

Figure 3.6.4: Strategy used for the construction and cloning of a G-M chimeric gene.

All DNA fragments were purified by polyacrylamide or agarose gel electrophoresis and phenol extracted prior to ligation. MLP, adenovirus major late promoter; DHRF, cDNA sequence coding for the dihydrofolate reductase gene.


Pst I
RI
BamHI
pG1
Sau 3A
Sau 3A
Pst I
Pst I
PURIFY 818 bp FAGMENT
Sau 3A
Purify 317 bp fragment
pM309
Bgl II
Pst I
Bgl II / Pst I
Purify 572 bp fragment
T4 LIGASE
Bgl II/Pst I
Purify large fragment
pCVSVe
Bgl II
Pst I
DHFR
SV40
pGM
47K
DHFR
SV40

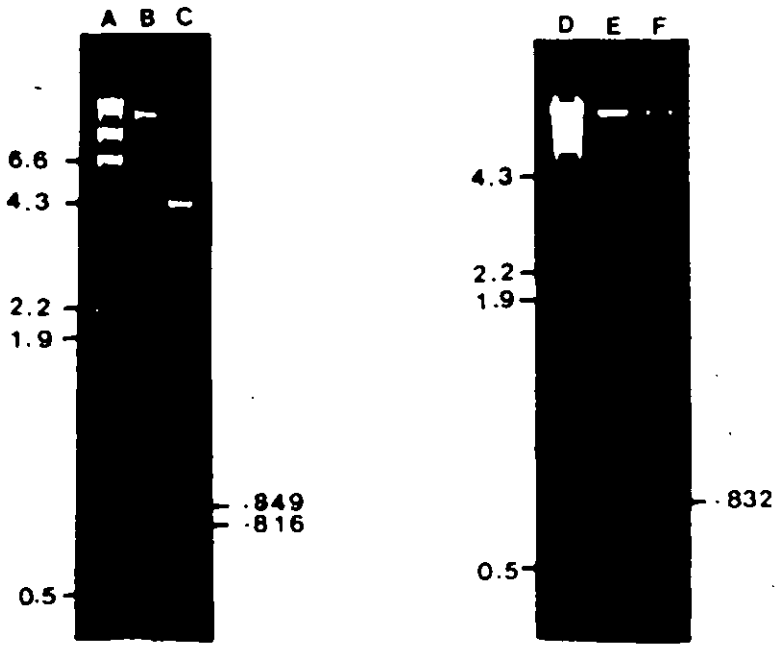


Figure 3.6.5: Pst I digestion of pG1 and pM309.

Plasmid DNAs were purified on CsCl gradients as described in Methods, digested with Pst I, and analyzed on a 1.5% agarose gel. Lanes B and C, pG1 before and after digestion with Pst I, respectively; lanes E and F, pM309 before and after digestion with Pst I, respectively. Lanes A and D are size markers (indicated in kilobases) obtained from Hind III digestion of λ DNA. The fast migrating band in lanes B and E represents the supercoiled form of the plasmid while the slower migrating species represents the relaxed form. The 4.3 kb (kilobase) fragment in lanes C and F represents the pBR322 vector DNA.

The isolation of the 317 bp Sau3A fragment of pG1 was accomplished as follows. pG1 was digested with Pst I, fractionated on a 1.8% agarose gel, and the 816 bp fragment was eluted and purified. This fragment was then digested with Sau3A and the 317 bp fragment was isolated and purified from a 6% polyacrylamide gel. This two step procedure was adopted since pG1 contains many Sau3A sites which would produce several DNA fragments comigrating with the desired one.

The 572 bp Bgl II/Pst I fragment from pM309 was isolated from a 2% agarose gel after digesting pM309 with these enzymes.

The pG1 and pM309 fragments were then ligated with agarose gel purified pCVSVe that had been digested to completion with Bgl II and Pst I. The ligation reaction was then used to transform E. coli LE293 to tetracycline resistance. This transformation resulted in numerous colonies exhibiting the desired tetracycine resistant phenotype.

Cloning into the Bgl II and/or Pst I sites of pCVSVe does not inactivate any selectable antibiotic marker gene thus screening clones which have an insert must be done either by sizing individually purified plasmid preparations on agarose gels or by in situ colony hydridization with a radioactivie probe. Since pCVSVe was cut with two different enzymes which produce dissimilar termini, the probability of the vector itself religating is minimal and, therefore, the vast majority of positive colonies should contain an insert. This, and the fact that purified DNA fragments were used made the possibility of isolating the correct clone by simple

size analysis feasible.

Two types of inserts were expected. Most of the colonies should contain the pM309 fragment ligated into pCVSVe since this ligation would only involve one fragment. These clones would have a size of 572 bp greater than that of Bgl II/Pst I cut pCVSVe alone. The desired clones, occurring at a lower frequency since it would involve the ligation of two separate DNA fragments into the vector, would consist of the Sau3A fragment fused to the Bgl II/Pst I fragment of pM309 and ligated into pCVSVe. These clones would have a size approximately 900 bp larger than Bgl II/Pst I cut pCVSVe.

Approximately 48 Tc^r colonies from this transformation were grown up and plasmids isolated from them. These plasmids were digested with Bam HI and analyzed on a 1% agarose gel. Some of these are shown in Figure 3.6.6. pCVSVe contains two Bam HI sites (see Fig. 3.6.3). Digestion with this enzyme produces two fragments of approximately 3450 and 2700 bp respectively, the large one which contains the Bgl II/Pst I cloning site. Since an approximately 100 bp DNA fragment was removed from pCVSVe for the cloning procedure, the correct clone should produce a 4300 bp fragment as well as the 2700 bp fragment following digestion with Bam HI.

As demonstrated in Figure 3.6.6, several clones displayed fragments of the expected size. These include the plasmids designated pGM 03, 04, 07, 09, 10, 29, 31, 36, and 42. These plasmids produced fragments of approximately 4300 and 2700 bp after Bam HI digestion. Other clones, pGM 12, 13, 14, 16, 17, 18, 19, 34,

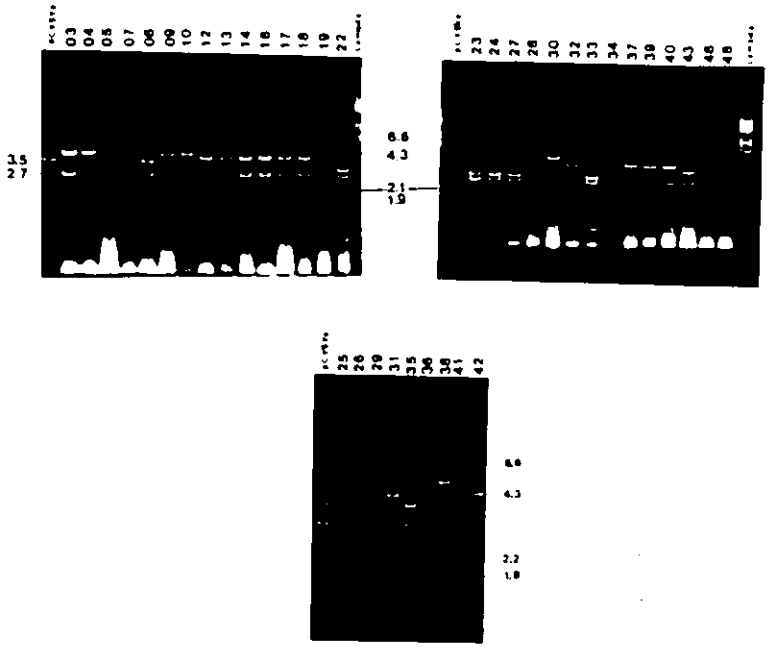


Figure 3.6.6: Digestion of recombinant plasmids with Bam HI.

Ligated recombinant plasmids constructed as described in figure 3.6.4 were used to transform E. coli LE392 to tetracycline resistance. Positive colonies were selected at random from the plates, grown to saturation in 5 mls of L broth containing the selective antibiotic, and plasmids purified by the rapid procedure described in methods. The plasmids were digested with Bam HI and analyzed on a 1% agarose gel. Markers (in kilobases) correspond to the Hind III digestion fragments of bacteriophage λ.

37, 39, 40, and 46 produced fragments of approximately 4000 and 2700 bp following digestion indicating that they contained just the pM309 Bgl II/Pst I fragment as an insert. Approximately 25% of the clones produced a Bam HI pattern not explainable by simple insertion. For instance, pGM 23 seems to have acquired an additional Bam HI site since three fragments were produced, while recombinants such as pGM05 and pGM48 have only one Bam HI site. These may have been produced by recombinational events or from vector plasmids ligating with each other. Still other plasmids such as pGM08, pGM35, and pGM43 appear to represent recircularized vector DNA. Since pCVSVe was linearized by double enzyme digestion, it is possible that a small amount of the linears were only digested with one or the other enzyme. These would recircularize during ligation and produce vector plasmids containing no insert.

Of the 48 colonies screened, 20% produced plasmids containing an insert of the expected size for the G-M hybrid while 34% were of the type indicative of plasmids just containing the M fragment insert.

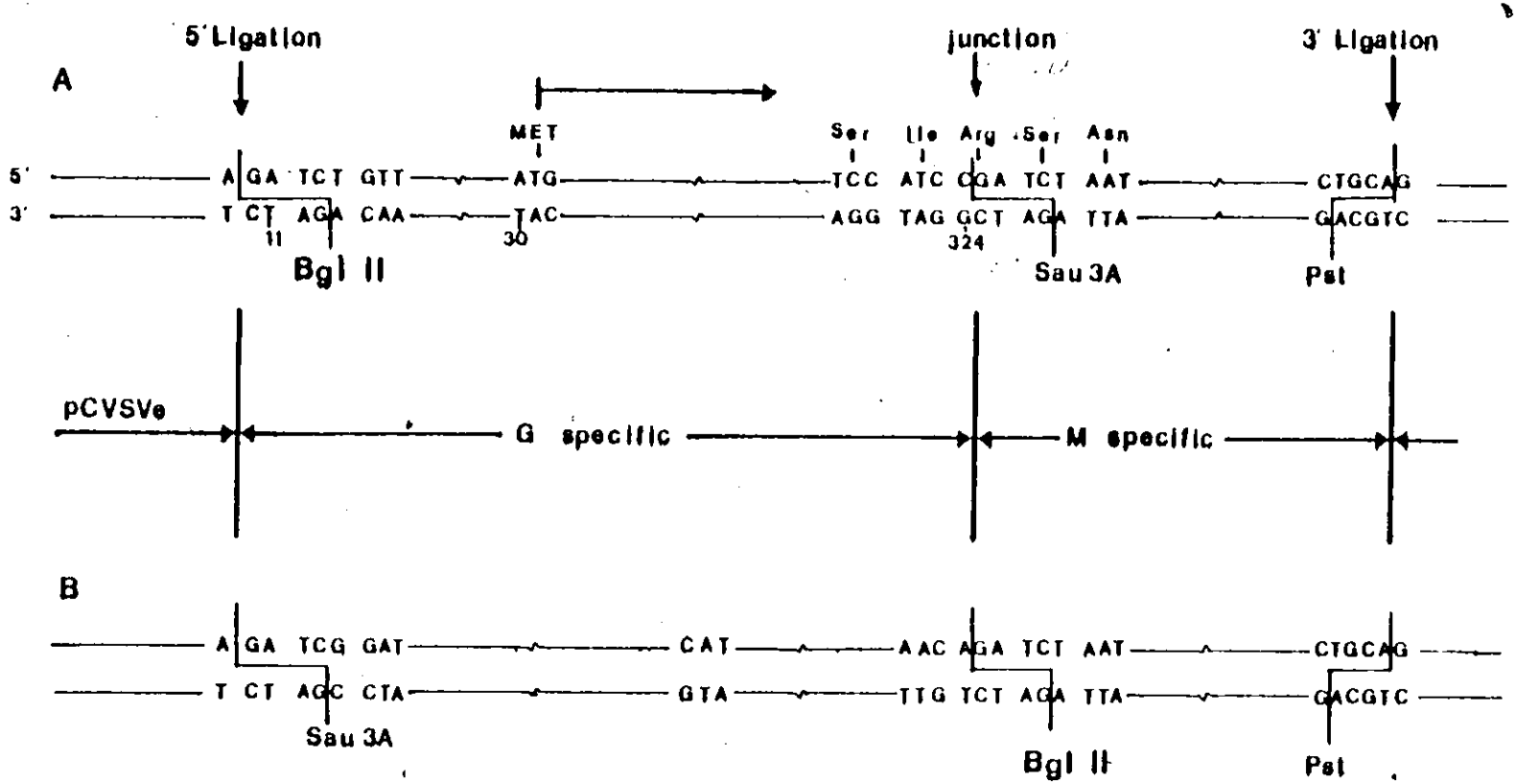
Since the pG1 Sau3A fragment used in the cloning contains homologous termini, it would be present in one of two possible orientations. Since only the orientation that maintains the continuity of the gene is desired, it is necessary to select those GM recombinants that contain the G fragment in the proper orientation.

The following observation, as outlined in Figure 3.6.7, was used to select the proper clone. Fusion of G to M in the correct

Figure 3.6.7: Determination of the orientation of the pG1 Sau3A fragment in the pGM recombinant.

The two possible orientations of the pG1 Sau3A fragment inserted into a pGM recombinant are illustrated in A and B respectively. In the correct orientation, as shown in A, the Bgl II site (AGATCT) is restored at the 5' ligation site. The initiator methionine and the direction of reading is indicated in the drawing. As can be seen, the correct reading frame is maintained at the G/M junction point. Digestion of such a recombinant with Bgl II and Pst I would generate a fragment having a size of approximately 900 bp (317 from pG1 + 572 from pM309).

Insertion of the pG1 Sau3A fragment in the opposite orientation, as shown in B, would generate a Bgl II site at the G/M junction while a Sau3A cleavage site (GATC) would be generated at the 5' ligation site. Thus digestion of such a recombinant with Bgl II and Pst I would produce just the 572 bp fragment derived from pM309.


5' Ligation
junction
3' Ligation
A
MET
Ser Ile Arg Ser Asn
5'
3'
A GA TCT GTT ATG TCC ATC CGA TCT AAT CTGCAG
T CT AGA CAA TAC AGG TAG GCT AGA TTA GACGTC
11
30
324
Bgl II
Sau 3A
Pst
pCVSVe
G specific
M specific
B
A GA TCG GAT CAT AAC AGA TCT AAT CTGCAG
T CT AGC CTA GTA TTG TCT AGA TTA GACGTC
Sau 3A
Bgl II
Pst

orientation and ligation into pCVSVe would regenerate a Bgl II site at the 5' ligation site while retaining the Sau3A site at the G/M junction. If the G fragment is in the opposite orientation, however, a Bgl II site is generated at the G/M junction. Thus, Bgl II/Pst I digestion of pGM clones containing the G fragment in the correct orientation would generate an approximately 900 bp fragment representing the entire G/M hybrid insert. Similar enzymatic digestion of pGM clones containing the G fragment in the opposite orientation would, however, produce a fragment on only 572 bp representing the M specific region.

Figure 3.6.8 shows the restriction pattern of the pGM clones of the proper size and digested with Bgl II/Pst I. As can be seen, pGM 07, 31, 32, and 34 generate a fragment coincident with the fragment generated from Bgl II/Pst I cut pM309. This is expected for pGM 32 and 34 since they were previously judged to contain only the M specific fragment (see Fig. 3.6.6). pGM03, 04, and 36, on the other hand, produced a fragment of higher mobility. Since pGM07 and 31 were shown to contain inserts of the size expected for the G-M hybrid, they must contain the G specific fragment in the incorrect orientation while pGM03, 04, and 36 contain the GM hybrid in the correct orientation.

pGM03 was further characterized by Bgl II/Bam HI digestion to get a more accurate size estimation of the insert. Bgl II/Bam HI digestion of pCVSVe produces three fragments having sizes of 2700, 2500, and approximately 900 bp respectively. The correct pGM clone

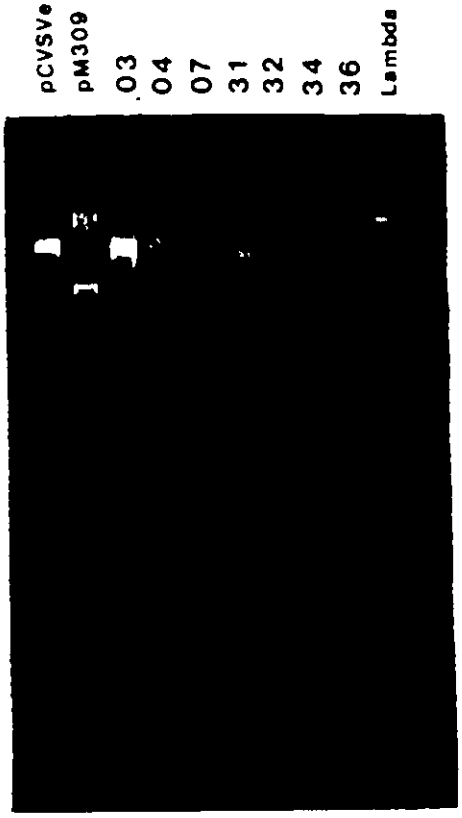


Figure 3.6.8: Analysis of pGM recombinants by digestion with Bgl II and Pst I.

Recombinant plasmids, which were shown to contain an insert of the expected size as determined by Bam HI digestion (Fig. 3.6.6) were digested with Bgl II and Pst I and analyzed on a 1.8% agarose gel. Also shown is the pattern obtained after digestion of pCVSVe and pM309 with these enzymes. The plasmids designated pGM07, 31, 32, and 34, generate a fragment corresponding to the one obtained from pM309 while pGM03, 04, and 36 produce a fragment which has a larger size. The sizes of the Hind III λ DNA markers are indicated.

(900 bp insert) should produce fragments of 2700, 2500, and 1700 bp (since approximately 100 bp was removed from pCVSVe by Bgl II/Pst I digestion prior to cloning). As demonstrated in Figure 3.6.9, this was in fact observed for pGM03. Digestion of pGM03 with Bam HI and Bgl II produced three fragments of sizes 2700, 2500, and 1700 respectively. Digestion of pGM12, which was judged to contain only the M specific sequence, produced fragments of 2700, 2500, and approximately 1400 bp after digestion with these enzymes as expected. Thus, pGM03 was used for further studies.

Expression of Cloned Genes in Mammalian Cells

COS-1 cells were chosen as the host system to attempt expression of the chimeric gene (Gluzman, 1981). T antigen is required for the replication of SV40. Since COS cells provide T antigen in trans, no helper virus is required for the replication of recombinant plasmids, such as pCVSVe, which contain the SV40 origin sequences (Mellon et al., 1981). In addition, it has been demonstrated that plasmids introduced into these cells replicate to a high copy number and that properly constructed foreign genes are transcribed to high levels within 48 hours after transfection (Mellon et al., 1981).

In order to insure that the COS-1 cells that were to be used were expressing the SV40 T antigen, [^{35}S]-methionine labeled cell extracts were prepared and immunoprecipitated with rabbit anti-T antiserum (kindly provided by Dr. M. Breitman, University of Toronto). As shown in Figure 3.6.10, a protein species having a

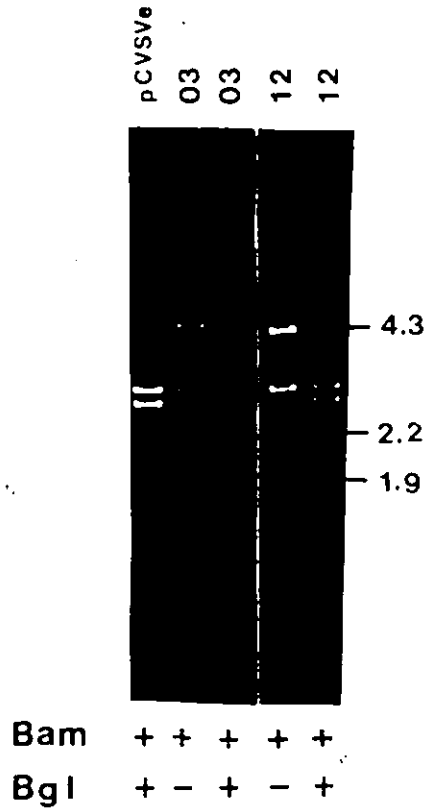


Figure 3.6.9: Digestion of recombinant plasmids pGM03 and pGM12 with Bam HI and Bgl II.

Recombinant plasmids pGM03 and pGM12 were digested with Bam HI or Bam HI + Bgl II and analyzed on a 1% agarose gel. Digestion of pCVSVe with both these enzymes generates three fragments having sizes of 2700, 2500, and 900 bp respectively. Similar digestion of pGM03 generates three fragments of 2700, 2500, and 1700 bp respectively. Digestion of pGM12, which was judged to contain only the pM309 specific insert (572 bp) produces fragments of 2700, 2500, and 1400 bp respectively. Size markers, in kilobases, are from Hind III digested λ DNA.

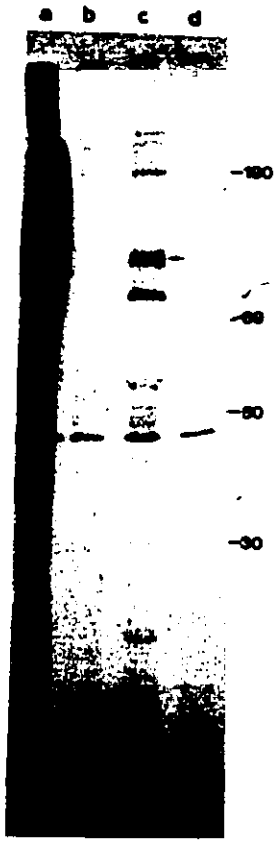


Figure 3.6.10: Detection of SV40 T antigen in COS-1 cells.

COS-1 cells were labeled with [^{35}S]-methionine for 4 hours in methionine deficient Dulbecco's modified MEM and harvested from the plate. The cells were washed three times with cold PBS, lysed with RIPA buffer, and aliquots (1 X 10^6 cpm) were immunoprecipitated with antisera as described in Methods. Samples were analyzed on a 12% polyacrylamide gel. Lane a, sample of whole cell extract; lanes b, c, and d, immunoprecipitates obtained with normal rabbit antiserum, anti-T antiserum, and anti-VSV M antiserum, respectively. The arrow in lane c indicates the position of the 90,000D SV40 T antigen. The molecular weight markers, in kilodaltons, were from a sample of [^{35}S]-methionine labeled VSV run in a separate lane.

molecular weight of approximately 90,000D is specifically precipitated with anti-T. This protein is not reactive with normal serum or antibody specific to the VSV M protein (Fig. 3.6.10, lanes b and c). Large T antigen has a molecular weight of 90,000D (Carrol and Smith, 1976), therefore, these COS cells constitutively express large T antigen.

DNA Mediated Gene Transfer

Introduction of plasmid DNA into the COS cells was carried out by the calcium phosphate transfection procedure (Graham and Van Der Eb, 1973; Wigler et al., 1979). By this procedure, DNA, in phosphate containing buffer, is mixed with $CaCl_2$ which results in a calcium phosphate - DNA coprecipitate. This complex then becomes absorbed to and taken up by cultured cells. Typically, 10-20% of the cells can become transfected by this technique (Chu and Sharp, 1981).

In order to examine the transfection capabilities in this system, a plasmid was required which was known to express detectable levels of a protein after transfection into COS cells. Recently, Sprauge et al., (1983) constructed an SV40 origin plasmid containing the entire cDNA copy of the N mRNA of VSV. COS cells transfected with this plasmid produced detectable levels of authentic N protein as judged by its mobility on SDS polyacrylamide gels and its reaction with viral specific antiserum (Sprauge et al., 1983). It was estimated that N protein represented 0.31-1% of the total proteins synthesized by those cells which were transfected with this plasmid.

The N gene containing plasmid, designated pJS223 was used in this study to determine if the transfection procedure was working before proceeding with the transfection of the chimeric clone. Thus, COS cells were transfected with pJS223 as described in Methods, and labeled cell extracts were prepared and immunoprecipitated. As demonstrated in Figure 3.6.11, lane h, a protein having a molecular weight identical to authentic N protein was specifically immunoprecipitated from pJS223 transfected COS cells reacted anti-VSV antiserum. This protein was not precipitated from pJS223 transfected cells immunoprecipitated with normal rabbit serum (lane g) or from nontransfected cells immunoprecipitated with normal serum or anti-VSV (lanes e and f). Thus, this product is authentic N protein, transcribed and translated only from those cells which were transfected with pJS223. Lane h also shows a high molecular weight protein immunoprecipitated with anti VSV. This probably represents a protein present in the transfected cells which is nonspecifically absorbed with the anti-VSV antiserum. This band is also precipitated with antiserum of different specificities (see below).

Assuming that transfection efficiencies are independant of the type of plasmid DNA, a number of possibilities could be expected upon transfection of cells with the hybrid plasmid pGM03. (1) No protein product is made or the level of synthesis is so low as to make detection impossible. This is a real possibility in that many cloned genes are not efficiently translated in COS cells (Mellon et al., 1981). This may be related to improper positioning of the gene

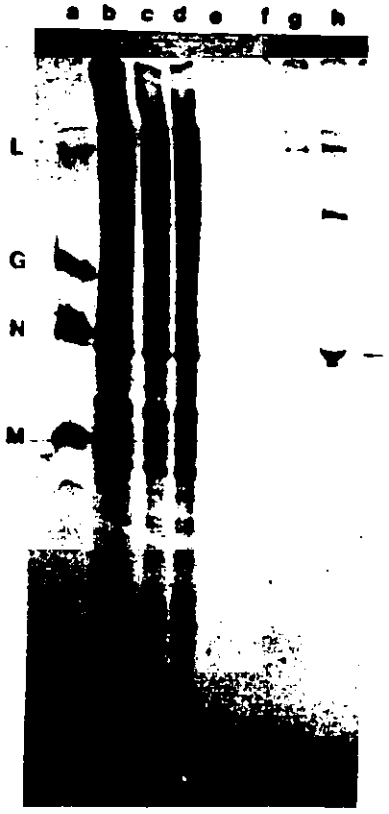


Figure 3.6.11: Detection of VSV N protein in COS-1 cells transfected with pJS223.

Cytoplasmic extracts were prepared from COS-1 cells and COS-1 cells transfected with pJS223, both labeled with [^{35}S]-methionine. Aliquots (approximately 1 x 10^6 cpm) were immunoprecipitated and analyzed on a 12% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV; lane b, sample of COS-1 cell extract; lanes c and d, samples of cell extracts from pGM03 and pJS223 transfected cells, respectively; lanes e and f, COS-1 cell extract immunoprecipitated with normal antiserum and anti-VSV antiserum, respectively; lanes g and h, pJS223 transfected cell extract immunoprecipitated with normal antiserum and anti-VSV antiserum, respectively. The arrow indicates the position of the VSV N protein which was only present in the pJS223 transfected cells.

with respect to the controlling elements, absence of DNA sequences necessary for the transcription and/or translation of this particular gene, or other unknown factors. (ii) The G-M hybrid gene may be synthesized but not inserted into membranes. In this case, the signal sequence would not be cleaved and the product not glycosylated. The complete polypeptide from a G-M hybrid would consist of a protein having a size of 255 amino acids and a calculated molecular weight of approximately 29000D. If we consider the actual molecular weight of the M protein as deduced from the nucleotide sequence (229 amino acids, 26,064D and the apparent molecular weight observed on SDS polyacrylamide gels (30,000D), then the true molecular weight differs from the apparent molecular weight by a factor of approximately 1.15. Thus, the apparent molecular weight of a G-M hybrid that is nonglycosylated and contains the signal sequence would be approximately 33,000D. (iii) The product may be discharged into membrane, have its signal sequence processed, but not glycosylated. In this instance, the product would have an apparent molecular weight of 31,000D if we take into account that the signal sequence of G protein is 16 amino acids long (Irving et al., 1979, Lingappa et al., 1978a). (iv) If the product is glycosylated but the signal sequence is not removed, it would have a molecular weight of 35-37,000D assuming it would contain only one Asn linked oligosaccharide having a molecular weight of approximately 2-3000D. (v) If the product is glycosylated and has the signal sequence removed, it would have an apparent molecular weight of approximately 34-35,000D.

In order to examine if a product was being synthesized from pGM03, COS-1 cells were transfected with this plasmid and labeled cell extracts were prepared and immunoprecipitated with antibody directed against the VSV M protein. As shown in Figure 3.6.12, a protein having an apparent molecular weight of approximately 36000D was precipitated with anti-M from pGM03 transfected cells. This protein was not precipitated from these cells with normal rabbit serum or from pCVSVe transfected cells immunoprecipitated with anti-M or normal serum (lanes e and f). Therefore, this protein appears to be specifically precipitated by anti-M only from cells containing the hybrid gene.

As well, there was a minor protein having a molecular weight of approximately 31-32,000D reacting specifically with anti-M and a larger species having a size of approximately 80,000D (lane h). This larger species is probably the result of nonspecific absorption as described earlier.

These results were confirmed by immunoblotting as shown in Figure 3.6.13. In this case, there were two protein species having molecular weights of approximately 32,000D and 35,000D immunoreactive with anti M which were present in pGM03 transfected cells but not present in cells transfected with pCVSVe or from nontransfected cells. As well, the 80K species, while present in all the samples, was much more apparent in the pGM03 transfected cells. The reason for this is not clear.

Given the size of the 35K species and its reaction with

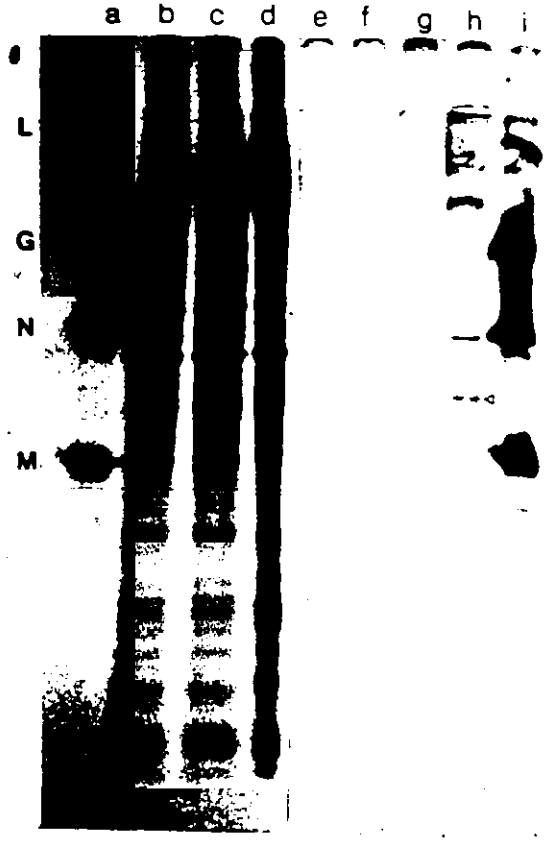


Figure 3.6.12: Polyacrylamide gel analysis of proteins immunoprecipitated from COS-1 cells transfected with pCVSVe or pGM03.

[^{35}S]-Methionine labeled cell extracts prepared from COS-1 cells transfected with pCVSVe or the recombinant plasmid pGM03 were immunoprecipitated and analyzed on a 12% polyacrylamide gel. Lane a, [^{35}S]-methionine labeled VSV; lane b, sample of whole cell extract from COS-1 cells; lanes c and d, whole cell extract from COS-1 cells transfected with pCVSVe and pGM03, respectively; lanes e and f, pCVSVe transfected cells (1 x 10^6 cpm) immunoprecipitated with normal antiserum and anti-M antiserum, respectively; lanes g and h, pGM03 transfected cells (2 x 10^6 cpm) immunoprecipitated with normal antiserum and anti-M antiserum, respectively. The open arrowheads indicate two protein species, having apparent molecular weights of approximately 35 kd and 32 kd respectively, specifically immunoprecipitated from pGM03 transfected cells by anti-M antiserum.

anti-M only in cells carrying the hybrid plasmid, it would appear that it is a product of the G/M hybrid gene. If this is the case, then this protein would be consistent with a G/M hybrid species that is glycosylated and has had its signal sequence removed. However, it is difficult to distinguish between products which have had their signal sequences removed and were either glycosylated or not because of the small molecular weight differences involved. For instance, a product which is glycosylated and has had its signal sequence processed would exhibit a similar molecular weight as a nonglycosylated product which retains the signal sequence. Certainly, if the product is glycosylated it must have been inserted into at least the endoplasmic reticulum where the initial en bloc transfer of the oligosaccharide chains occurs. The 32K species that is reactive with anti M only in pGM02 transfected cells could represent a G/M hybrid protein which is not glycosylated but has had the signal sequence removed.

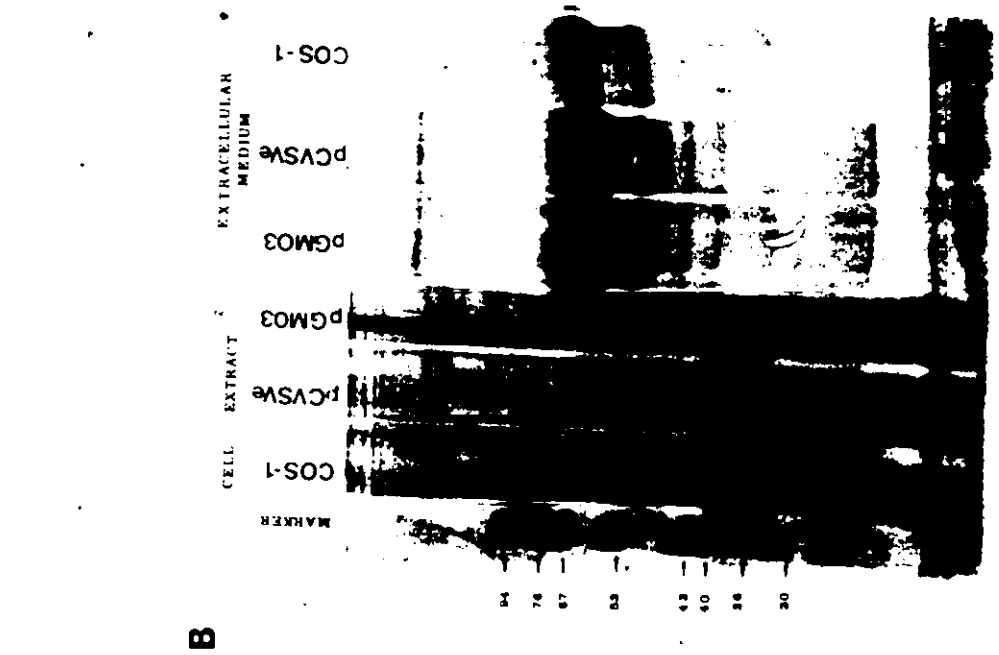
No M specific polypeptides seemed to be secreted by cells transfected with pGM03 as judged by immunoblotting experiments (Fig. 3.6.13). This is not surprising given that M protein has a strong affinity for membranes as demonstrated in Section 3.5. If a G/M hybrid protein is being synthesized and translocated across the endoplasmic reticulum, it may very well become associated with intracellular membranes and thus not be secreted as expected.

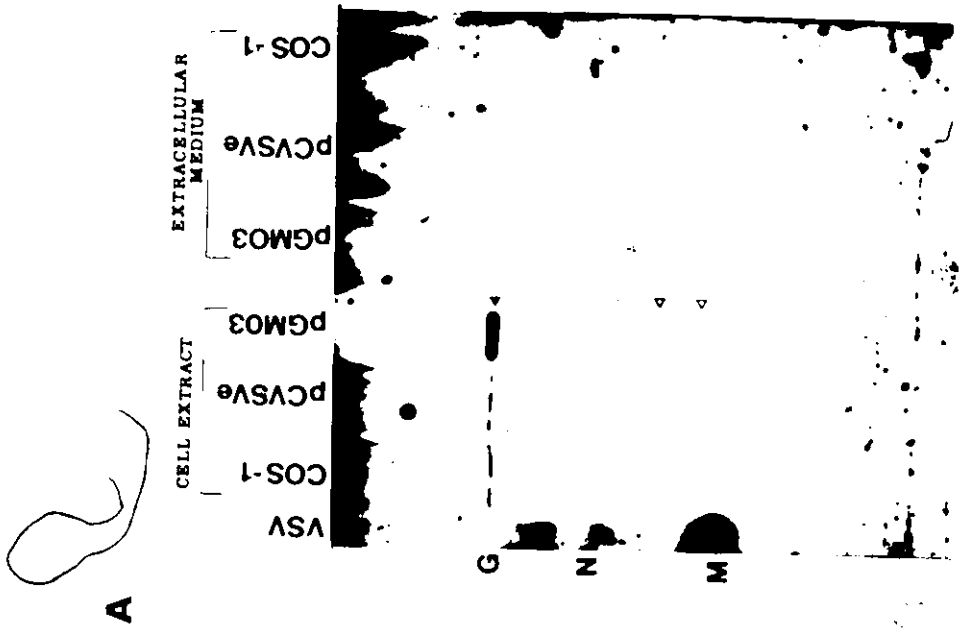
These important questions can be more clearly addressed once an <u>in vitro</u> coupled transcription-translation system is developed.

Figure 3.6.13: Immunoblotting of COS-1 transfected cells with anti-M antibody.

A) COS-1 cells, grown on 60 mm culture plates, were transfected with pCVSVe or pGM03. The cells were harvested 48 hours post-transfection, washed, and lysed with sample buffer. One half of each sample (approximately 1 x 10^6 cells) was run on a 10% polyacrylamide gel. For the analysis of material present in the extracellular growth medium, the medium was collected, precipitated with acetone, and the precipitated proteins were collected, washed, and solubilized in sample buffer. One half of the total precipitated material from each plate was run on the same gel. Following electrophoresis, the proteins were transferred onto nitrocellulose and reacted with anti-M antiserum (1:1000 dilution) followed by reaction with [^{125}I] protein A as described in Methods. Shown here is the autoradiogram from a 1 hour exposure of the nitrocellulose membrane using an intensifying screen. The open arrowheads indicate protein species reactive with anti-M antiserum which are only present in the pGM03 transfected cells.

B) Aliquots of the above samples were run on a parallel gel which was stained with Coomassie blue. This gel included the following molecular weight markers; phosphorylase b (94 kd), transferrin (76 kd), albumin (76 kd), glutamine dehydrogenase (53 kd),ovalbumin (43 kd), aldolase (40 kd), lactate dehydrogenase (36 kd), and carbonic anhydrase (30 kd).


B
MARKER
CELL EXTRACT
EXTRACELLULAR MEDIUM
COS-1
pCVSVe
pGMO3
pGMO3
pCVSVe
COS-1


A
CELL EXTRACT
EXTRACELLULAR MEDIUM
VSV
COS-1
pCVSVe
pGMO3
pGMO3
pCVSVe
COS-1
G
N
M

In this manner, the effect of the presence or absence of membranes on the expression of the G/M chimeric gene can be properly examined.

DISCUSSION

During the course of this work, a great deal of information has accumulated on the structure and assembly of the various components of vesicular stomatitis virus. This has led to a broader understanding, at the molecular level, of the mechanisms involved in these processes.

In this discussion, I will attempt to correlate the observations reported in this thesis with observations of other investigators into a coherent picture of how these findings relate to the structure, morphogenesis, and function of VSV.

4.1 Topology of G Protein In the Viral Envelope

The results presented in Section 3.1 on the orientation of G protein in the viral membrane confirms and extends previous findings regarding the asymmetrical distribution of G protein. That is, the bulk of G protein is external to the bilayer and thus susceptible to proteolytic attack, and that G is anchored to the membrane by a short polypeptide segment (Mudd, 1974; Schloemer and Wagner, 1975a). Schloemer and Wagner (1975a) isolated a short polypeptide fragment (5500 D) from VSV after protease treatment and suggested that this fragment was derived from the amino terminus of G protein and that it represented the membrane interacting domain of the molecule. This conclusion was based on the finding that both the protease resistant fragment as well as mature G protein contained alanine as the amino

terminal amino acid. In contrast, protein sequence analysis of G protein as well as nucleic acid sequencing of the G protein mRNA showed that lysine is the amino terminal amino acid of G protein (Katz et al., 1978; Irving et al., 1979; Rose et al., 1980; Rose and Gallione, 1981; see Appendix).

In addition, Schloemer and Wagner (1975a) reported that the protease resistant fragment obtained from VSV was devoid of tyrosine residues as judged by amino acid analysis. However, it has been determined that both the amino terminal and carboxy terminal proximal regions of G protein contain tyrosine residues (Rose et al., 1980; Rose and Gallione, 1981). In all probability, Schloemer and Wagner (1975a) were observing a heterogenous mixture of small polypeptide fragments. The results put forth in Section 3.1 demonstrate that degradation products, including a small molecular weight fragment (Fragment B) unrelated to G protein, remain associated with spikeless virions following enzymatic treatment.

In contrast to the study reported by Schloemer and Wagner (1975a), the results presented in Section 3.1 demonstrate that G protein is anchored in the membrane by a region at or close to the COOH-terminal end of the molecule. This assignment was based on comparative tryptic peptide analysis of a protease resistant fragment recovered from spikeless virions with intact as well as with carboxypeptidase treated G protein.

The size of the protected fragment (9,000D) suggested that approximately 13% of G protein was protected from proteolysis by

being buried in or in close association with the viral envelope while the remainder was external to the bilayer. This protected region of G protein was shown to contain tryptic peptides common to the carboxy terminal region of G protein that remains exposed to the cytoplasm when G protein is synthesized and inserted co-translationally into membrane vesicles in vitro (H. P. Ghosh, personal communication).

Thus, G protein is anchored in the membrane by a domain near the COOH terminus. The orientation of G protein in the viral envelope is in agreement with the topology observed when G protein is synthesized in vitro in the presence of membranes (Toneguzzo and Ghosh, 1978; Katz et al., 1977; Katz and Lodish, 1979) or in microsommal vesicles isolated from VSV infected cells (Chatis and Morrison, 1979).

Primary structural studies on the membrane interacting region of G protein (Section 3.2) in conjunction with the recently reported nucleic acid sequence of G mRNA (Rose et al., 1980; Rose and Gallione, 1981) has allowed the unambiguous alignment of the protease resistant fragment to the carboxy terminal end of G protein. Thus, it appears that G protein belongs to the same class of membrane proteins as glycophorin (Tomita and Marchesi, 1975), IgM (Rogers et al., 1980), influenza A haemagglutinin (Gething et al., 1978; Min Jou et al., 1980), HLA A2/A7 (Robb et al, 1978; Ploegh et al., 1980), and the Semliki Forest Virus E1 and E2 glycoproteins (Garoff et al., 1980). These proteins are all bound to the membrane by a region near the COOH terminus.

Features of the Membrane Interacting Region of G Protein

The predicted amino acid sequence of the carboxy terminal proximal region of G protein (Rose et al, 1980) reveals that this region exhibits a tripartite domain structure. The most prominant feature of this region is the presence of a contiguous stretch of 20 principally hydrophobic amino acids which presumably represents the region of the protein which actually transverses the lipid bilayer. That a region near the carboxy terminus of G is in fact in intimate contact with the hydrophobic core of the lipid bilayer was demonstrated by the photolabeling studies described in Section 3.4 and dealt with in more detail below. The stability of G protein in the membrane is no doubt partly a function of the favourable hydrophobic interactions between the phospholipids and this highly lipophilic domain of the molecule.

On either side of this membrane transversing region are domains containing hydrophilic as well as charged amino acids. In particular, the 29 amino acids carboxy terminal to the membrane spanning domain contain a high proportion of basic amino acids. This region of the molecule is thought to reside in the interior of the virus particle since its highly basic nature would likely preclude its ability to interact with the hydrophobic lipid bilayer. This would agree with the observation that approximately 3,000D at the extreme carboxy terminus of G protein is susceptible to proteolytic digestion when G protein is synthesized and inserted into membranes in vitro (Toneguzzo and Ghosh, 1978; Katz et al., 1977;

Katz and Lodish, 1979).

The assumption that this region of G protein is actually within the virus is suggested by the size of the protected fragment obtained after proteolysis of VSV. Fragment A was shown to have an apparent molecular weight of approximately 9,000D as judged by polyacrylamide gel electrophoresis. This agrees very well with the actual molecular weight of Fragment A derived from partial amino acid sequencing and alignment with the predicted amino acid sequence. These results suggested that Fragment A encompassed the entire carboxy terminal domain of G protein starting at Asn_{433} (see Section 3.2) and extending to the carboxy terminal lysine. If the carboxy terminus were exposed and therefore susceptible to proteolytic attack, for example by refolding back into the membrane, a smaller protected fragment(s) than the one actually observed would be expected.

Other evidence that G actually spans the viral envelope comes from early crosslinking studies using glutaraldehyde which suggested that interaction can occur between G and the internal M protein (Brown et al, 1974). More recently, evidence that the COOH terminal distal segment does indeed protrude beyond the inner membrane surface comes from studies in which spikeless virions were labeled with ^{125}I by lactoperoxidase or chloramine T iodination after disrupting the viral envelope with pardaxin (Pal et al., 1981). It was found that in the presence of pardaxin, the G tail fragment could be labeled with ^{125}I. The labeling presumably occurred at the

tyrosine located at position 485 and to a lesser extent at the His_{478} and His_{471} residues. However, substantial labeling was also observed in the protected fragment even without prior disruption of the membrane. More recently, labeling of envelope fragments present in spikeless virions with radioactive isethionyl acetamide indicated that a small fragment could be amidinated from the inside of the virus (Taube and Braun, 1982), however, the precise origin of this fragment was not determined.

A model of G protein in the viral envelope based on results presented in this thesis as well as by other investigators is presented in Figure 4.1.

As pointed out by Rose (Rose et al, 1980), the 20 amino acid long nonpolar domain present at the COOH-terminal region of G protein is sufficiently long to span the membrane in an α helical conformation having a residue rise of 0.15 nm (Dickerson and Geis, 1969). The partial amino acid sequence of the membrane anchoring domain of G protein reported in this thesis as well as the partial characterization carried out by Rose (Rose et al., 1980) were among the first examples of direct amino acid sequencing of such a polypeptide segment isolated from a viral glycoprotein. This attests to the difficulty of sequencing highly hydrophobic peptides present in small amounts by conventional sequencing techniques. Recently, Rice et al., (1982) were able to obtain extended sequence information on the hydrophobic membrane embedded domain of the Sindbis E1 and E2 glycoproetins by using an ultrasensitive microsequencing procedure

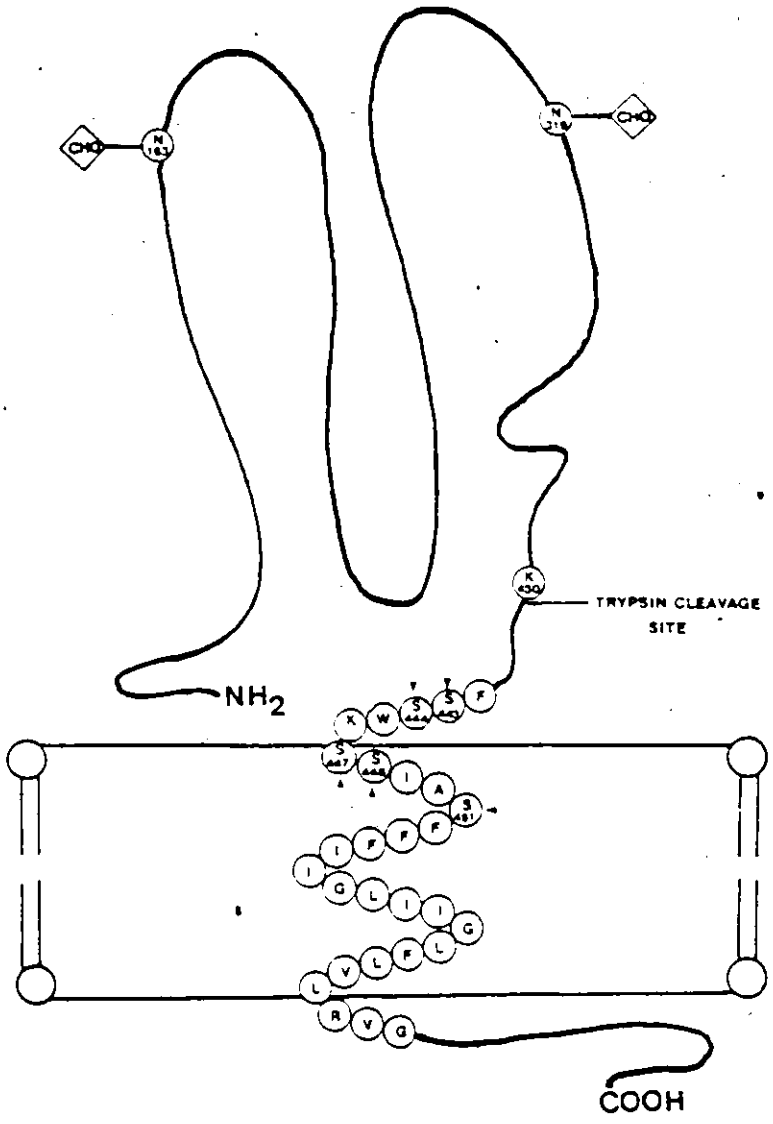


Figure 4.1: Proposed model of G protein in the viral envelope.

This model is based on the results of this work as well as from the predicted primary structure of G protein (Rose and Gallione, 1981). Potential sites of fatty acid acylation are indicated by the arrows and the numbering is from the amino terminal lysine. This diagrammatic representation is not intended to portray any higher order secondary structure.

and computer assisted HPLC analysis of the PTH amino acids.

In the past 1-2 years, extended primary sequencing data has been obtained predominantly by molecular cloning and DNA sequencing analysis, from which amino acid sequences are predicted directly from the corresponding nucleic acid sequence. These new techniques have resulted in a rapid accumulation of data concerning the primary structure of a number of membrane proteins. Figure 4.2 illustrates the amino acid sequence of the membrane interacting domain of a number of such proteins.

The most striking feature is the presence of a contiguous stretch of hydrophobic or nonpolar amino acids (20-45 residues in length) which presumably span the membrane. Another common feature is that the membrane spanning segment is flanked on both sides of the membrane with acidic or basic amino acids. These amino acids may serve to juxdapose the protein in the membrane thereby providing an additional means of maintaining the proper disposition of the protein.

There is no concensus sequence present in the membrane spanning segments that may suggest an important function. The only common feature is the overall high degree of hydrophobicity. This situation is analagous to the amino terminal transient signal sequences present in secretory and membrane proteins. All the signal sequences thus far characterized are rich in hydrophobic amino acids, however, there is little exact sequence homology among them (Sabatini et al., 1982).

MEMBRANE PROTEIN	TRANSMEMBRANE	CYTOPLASMIC	REFERENCE
VSV G PROTEIN	SSHK†SSIASFFFIIGLIIGLFLVL†RVGIHLCIKLKH$_{(aa)17}$		(ROSE ET. AL., 1980)
SFV E_1	QK†ISGGLGAFAIGAILVLVVVTCIGL†RR		(GAROFF ET. AL., 1980)
SINDBIS E_1	K†TSWSMLFALFGGASSLLIIGLMIFACSMWLTST†RR		(RICE AND STRAUSS, 1981)
HAEMAGGLUTININ (FPV)	YKD†VILWESFGASCFLLLATAVGLVFICV†KNGNMRCTICI		(PORTER ET. AL., 1979)
HAEMAGGLUTININ (VICTORIA)	CD†WILWISFAISCFLLCVVLLGFIMWACQ†KGNIRCNICI		(MIN JOU ET. AL., 1980)
μ CHAIN MEMBRANE IgM	FE†NLWTTASTFIVLFLLSLFYSTTNTLF†KVK		(ROGERS ET. AL., 1980)
HAEMAGGLUTININ (AICHI)	KD†WILWISFAISCFLLCVVLLGFIMWACQ†RGNIRCNICI		(VERHOEYEN ET. AL., 1980)
GLYCOPHORIN	EIE†ITLIVFGVMAGVIGTILLISYGI†RRLIKKSPS$_{(aa)27}$		(TOMITA AND MARCHESI, 1975)

Figure 4.2: Amino acid sequences of the putative membrane spanning segments of membrane proteins.

The arrows indicate the borders between the intra and extramembrane domains

Function of the COOH-Terminal Domain

The most apparent function of the COOH-terminal hydrophobic domain of G protein is that it ensures a proper membrane interaction. This has been elegantly demonstrated in recent studies involving the expression of cloned glycoprotein genes in eukaryotic cells. Using a cDNA clone containing the entire coding region of G protein introduced into monkey cells by transfection, it was found that G protein was synthesized, glycosylated, and inserted into the plasma membrane. However, using a clone in which the COOH-terminal coding region was deleted, a truncated form of the G protein was synthesized which was glycosylated but was secreted by the cell (Rose and Bergman, 1982). None of the truncated form of G protein was detected in association with the plasma membrane. Similar results were obtained with the influenza haemagglutinin, demonstrating that cell surface expression requires the carboxy terminal hydrophobic domain (Sveda et al., 1982).

VSV infected cells secrete a soluble glycoprotein antigen, G_s, related to G protein but having a lower molecular weight of 57,000D (Kang and Prevec, 1970; Little and Huang, 1978). Analysis of G and G_s indicate that they contain the same glycosyl residues (Little and Huang, 1978) and the same amino terminal amino acid sequence (Irving and Ghosh, 1982). Tryptic peptide analysis has revealed that G_s is missing the COOH-terminal hydrophobic membrane anchoring domain that is present in G protein. Thus, the absence of this domain results in the formation of a secreted form of the VSV

glycoprotein (Irving and Ghosh, 1982).

The importance of the membrane anchoring domain in the function of membrane proteins has been demonstrated in the case of the middle-sized tumor antigen of polyoma virus. Using a defective viral mutant in which the carboxy terminal region of the middle t antigen was deleted, it was demonstrated that the hydrophobic region was necessary not only for membrane attachment, but also for the associated protein kinase activity as well as for cell transformation (Carmicheal et al., 1982).

The natural system most analagous to this situation is the capacity of β-lymphocytes to synthesize both secretory and membrane associated forms of IgM molecules. It has been demonstrated that the only difference between the secretory and membrane bound forms of the IgM heavy chain is the presence of a hydrophobic COOH-terminal extension in the latter (Vassalli et al., 1979; Kehry et al., 1980; Atl et al., 1980; Cheng et al., 1982).

The results in Section 3.1 demonstrated that G protein is anchored in the membrane by the same region of the molecule by which G is bound in intracellular membranes. This region of the molecule as well as analagous ones present in other transmembrane proteins have been collectively called "stop transfer sequences", reflecting the presumed function that the COOH-terminal hydrophobic domain halts the vectorial transfer of the polypeptide into the rough endoplasmic reticulum which was initiated by the signal sequence (Blobel, 1980). This is thought to occur by dissociation of a

protein tunnel in the RER or by the same lipophilic features of the hydrophobic domain that promote a stable interaction with the phospholipid bilayer (Sabatini et al., 1982). In the case of G protein, this may be facilitated by the high proportion of basic amino acids present in the cytoplasmic portion which would not be expected to be able to penetrate the lipid bilayer.

In contrast, it has been postulated that G protein remains transmembranal simply due to the cessation of protein synthesis (Lodish et al., 1981; Lodish, 1980). It is suggested that the 30 carboxy proximal amino acids remain on the cytoplasmic side of the membrane because they would simply represent that region of the polypeptide remaining in the large ribosommal subunit after protein synthesis is complete. This model does not explain how secretory proteins are completely extruded into the cisternal space of the RER. In addition, the model is inconsistent with the findings that for some transmembrane proteins, the cytoplasmic domain can be as short as two amino acids, as is the case for the E1 glycoprotein of Semliki Forest Virus and Sindbis (see Fig. 4.2).

The other important implication of the finding that the orientation of G protein in the viral envelope is topologically equivalent to that observed in the RER is that this orientation is maintained during the intracellular transport of G protein to the plasma membrane without any re-orientation of the polypeptide. This would support Palade's model of protein migration which involves the fusion and pinching off of protein containing vesicles along the

intracellular membrane pathway (Palade, 1975). Since the fundamental difference between secretory and transmembrane proteins is the presence of a hydrophobic transmembrane domain and a cytoplasmically exposed segment in the latter, the COOH-terminal domain of G protein may serve as a sorting out or addressing signal whereby the cell can recognize this signal and determine its subsequent pathway within the cell.

Amino Terminal Domain

As demonstrated in Section 3.1, G protein was found to be resistant to the actions of aminopeptidase in the absence but not in the presence of detergent. This result may arise from the fact that the amino terminus of G protein is somehow buried in the tertiary conformation of the molecule or that the amino terminus is in close proximity or actually inserted in the lipid bilayer. For instance, Band 3, the anion transport protein present in the red blood cell membrane, appears to be anchored in the membrane by a COOH-terminal domain as well as by another region closer to the amino terminus of the molecule (Rothstein et al., 1977). It is unlikely that the amino terminus of G protein is inserted into the membrane, at least to a significant extent, since a protease resistant fragment derived from the amino terminus of G protein was not detected in spikeless virions. However, it is possible that a small fragment was protected but was, however, beyond the resolving power of the gels used. More convincingly, amino acid sequence analysis of G protein shows that G protein does not contain any significant stretches of hydrophobic

amino acids near the amino terminus that might promote phospholipid attachment (See Appendix; Kotwal et al., in press). The amino terminal region of G protein does, however, contain a clustering of basic amino acids. It is conceivable that these basic amino acids can interact with the phosphate head groups of the lipid bilayer and in this way confer resistance to aminopeptidase.

Recently, the 3 dimensional structure of the haemagglutinin membrane glycoprotein of influenza virus has been solved to a 3 A resolution (Wilson et al., 1981). The structure reveals that the haemagglutinin is anchored to the membrane by a COOH-terminal domain and, in addition, that the NH_2-terminus of the protein is in very close proximity to the external surface of the membrane.

Thus, the VSV G protein may exhibit a similar type of orientation. Work from this laboratory has shown that the G protein from the VSV serotypes New Jersey and Cocal is also protected by the intact envelope from aminopeptidase digestion (H. P. Ghosh, unpublished observations) as is the G protein from the Indiana serotype used in this work. Amino acid sequence analysis of the G proteins purified from these various serotypes reveals that there is a strong degree of conservation in the amino terminal region of these different proteins (see Appendix). This is in sharp contrast to the high degree of non-homology observed in the primary structure of the transient signal sequence of the G protein from these three different serotypes (Kotwal et al., in press).

Thus, there appears to be a strong selective pressure to

resist amino acid changes in the amino terminal region of the mature glycoprotein of VSV. It has been demonstrated that the NH_2-terminal sequences of the HA_2 membrane glycoproteins of different strains as well as different groups of influenza viruses are also highly conserved (Shekel and Waterfield, 1975; Winter et al., 1981; Air et al., 1981; Verhoegen et al., 1980).

The conservation of the amino terminal domain, even in weakly related serotypes, suggests an important but as yet undetermined biological role for this region of the molecule.

4.2 Fatty Acid Acylation of G Protein

Schmidt and Schlesinger were the first to demonstrate that the VSV G protein contains tightly bound fatty acid residues (Schmidt and Schlesinger, 1979). This was based on the observation that labeling VSV with [^{3}H]-palmitate resulted in incorporation of the label not only into the viral phospholipids but also into G protein. The label in G protein could not be removed by denaturation, extraction with organic solvents, or SDS polyacrylamide gel electrophoresis. The label could, however, be released by mild alkaline treatment and recovered as the palmitate methyl ester. This suggested that the fatty acids were covalently attached by ester linkage directly to the polypeptide backbone of G protein. Using this criterion, it has subsequently been shown that a wide variety of viral as well as non-viral membrane proteins contain tightly bound lipid moieties (Table VII).

TABLE VII: FATTY ACID ACYLATION OF MEMBRANE PROTEINS

SOURCE	ACYLATED PROTEIN
VESICULAR STOMATITIS VIRUS[1]	G PROTEIN
SINDBIS VIRUS[2]	E_1 , E_2
SEMLIKI FOREST VIRUS[3]	E_1 , E_2
INFLUENZA VIRUS[3]	HA (HA_2)
NEWCASTLE DISEASE VIRUS[3]	F PROTEIN
SENDAI VIRUS[3]	F PROTEIN
BOVINE CORONA VIRUS[3]	E_2
MURINE HEPATITIS VIRUS A59[3]	E_2
ROUS SARCOMA VIRUS[4]	PP60SRC
ABLESON VIRUS[4]	P120
HARVEY SARCOMA VIRUS[4]	P21
RAUSCHER MLV[5]	P15
MOLONEY MLV[5]	P15
SIMIAN VIRUS 40[6]	LARGE T ANTIGEN
HERPES SIMPLEX VIRUS[7]	GE
HUMAN CELLS[8]	TRANSFERRIN RECEPTOR
BOVINE CARDIAC MUSCLE[9]	cAMP DEPENDENT PROTEIN KINASE
SARCOPLASMIC RETICULUM[10]	CALCIUM ATPASE
B. LICHENFORMIS[11]	PENICILLINASE
E. COLI[12]	MUREIN LIPOPROTEIN

References; 1) Schmidt and Schlessinger, 1979; 2) Schmidt et. al., 1979; 3) Schmidt, 1982; 4) Sefton et. al., 1982; 5) Henderson et. al., 1983; 6) Klockmann and Deppert, 1983; 7) Johnson and Spear, 1983; 8) Omary and Trowbridge, 1981; 9) Carr et. al., 1982; 10) MacLennan et. al., 1973; 11) Neilson et. al., 1981; Lai et. al., 1981; Smith et. al., 1981; 12) Hanke and Braun, 1973.

Localization of Fatty Acids

The results put forth in Section 3.3 demonstrate that all the fatty acid molecules present in G protein are localized exclusively to the membrane anchoring COOH-terminal domain. Proteolytic digestion of [^{3}H]-palmitate labeled VSV resulted in the ^{3}H label remaining associated with Fragment A. The label could be released from Fragment A by transesterification and was identified by TLC as the methyl ester derivative of palmitate.

The presence of fatty acids linked to Fragment A helps explain a number of observation made with this peptide. According to the amino acid sequence reported in Section 3.2 and the alignment with the COOH terminal sequence of G protein, the protected fragment should have a molecular weight of approximately 7,500D based on the site for trypsin cleavage. Fragment A had, however, an apparent molecular weight of approximately 9,000D as judged by polyacrylamide gel electrophoresis. The differences in the observed and predicted molecular weight may be due to the bound fatty acid residues. The fatty acids, in addition to increasing the molecular weight of Fragment A (by approximately 255D/palmitate residue) may also result in aberrant or non-uniform binding of SDS, thereby resulting in anamolous electrophoretic behaviour.

It has subsequently been demonstrated that a number of other viral membrane glycoproteins contain covalently linked fatty acid residues within their membrane interacting domains. These include the influenza haemagglutinin (Schmidt, 1982a), and the Sindbis virus

E1 and E2 glycoproteins (Rice et al., 1982). As well, the transforming protein pp60src of Rous Sarcoma virus (Sefton et al., 1982) and the transferrin receptor from cultured cells (Omary and Trowbridge, 1981) have been shown to contain covalently linked fatty acid residues in their membrane interacting regions.

Despite intensive effort, the actual linkage site(s) of the fatty acid residues is not known for any glycoprotein. This is largely due to the unusual properties of the acylated peptides which makes localization of fatty acid residues by direct sequence analysis difficult (Schmidt, 1982b; Magee and Schlesinger, 1982).

Schlesinger has suggested that in the case of G protein, the fatty acids may be esterfied directly to serine residues (Schmidt and Schlesinger, 1979). This observation was based on the finding that a fatty acid containing peptide recovered from a proteolytic digestion of [^{3}H]-palmitate labeled G protein contained a high proportion of serine residues. However, esterfication to the side chains of other hydroxy amino acids such as threonine and tyrosine cannot be ruled out.

Amino acid sequence analysis of the membrane interacting domain of G protein reveals that there are five serine residues located within the presumed transmembrane segment or just outside the membrane surface that could serve as fatty acid acceptor sites (Fig. 4.1). These include Ser 443, 444, 447, 448, and 451. Acylation of serine residues within the transmembrane segment would serve to render these residues more hydrophobic, thereby fixing

the protein more tightly within the membrane.

Other potential esterification sites include the threonine residues at position 479 and 486 and the tyrosine residue located at position 485. These latter residues are unlikely acylation sites based on the carboxypeptidase digestion experiments reported in Section 3.3. It was found that [^{3}H]-palmitate labeled G protein retained the fatty acid residues even after approximately 2,500D was removed from the carboxy terminus by digestion with carboxypeptidase. This size reduction of G protein would result in the loss of the hydroxy amino acids reported above, thus, the fatty acids residues must be confined within the transmembrane domain or within the external portion not degraded by proteases.

The peptide mapping data described in section 3.3. suggests that the fatty acids are located at the serine 443 and/or 444 residues. This was based on the quantitation of radioactivity present in tryptic peptides of amino acid as well as fatty acid labeled Fragment A. Although the assignment is at best tentative, it is supported by the observation that a small fatty acid labeled peptide containing an amino terminal phenylalanine residue could be recovered from [^{3}H]-palmitate labeled G protein after tryptophan specific cleavage (Schlesinger et al., 1981). This would place the fatty acids at the serine 443 and 444 positions. The precise boundary of the transmembrane domain is not known with certainty, thus, it is possible that the Ser 443 and 444 residues which are depicted as being located external to the surface of the bilayer

in Figure 4.1 could, in fact, be within or in close association with the bilayer.

It should be noted that many acylated proteins contain hydroxy amino acids, in particular, serine residues, within their membrane anchoring domains as well as in the extracytoplasmic segment (see Table VII). However, not all fatty acid containing proteins contain esterified lipid in their carboxy terminal regions. For instance, pp60src, the transforming protein of RSV, contains fatty acid apparently in amide linkage to an amino terminal region of the polypeptide, the region of the molecule which is thought to be involved in membrane association (Sefton et al., 1982). Recently, the membrane associated p15 protein of both Rauscher and Moloney leukeamia virus has been shown to contain a myristyl fatty acid ($CH_3(CH_2)_{12}COOH$) amidated directly to the amino terminal glycyl residue (Henderson et al., 1983). The murine lipoprotein of E. coli contains a fatty acid amidated to the α amino group as well as fatty acids esterfied to a glycerol moiety that is in thioether linkage directly to the amino terminal cysteine (Hanke and Braun, 1973).

Function of Covalently Attached Fatty Acids

The biological function of the covalently linked fatty acids is not clear, however, a number of roles can be suggested. The most obvious role, in lieu of any experimental data, for the presence of fatty acids in the membrane interacting region of G protein is that they enhance the lipid solubility of this region of the molecule in the lipid bilayer and thus augment the forces that stabilize

G protein in the membrane. It has been demonstrated recently that the lipid free neuraminidase of influenza and the neuraminidase-haemagglutinin of New Castle Disease (NDV) virus are inefficiently incorporated into liposomes whereas the fatty acid acylated F protein or the haemagglutinin of Fowl Plaque virus can be incorporated into liposomes much more readily (Huang et al., 1979). In addition, the normally non-acylated IgG, which has no affinity for liposomes, was found to incorporate into liposomes following chemical acylation (Huang et al., 1980).

Certainly not all viral glycoproteins, even within the same virus, require fatty acid residues. In the case of Bovine Corona virus and Mouse Hepatitis A59, which contain two spike glycoprotein species E1 and E2, only E2 was shown to contain fatty acid (Schmidt, 1982a). Similarily for NDV and Sendai virus, only the F protein and not the HN protein was found to be acylated (Schmidt, 1982a).

Schmidt (1982b) has suggested that the fatty acids present in viral glycoproteins are involved in membrane fusion. For instance, the F protein of Sendai, the HA_2 of influenza, and the Sindbis virus E1 and E2 have strong fusinogenic properties and it is thus conceivable that the fatty acids may promote fusion between closely associated membrane bilayers. In the case of VSV G protein, this activity may be important in the uptake of the virus. This is intriguing in light of the recent observation that phospholipid may be an important part of the VSV cell surface receptor (Schlegel et al., 1983). This hypothesis is, of course, not applicable to those

fatty acid containing proteins that do not have fusinogenic activity.

It appears that the membrane affinity of certain proteins is modulated by the presence or absence of fatty acids. In the case of the RSV pp60src protein, which can exist as a cytoplasmically soluble form or in association with the plasma membrane, it has been demonstrated that only the plasma membrane associated form contains covalently bound lipid moieties (Garber et al., 1983). This may be an important observation considering the fact that pp60src is a transformation specific protein which is thought to exert its effects on the plasma membrane (Courtneidge et al., 1980). Similarily, the large T antigen of SV40 is largely present in the nucleus but a small amount is associated with the plasma membrane (Soule and Butel, 1979). Recently, the plasma bound form but not the nuclear form of the SV40 large T antigen has been shown to contain covalently linked fatty acids (Klockman and Deppert, 1983). Since the large T antigen does not contain any long stretches of nonpolar amino acids that would promote membrane association, it is possible that the fatty acids provide this function. In the case of the VSV G protein, the fatty acids appear to be added after the protein is inserted into the RER (Schmidt and Schlesinger, 1980) thus, the fatty acids are probably not required for inherent membrane affinity, however, this affinity may be augmented once the fatty acids are attached.

The essentiality of fatty acid acylation, even in the case of G protein, is not clear. Results from this laboratory indicate that while the G proteins from the VSV serotypes Indiana, Chandipura, and

Piry contain covalently linked fatty acid residues, the G proteins from the Cocal as well as three different strains of the New Jersey serotype do not (Kotwal and Ghosh, unpublished observations). This implies that fatty acid acylation is not critical for the maturation and function of all VSV subtypes. It cannot, however, be completely ruled out that these different G proteins contain fatty acids other than palmitate, or have some other type of modification that provides an analogous property. The finding that not all G proteins are acylated indicate that acylation is not a random event and that it must be specified by information present in the polypeptide backbone.

The lack of fatty acid acylation in the case of New Jersey G protein (Kotwal and Ghosh, unpublished observations) is especially interesting in light of the recently published amino acid sequence of this molecule (Gallione and Rose, 1983). The sequence of the membrane anchoring region of the New Jersey G protein shows extensive homology with that of Indiana G protein, however, the serine residues present in position 444 and 451 of Indiana G are changed to glycine residues in New Jersey G protein. Thus, the observed lack of fatty acid labeling of New Jersey G protein may be a direct result of the lack of these esterification sites.

It has been suggested that fatty acid acylation of G protein may serve as a marker for the intracellular transport of this protein to the plasma membrane. Pulse labeling experiments have demonstrated that fatty acid acylation is a post-translational modification that

occurs approximately 10-15 minutes after completion of protein synthesis and about 1-2 minutes before trimming of the high mannose oligosaccharides (Schmidt and Schlesinger, 1980). This indicates that fatty acid acylation occurs before or shortly after G protein reaches the Golgi complex. Monensin, an ionophore which results in the blockage of the transport of G protein from the Golgi complex to the plasma membrane had no effect on fatty acid acylation (Johnson and Schlesinger, 1980). Rothman and co-workers, using a reconstructed in vitro system capable of processing G protein have provided evidence that acylating activity is present in transitional elements of the endoplasmic reticulum or in the early acting Golgi membranes (Dunphy et al., 1981). It has recently been demonstrated for Semliki Forest virus that fatty acid addition occurs in the cis or medial cisternae of the Golgi complex (Quinn et al., 1983; Griffiths et al., 1983).

The observation that the RSV pp60src and the SV40 large T antigen, proteins that are synthesized in the cytoplasm on free ribosomes, contain covalently linked fatty acid residues indicates that fatty acid acylation is not restricted to those proteins which are transported to their final destination via the Golgi complex.

Using a temperature sensitive mutant of VSV that was defective in fatty acid acylation of G protein, Lodish and co-workers have shown that the intracellular transport of G protein from the Golgi complex to the plasma membrane was inhibited in the absence of acylation (Zilberstein et al., 1980). It was suggested that the

fatty acids were required for the proper signaling of G protein to the plasma membrane, however, the mutant glycoprotein has a number of other temperature sensitive defects, including aberrant glycosylation, that may have contributed to the observed effects. Recently, using more refined detection techniques, it has been demonstrated that a small amount of the mutated G protein from this virus does, in fact, reach the plasma membrane although it was not demonstrated whether these plasma membrane bound molecules were acylated or not (Lodish and Kong, 1983).

Schlesinger and Malfer (1982) have examined VSV particle formation in the presence of cerulenin, an antibiotic shown to inhibit fatty acid acylation of G protein. In the presence of cerulenin, G protein was not acylated and VSV particle production was reduced by up to 90%. Non-acylated G protein was found to accumulate on the cell surface but was not incorporated into budding virus. This suggests that while acylation is not required for G protein migration it may play a role in virus assembly and budding, possibly by providing an interaction with other viral components. For instance, it is possible that the fatty acids may alter the conformation of G protein and thereby establish new sites for protein-protein or protein-lipid interactions. These results must be interpreted with caution since cerulenin is known to have a wide effect on a number of cellular functions, including the inhibition of protein and RNA synthesis (Goldfine et al., 1978).

It is evident from all the above studies that the function of

covalently linked fatty acid residues cannot be generalized to include each case and that it may be necessary to examine features particular to individual types or classes of membrane proteins. Whatever the function, fatty acid acylation probably plays some important biological role given the widespread occurrence of this modification.

4.3 Membrane Organization of the VSV G and M Proteins

The development of photoreactive fatty acids has provided an elegant approach to the study of the nature of membrane organization. By placing the reactive group in the hydrophobic core of the membrane, one can identify those proteins that are in intimate contact with the bilayer as well as examine the nature of protein-protein and protein-lipid interactions (Gupta et al., 1977; Gupta et al., 1979; Brunner et al., 1980; Quay et al., 1981; Rose et al., 1982). The potential and applicability of this approach has been greatly extended by the synthesis of highly radioactive fatty acid derivatives containing the diazirine photoreactive group (Leblanc et al., 1982).

The results presented in Section 3.4 described the use of one such compound, [^{3}H] diazirinophenoxy nonanoate, to examine the protein topography of the VSV viral membrane as well as to produce a photoaffinity derivative of G protein.

Biosynthetic Incorporation of [^{3}H] DAP-nonanoate into VSV

The results described in Section 3.4 demonstrated that

[^{3}H] DAP-nonanoate was efficiently incorporated into both viral lipids as well as G protein. Thus VSV can utilize this fatty acid derivative to acylate G protein and, as well, the virus particle is able to bud from the plasma membrane containing this modified fatty acid. No other viral protein was detectably labeled by DAP-nonanoate, indicating that the observed incorporation into G protein was not due to the metabolic conversion of the fatty acid into protein synthetic precursors.

The immediate consequence of this observation is that it is possible to biosynthetically incorporate a fatty acid other than palmitate into G protein. Thus, G protein is able to accept fatty acids of various chain lengths and, therefore, the chain-length of the fatty acid may not be critical for its biological activity.

In a related study, Petri et al. (1981), using VSV infected BHK cells, were able to biosynthetically incorporate the fluorescent fatty acid 16-(9-anthroloxy) palmitate into G protein but were, however, unable to label G protein with shorter chain length fluorescent fatty acids such as anthracene-9-proprionic acid or anthracene-9-decanoic acid. Although the fluorescent fatty acid incorporated into G protein was not characterized, they suggested that only longer chain fatty acids could be incorporated in vivo into G protein. In contrast, the results presented in Section 3.4 clearly demonstrate that fatty acids with chain lengths as short as nine carbons can still be recognized by the acylating machinery and incorporated into G protein. The differences reported may reflect

differences in the chemical structures of the fatty acid derivatives. The fluorescent anthroloxy derivative was linked near the middle of the fatty acid molecule (at the C9 position) whereas the photoreactive diazirine was coupled to the end of the molecule. Thus, the position of the photoreactive group may, in fact, result in a recognizable increase in the effective chain length of the fatty acid.

Of the total radioactivity incorporated into VSV, approximately 1% was tightly bound to G protein. This is in comparison to the level of approximately 2% when VSV was grown in the presence of radioactive palmitate. The differences in the level of labeling could be due to the fact that the fatty acid derivative is not as efficiently recognized as the normal C16 fatty acid or that only a percentage of the G protein molecules become acylated with the photoreactive fatty acid.

By all criteria examined, the DAP nonanoate fatty acid behaved in a manner analogous to palmitic acid. All of the [^{3}H] DAP-nonanoate incorporated into G protein was present in the membrane interacting domain of the molecule. The label present in G protein could not be removed by boiling in SDS, polyacrylamide gel electrophoresis, treatment with DMSO, or exhaustive extraction with organic solvents, but could be quantitatively released by treatment with hydroxylamine or by mild alkaline hydrolysis. Thus, the DAP nonanoate appears to be covalently attached by an ester linkage directly to the polypeptide backbone of G protein.

This was confirmed by transesterification of the [^{3}H] DAP-nonanoate labeled G protein and identification of the released radioactivity by HPLC. Most of the radioactivity migrated as the methyl ester derivative of DAP nonanoate, however, significant amounts of longer chain length species were also present. These included the methyl ester derivatives of DAP undecanoate, tridecanoate, and pentadecanoate. While the elution profile demonstrated that the photoreactive diazirine group was retained after biosynthetic incorporation, it also directly demonstrated that DAP-nonanoate could be metabolically converted to longer chain length species by the cell prior to incorporation into G protein. In agreement with this observation, it has been reported that fatty acids other than palmitate, such as steareate and oleate, are present in G protein (Schmidt et al., 1979).

VSV Membrane Topography

(a) Phospholipid-Protein Interaction

Photolysis of [^{3}H] DAP-nonanoate labeled VSV resulted in a tenfold increase in the amount of label present in G protein as compared to the amount of radioactivity present prior to irradiation. In a similar type of study, photolysis of VSV grown in the presence of 16-azido-[9,10 - ^{3}H] palmitic acid also resulted in extensive labeling of G protein (Stoffel et al., 1978). This increase resulted from the crosslinking of surrounding DAP containing phospholipids to the G protein following photoactivation. This photoinduced increase in labeling of G protein was restricted to

the hydrophobic membrane interacting region of the glycoprotein.

Thus, the fatty acid probe, incorporated into the viral phospholipids, specifically labels the hydrophobic core of the viral envelope. This result confirms and extends previous findings using exogenous probes to show that G protein is an integral membrane protein and in addition, unambiguously demonstrates that the COOH terminal domain of G is in intimate contact with the hydrophobic core of the lipid bilayer.

The hydrophobic aryl azide probes [^{125}I]-iodonaphthyl-1-azide and [^{3}H] pyrenesulfonylazide were shown to partition into the VSV viral envelope and label the membrane interacting region of G protein (Zakowski and Wagner, 1980). However, in certain cases, internal components such as the N protein also became labeled upon photolysis indicating that these probes were not reacting exclusively with hydrophobic regions.

Recent studies have indicated that photoreactive compounds containing nitrene precursors (such as the aryl azides described above) are not suitable as reagents for labeling hydrophobic regions in the bilayer (Bayley and Knowles, 1978a, 1978b; Chowdhry and Westheimer, 1979). This is due to the low reactivity and long half life of the photolytic intermediates which results in a low yield of crosslinking. In addition, these compounds undergo a high degree of intramolecular rearrangement which results in the production of reactive byproducts upon photolysis. Of more concern is the observation that the photoactivated species can be almost completely

scavanged by externally added thiols such as glutathione, indicating that the probes are not specifically reacting with the hydrophobic core of the membrane.

In contrast, carbene generated precursors, such as diazirine, are much more suitable for membrane studies since they are much more reactive, especially to chemically inert polypeptide regions, the efficiency of crosslinking is substantially higher, and they cannot be scavanged by externally added thiols (Bayley and Knowles, 1978a, 1978b). In addition, the reactive species has an extremely short half life. The studies with the L cells described in Section 3.4 indicated that crosslinking was essentially complete within 2 seconds, the shortest time interval examined. As such, the diazirine containing photoreactive probe provides a much more accurate representation of the VSV membrane structure.

M Protein

In contrast to the heavy irradiation dependant labeling observed with G protein, the M protein was not labeled to an observable extent following photolysis. This demonstrates the high degree of specificity displayed by the photoreactive probe in that only those proteins that are in direct contact with the hydrophobic core of the membrane undergo phospholipid-protein crosslinking following irradiation. More importantly, this result directly demonstrates that the M protein does not penetrate the phospholipid bilayer to a significant extent, if at all.

Studies utilizing exogenous probes have indicated, however,

that the M protein must be in close association with at least the phospholipid head groups present on the inner monolayer since the M protein can be crosslinked to phospholipids with bifunctional reagents that are 1 nm in length (Pepinsky and Vogt, 1979). More recently, it has been demonstrated that while the hydrophilic crosslinker tartyl diazide can crosslink phospholipids to M protein the hydrophobic reagent 4,4' dithiobisphenylazide does not (Zakowski and Wagner, 1980).

Thus, the M protein, although in close association with the lipid envelope, does not penetrate into the hydrophobic region of the bilayer.

Protein-Protein Crosslinking

A critical aspect of membrane organization is the nature of protein-protein interactions. By incorporating a photoreactive fatty acid *in vivo* into the VSV G protein we have, in effect, produced a photoaffinity derivative of a specific integral membrane protein. Thus, one can examine the spatial relationships of G protein with itself and with other components present in the membrane.

Photolysis of [^{3}H] DAP-nonanoate labeled VSV resulted in the appearance of several discrete new protein species having apparent molecular weights both higher and lower than G protein (compounds I-VI, Fig. 3.4.11).

The presence of these irradiation dependant species may be explained in several ways. (i) They may be bona fide protein-protein crosslinks mediated by the fatty acid attached to G protein.

(ii) They may represent minor species present in the viral envelope that undergo phospholipid-protein crosslinking. (iii) They may represent minor contaminating, protein containing membrane fragments co-purifying with VSV during isolation. This latter possibility is unlikely since the virions were purified in the presence of salt and the photolytic pattern was consistent with different virus preparations.

Immunoblotting was used to distinguish between the first two possibilities. Compound I (140,000D), based on its molecular weight and its specific reaction with anti-G antibody, was shown to be a G-G homodimer. Its presence was dependant on prior photolysis. The formation of this dimer must be the result of crosslinking mediated by the photoreactive fatty acid attached to G protein in vivo since the diazirine containing fatty acids are incorporated into only one of the positions of glycerol backbone during phospholipid synthesis (G. D. Gerber, personal communication). In contrast, Stoffel et al. (1978) claimed that G protein aggregates produced upon photolysis of VSV grown in the presence of 16-azido-[9,10-^{3}H]-palmitic acid were the result of crosslinking mediated by phospholipids which were presumably disubstituted with the azido labeled fatty acid. As explained above, this would not be the case with the diazirino fatty acid. In addition, the extremely short half life of the carbene precursor, as opposed to the nitrene precursor, would further reduce the likelihood that G-G crosslinking was mediated by disubstituted phospholipid molecules.

The crosslinking event that produces the G-G homodimer must be occurring in the hydrophobic region of the bilayer since if the photoreactive group was positioned external to the bilayer it would be rapidly scavanged by water molecules (G. E. Gerber, personal communication).

The intensity of the G-G dimer (approximately 2-3% relative to the [^{125}I] labeling of the G protein monomer), the short half life of the carbene precursor, and the rigidity of the viral envelope (Lansberger and Compans, 1976; Landsberger and Aitstiel, 1980; Aitsteil and Landsberger, 1981) precludes the possibility that the existence of the G-G dimer is the result of collisional interactions occurring in the membrane. Petri et al. (1981) incorporated a fluorescently labeled fatty acid *in vivo* into G protein and examined the fluorescence anisotropy of this G protein incorporated into artificial phospholipid vesicles. The results indicated that the lipids surrounding the membrane interacting region of G protein are relatively immobile. Thus, the G protein would be expected to have little lateral mobility in the viral envelope.

The conclusion from the crosslinking studies with the photoreactive fatty acid then is that the VSV spike is composed of more than one molecule of G protein. This conclusion is supported by several studies that have used exogenously added bifunctional crosslinking reagents (Mudd and Swanson, 1978; Dubovi and Wagner, 1977; Zakorski and Wagner, 1980). Dubovi and Wagner (1977) have suggested in fact that the G protein may exist in the form of trimers

as well as dimers. However, it could not be completely discounted whether the dimerization or trimerization represented true interactions since high concentrations of exogenously added crosslinking compounds can cause non-specific protein aggregation (Peters and Richards, 1977). In contrast, there are no concentration dependant artifacts with the photoreactive fatty acid since the probe is incorporated intrinsically during virus assembly.

Several other enveloped animal viruses have been shown to contain surface glycoproteins in oligomeric structures. These include SVP (Garoff, 1974), influenza (Laver, 1971), and Newcastle Disease virus (Scheid and Choppin, 1973).

Whether G protein may exist in the intracellular membrane in an oligomeric form is not known. It is possible that this association may be required for the migration of G protein to the plasma membrane. This is intriguing in light of the observations that some membrane proteins require associated proteins for their transport. For instance, the transport of HLA heavy chain requires the presence of β-2-microglobulin (Sege et al., 1981) and the transport of HLA-RR antigen requires the presence of the gamma chain (Kvist et al., 1982).

Another protein species (160,000D molecular weight) was found to be reactive with anti-G antibody, however, its appearance was independent of irradiation since it was also present in the non-photolyzed sample (Fig. 3.4.16, lane b). This may represent a

G protein aggregate not completely dissociated during electrophoresis whose appearance becomes evident because of the high sensitivity of the immunoblotting technique. A minor protein having this molecular weight was also occasionally observed on very long exposures of gels containing samples of [^{3}H]-palmitate labeled VSV (data not presented).

The photoinduced products exhibiting a molecular weight lower than G protein (compounds V and VI) were not the result of degradation of G protein since they were not reactive with anti-G antibody.

None of the photoinduced species were reactive with antibodies specific to the M or N proteins. Thus, there is no formation of a G-M heterodimer. This suggests that, although the G and M proteins may be in close association with each other, the position and length of the fatty acid is insufficient to span the distance between them. This would be the case if the fatty acids were acylated to the serines located at positions 443 and 444 as indicated in Figure 4.1. In these positions, the probe would only be able to extend approximately half way into the lipid bilayer thus excluding the possibility of forming G-M heterodimers since the M protein does penetrate the bilayer to that extent. A photoreactive fatty acid acylated somewhere in the cytoplasmic extension of G protein would, however, be more likely to form G-M dimers. This would also be true for the formation of G-N heterodimers, however, none were detected. This result further suggests that the fatty

acids are not present in the cytoplasmic extension of G protein.

Thus, although the fatty acids present on G protein may be important for viral assembly and function as has been suggested (Schlesinger and Malfer, 1982), their actions are not mediated by a close association with either M or N protein in the virus.

Minor Components in the Viral Envelope

The lack of reactivity of the photoinduced compounds II (MW 107,000D), II (MW 100,000D), IV (MW 90,000D), V (MW 53,000D), or VI (MW 24,000D) with any of the viral specific antibodies tested suggests that they represent minor integral membrane proteins present in the viral envelope.

Lactoperoxidase catalyzed radioiodination of VSV grown in L cells and Vero cells has shown that VSV incorporates several specific host cell surface proteins into its envelope during budding (Lodish and Porter, 1980c). The two major species incorporated had molecular weights of 110,000D and 20,000D. Thus, they could be analogous to the compounds II and VI that are observed following photolysis of VSV. Lodish and Porter (1980c) showed that the incorporation of these proteins was specific and not due to contamination with cellular debris. They further demonstrated that each virion incorporated approximately 10 copies of the 110,000D species and 80 copies of the 20,000D species as opposed to 500 copies of G protein. This ratio (1:8) is similar to that observed between compounds II and VI following irradiation (Fig. 3.4.11). Mancarella and Lenard (1981) also observed species corresponding to these molecular weights after

labeling VSV with [^{125}I] iodonaphthylazide. Recently, Little et al. (1983) have reported that VSV grown in HeLa cells specifically incorporates several cell surface antigens with molecular weights between 75,000D and 100,000D into the viral envelope.

Based on immunological studies, Hetch and Summers (1972, 1976) have demonstrated the existance of mouse H-2 histocompatability antigens in highly purified preparations of VSV, however, this interpretation has been questioned (Haspel et al., 1977).

Various enzymatic activities have been detected in VSV that appear to be an integral part of the virion but of cellular origin. These include a protein kinase activity (Imblum and Wagner, 1974; Moyer and Summers, 1974; Clinton et al., 1982) as well as an enzymatic activity that synthesizes cytidyl (5'-3') guanosine 5' triphosphate (Chanda and Banerjee, 1981).

Thus, VSV does not completely exclude pre-existing host proteins from the viral envelope during assembly. The specificity of incorporation of these proteins to the exclusion of the many proteins present in the cellular membrane suggests thay they may play some important but as yet undefined function. The use of a radioactive photoreactive fatty acid as a precurosr for viral phospholipid synthesis provides a novel and sensitive means of detecting such minor proteins.

L Cells

The use of this photoreactive fatty acid is, of course, not restricted to VSV. The results using L cells (Section 3.4) was the

first demonstration of the use of such compounds to examine membrane structure in eukaryotic cells. Khorana and co-workers were able to biosynthetically incorporate diazirine containing fatty acids into E. coli and identify membrane proteins after photolysis, however, these fatty acids were not radioactive (Quay et al., 1982). The studies with L cells indicate that eukaryotic cells can recognize these modified fatty acids and incorporate them into both phospholipids and proteins. The fatty acids so incorporated retain the photoreactive group and remain capable of undergoing irradiation dependent crosslinking to both phospholipids and proteins.

Analysis of DAP nonanoate labeled L cell proteins by gel electrophoresis (Fig. 3.4.4) demonstrated that a reproducible subset of proteins were labeled prior to photolysis. The fact that the labeling pattern was different from the pattern observed with [^{35}S] methionine labeled proteins and that the label was sensitive to hydroxylamine (unpublished observations) indicated that the photoreactive fatty acid was acylated directly to these proteins. Thus, L cells contain a subset of proteins that have lipid moieties covalently attached to them. The results obtained with L cells agrees with that reported with chick embryo fibroblasts labeled with [^{3}H]-palmitate (Schlesinger et al., 1980) and suggests that fatty acid acylation of proteins is a general phenomena.

The level of radioactivity present in some of these fatty acid acylated proteins increased dramatically on irradiation indicating that these were integral membrane proteins. Other

proteins were only observable after photolysis, indicating that these integral membrane proteins are not fatty acid acylated. Proteins that were labeled prior to photolysis but did not exhibit an irradiation dependant increase in labeling may represent soluble cytoplasmic proteins which covalently interact with fatty acids.

The photoreactive fatty acid may thus be used to identify integral membrane proteins present in mammalian cells and produce photoaffinity derivatives of those proteins which normally contain covalently attached fatty acids.

4.4 Association of M Protein with Lipid Vesicles

Several reports have demonstrated that M protein is associated with membranes in infected cells (David, 1973; Hunt and Summers, 1976; Atkinson et al., 1976; Morrison and McQuain, 1977) but very little is known about the nature of this interaction or whether M protein has an intrinsic affinity for membranes. In addition, despite the well documented importance of M protein in the life cycle of the virus, its relationship with the viral envelope and other structural components is unclear. This is in large part due to its inaccessable location in the virion and its insolubility in the absence of detergents.

Since our knowledge of the nature of membrane structure and phospholipid-protein interaction has benefited from the use of artificially reconstituted membranes as model systems, this approach was taken to examine the affinity of M protein for membranes.

The M protein used was purified by gel electrophoresis, salt

extraction of virus followed by column chromatography, or synthesized *in vitro*. In each case, M protein was shown to have a strong affinity for lipid vesicles when reconstitution was performed by the detergent dialysis procedure. In addition, detergent free M protein was shown to partition into preformed phospholipid vesicles.

Proteolytic digestion of the reconstituted vesicles demonstrated that the association of M protein was not due to physical entrapment during vesicle formation. In contrast, when N protein was reconstituted with lipid vesicles under identical conditions, no association was observed. The association of M protein with lipid was maintained even after including 0.5 M NaCl in the flotation gradients to eliminate any non-specific electrostatic attraction.

Similar results were independantly reported by Zakowski et al. (1981) who also demonstrated that a specific lipid-protein interaction with M protein and artificial vesicles was maintained in the presence of high salt. Their study was extended to show that while the association of M with lipid was unaffected by the presence of salt after reconstitution, the presence of high salt concentrations before and during detergent dialysis inhibited the reconstitution of M protein with the lipid vesicles. Thus, although high salt concentration can prevent reconstitution, it cannot dissociate M protein once it is bound to the vesicles. These results suggest that the binding of M protein to lipid is a two step process involving an initial electrostatic attraction which is salt sensitive

followed by an irreversible hydrophobic interaction once the M protein binds to the membrane.

The initial electrostatic association was demonstrated by examining the reconstitution of M protein after derivatization with succinic anhydride or acetic anhydride to block lysine groups (Zakorski et al., 1981). Derivatization of M protein, which is highly basic (pI of 9.1) greatly reduced the amount of M protein binding to vesicles. However, it was not clear whether this was the result of partial denaturation of the protein.

The nature of the M-lipid association in the reconstituted in vitro system is not clear. The inability of high salt concentration to remove M protein once bound to the vesicles suggests a lipophilic component is involved, however, the experiments described with the photoreactive probe (Section 3.4) clearly demonstrate that in the virus at least, the M protein is not in contact with the hydrophobic core of the membrane.

Zakowski et al. (1981) reported efficient reconstitution of M protein only with vesicles containing negatively charged phospholipids and not in vesicles composed of only phophatidylcholine. In contrast, the results presented in Section 3.5 show that reconstitution of M protein with phospholipid vesicles containing only PC was highly efficient. Some of the possible reasons for this difference have been outlined in Section 3.5. Other differences may be related to the method of purification of M protein or the conditions of reconstitution. Zakowski et al. (1981) purified

N protein by phosphocellulose chromatography while in the situation reported in this thesis, M protein was purified by gel filtration or synthesized *in vitro*.

The results reported here are similar to that obtained with influenza virus M protein where it was demonstrated that influenza M protein incorporates efficiently into artificial phospholipid vesicles composed entirely of phosphatidylcholine (Gregoriades, 1980; Bucher et al., 1980).

The nature of the interaction of the influenza M with lipid vesicles is, however, much clearer than is the case for VSV M. Although the overall composition of the influenza M protein is not hydrophobic, there are several distinct regions where there is a clustering of hydrophobic and/or neutral amino acids (Winter and Fields, 1980). Two of these regions are situated near the amino terminus while two other clusterings are near the middle of the molecule. Fluorescent labeling with the photoreactive hydrophobic probe pyrenesulfonyl azide was used to demonstrate that one of the middle domains as well as one of the amino terminal proximal hydrophobic domains actually penetrate the lipid bilayer (Gregoriades and Frangione, 1981).

The predicted amino acid sequence of the VSV M protein on the other hand does not show any extended stretches of hydrophobic or neutral amino acids that might be expected to preferentially associate with lipid (Rose and Gallione, 1981). The VSV M protein is in fact highly basic and contains a large number of lysine residues,

many of which are clustered near the amino terminal end of the molecule.

As demonstrated in Section 3.5, digestion of M containing lipid vesicles with proteolytic enzymes resulted in the complete degradation of M protein, however, several small protease resistant fragments having a molecular weight of 3-5000D remained associated with the vesicles following repurification. This result is similar to that reported with reconstituted influenza M protein treated in an identical manner (Gregoriades, 1980). In the case of VSV M protein reported here, it is not clear whether the fragments protected from proteolytic digestion are protected by virtue of being embedded in the membrane or simply because of close association with the surface of the vesicles. The results obtained with the photoreactive fatty acid probe would tend to indicate the latter.

The heterogeneity of these small molecular weight fragments may be a reflection of the heterogeneity of the lipid-protein interaction or that more than one region of the molecule is in contact with the membrane, as appears to be the case for influenza M protein. Tryptic peptides of VSV M protein had no affinity for the phospholipid vesicles suggesting that tryptic peptides on their own have no lipophilic properties. It is possible, however, that the tertiary conformation of M protein, bringing small nonpolar regions of the molecule in close proximity to each other, may be important in the lipid association observed. The predicted amino acid sequence of the VSV M protein (Fig. 3.6.2) indicates that while there are no long

hydrophobic stretches, there are several clusers of nonpolar or hydrophobic amino acids of 5-10 residues in length that could conceivably interact with lipid. In the case of the influenza M protein, a neutral amino acid cluster of only 13 amino acids was shown to be sufficient to penetrate the bilayer (Gregoriades and Frangione, 1981).

The reconstitution of M protein into phospholipid vesicles reported in this study as well as by Zakowski et al. (1981) provides the first direct demonstration that VSV M protein has an intrinsic affinity for lipid in vitro. It suggests, but does not prove, that this same affinity is manifested in vivo in the infected cell as well as in the completed virus. Studies with artificial membrane systems reconstituted with viral glycoproteins show that the glycoproteins interact with these vesicles in a manner indistinguishable from that of the viral or host cell membrane. In addition, in the cases where a functional assay is available (such as fusion activity) it has been shown that the reconstituted proteins still maintain and exhibit that function (Hsu et al., 1979; Huang et al., 1980; Helenius et al., 1977; Petri and Wagner, 1979; Miller et al., 1980; Miller and Lenard, 1980; Huang et al., 1980). Thus, it is likely that this observed in vitro affinity of M protein for lipid is also manifested under normal conditions in vivo.

It has been postulated that in the virus M protein may interact with G protein, possibly via the cytoplasmic extension of G, and that this interaction is important for viral assembly. The basic

nature of the cytoplasmic domain of G protein may facilitate this interaction (Rose et al., 1980). Evidence for this has come from crosslinking studies which have shown that it is possible to form G-M heterodimers (Dubovi and Wagner, 1977; Mudd and Swanson, 1978). The level of these dimers was, however, extremely small. More recently, using the technique of fluorescence recovery after photobleaching, it was demonstrated that M protein is able to reduce the lateral mobility of G protein in the infected call membrane (Reidler et al., 1981). It was suggested that M protein immobilizes a small amount of cell surface G protein molecules into budding sites.

The results presented here, in regard to the topology of G protein in the membrane and the observation that M protein has an affinity for lipids supports the concept of M interacting with G protein. The results of the recontitution experiments clearly show, however, that G protein is not necessary to promote M-lipid association since efficient reconstitution of M protein with lipid vesicles occurred in the absence of G protein. Thus, the postulated association of M protein with the cytoplasmic tail of G protein is not a prerequisite for the interaction of M protein with lipid. This is supported by studies with mutants of VSV which have demonstrated that G protein is not essential for virus budding to occur (Schnitzer and Lodish, 1979).

Association of N Protein With Lipid Vesicles Containing M

The demonstration that M protein can be reconstituted into artificial lipid vesicles in an external orientation provides a

unique system with which one can directly examine the interaction of M protein not only with lipid but also with other viral components.

To test the idea that the association of viral nucleocapsids with the envelope is mediated through an interaction between M protein and the nucleocapsid protein N (Wagner, 1975; Mudd and Swanson, 1978; Schnitzer and Lodish, 1979; Wilson and Lenard, 1981) reconstitution of N protein with phospholipid vesicles in the presence or absence of M protein was examined.

As demonstrated in Section 3.5, N protein on its own has no affinity for lipid vesicles. In the presence of M protein, however, both M and N protein were found to associate with the lipid vesicles. As a control, when an equal mixture of G and N proteins were reconstituted, only the G protein was found to associate with the lipid vesicles. It has previously been demonstrated that G protein can be incorporated into synthetic phosphatidylcholine containing vesicles either by the detergent dialysis method or by using preformed vesicles (Petri and Wagner, 1979, 1980). Thus, the association of N protein with M containing vesicles is not simply the result of a non-specific protein-lipid coaggregation resulting from the presence of a protein that normally has lipid affinity.

It is concluded then that the association of N protein with lipid vesicles, only in the presence of M protein, must be mediated by an interaction occurring between the N protein and the membrane associated M protein. In one experiment where N protein and

detergent purified M protein was used for reconstitution (Fig. 3.5.10) the N:M ratio was found to be 1.1:1 after correcting for the number of methionines present in each protein. Since the molecular weight of N protein (50,000D) is approximately half that of M protein (29,000D), it can be said that one molecule of N protein is associated with two molecules of M protein. This agrees with the estimation that the number of N and M molecules in the virion is approximately 2300 and 4700, respectively (Wagner, 1975).

These observations were confirmed with proteins synthesized in vitro. These results are probably more significant since they depend on proteins whose conformation more closely resembles their respective native state. As with the virion purified proteins, reconstitution of N and M proteins synthesized in vitro resulted in the association of both these proteins with lipid vesicles. This interaction does not seem to be non-specific. In one experiment (Fig. 3.5.12) the starting ratio of N/M in the reaction mixture was 3:1, however, after reconstitution these proteins were present in the lipid vesicles in a ratio of 1:2.3. If non-specific lipid-protein aggregation was occurring, one would expect to find a similar N/M ratio in the vesicles as was present in the starting material.

N protein synthesized in vitro requires the presence of M protein for association with lipid vesicles since N protein was not found in the lipid fraction in the case where the synthesis of M protein was supressed by hybridization. N protein was, however, found to associate with lipid when the reaction was supplemented

with unlabeled purified M protein. Thus, as with virion purified N protein, N protein synthesized in vitro will associate with lipid vesicles only in the presence of M protein.

This data is taken to indicate that M protein contains two functionally distinct regions or domains, one which can recognize and interact with the lipid bilayer, and one that can recognize and interact with the nucleocapsid protein N.

The mechanism by which N protein may associate with M is not known. The highly basic nature of M protein (Rose and Gallione, 1981) may be involved in this process. As stated above, in one experiment in which N and M were reconstituted with lipid vesicles, the N:M ratio was found to be 1:2.3. When these reconstituted vesicles were isolated and re-purified through a flotation gradient in the presence of 0.5 M NaCl, the ratio of N:M was found to be approximately 1:4.5. Thus, while approximately half of the N molecules were removed by salt treatment, the remainder remained tightly associated with the M containing vesicles. This suggests that the interaction of a portion of the N molecules with M protein is not through simple electrostatic association.

Recently, it has been demonstrated that fractionation of VSV on SDS polyacrylamide gels under non-reducing conditions results in the appearance of a new species having an apparent molecular weight of 87,000D in addition to the other viral proteins (Mancarella and Lenard, 1981). Analysis in a second dimension under reducing conditions demonstrated that this new species was composed of N and M

proteins. The appearance of the 87,000D species was dependant upon the pH of the dissociation buffer and was abolished with reducing agents such as 2-mercaptoethanol. These results suggested that the N-M complex arose from a process of disulfide exchange. The amount of this N-M complex was reduced in temperature sensitive mutants defective in M protein (Mancarella and Lenard, 1981).

The results reported here, namely that N and M protein can interact in vitro, does not prove that the same interaction occurs in vivo during viral assembly. N protein is normally found tightly complexed in a RNP structure, thus, it may have a conformation totally different from that exhibited in an RNA free state. There is no way to discriminate between biologically important conformational structures of N protein and other conformational states since there is no functional assay available for N protein. However, the results with purified RNP particles demonstrated that M protein can interact with nucleocapsids. As shown in Figure 3.5.15, when M protein was reconstituted with phospholipid vesicles in the presence of RNP particles, the M protein remained at the bottom of the tube as opposed to its upward flotation in the absence of RNP. This was teken to indicate that M protein was, in fact, associating with the RNP but was prevented from floating upwards because of the large size and density (1.31 gm/ml) of the RNP.

Further indication that M can associate with the RNP was obtained from the observation that a small amount of M protein was consistently observed to associate with in vitro synthesized RNP

particles containing the 42S RNA and N protein, even after centrifugation through 30% CsCl (Ghosh and Ghosh, 1982) or when these RNP particles were centrifuged through a sucrose gradient containing EDTA to dissociate polysomes (Capone and Ghosh, unpublished observations). It appears, therefore, that M protein has a strong affinity for nucleocapsids.

As stated in the Introduction, M protein regulates virus directed RNA synthesis and inhibits VSV transcription in vitro (Clinton et al., 1978; Martinet et al., 1979; Carroll and Wagner, 1979; Combard and Printz-Ane, 1979; De et al., 1982). Recently, Wilson and Lenard (1981) have used M protein mediated inhibition of VSV transcription in vitro to examine interactions between wild type and mutant M protein with nucleocapsids. It was shown that the inhibition demonstrated by M protein was dependent on the ionic environment. Wild type as well as mutant M protein was inhibitory under conditions of low ionic strength, however, under high ionic strength, the wild type but not the mutant M protein still displayed some inhibitory activity. This suggested that the mutant M protein loses its capacity to interact with nucleocapsids under conditions of high ionic strength.

Recent experiments have indicated that M protein plays a key role in maintaining the condensed structure of the RNP (Newcomb and Brown, 1981; Newcome et al., 1982). Disruption of VSV by non-ionic detergent in the presence of 0.5 M NaCl to solubilize G and M proteins led to the disassembly of the RNP into an extended form.

Dialysis to remove salt led to the complexing of M to the RNP and the reassembly of the nucleocapsids into condensed structures. The nucleocapsids were shown to have a high affinity for M protein since virtually all the nucleocapsids were in the form of condensed structures complexed with M protein.

Recently, De et al. (1982) using purified M protein to examine inhibition of transcription in vitro showed that a ratio of M:N of 2:1 (w/w) resulted in a greater than 80% reduction in the level of VSV RNA synthesis. In addition, it was demonstrated that this inhibition of transcription was accompanied by a conformational change in the transcribing nucleocapsid from an extended form to a compact structure. The purified M protein did not interact with these nucleocapsid structures at an ionic strength greater than 0.12 M. Thus, the association of M protein with the N protein contained in the nucleocapsids may be important in regulating transcription.

Thus, the work described by other investigators as well as the work described in this thesis indicates that M protein can interact with both the lipid bilayer as well as the N protein. That these two sites are functionally distinct is supported by recent work with temperature sensitive mutants that have shown that M protein displays co-ordinate affinity for membranes and nucleocapsids (Reidler et al., 1981; Wilson and Lenard, 1981; Mancarella and Lenard, 1981). That is, mutants that displayed a decreased association with nucleocapsids at the same time exhibited an

increased modulation of G protein mobility at the cell surface, suggesting an increased association of M with the membrane. Phosphorylation of M protein may play a role in these interactions, however, the functional role of phosphorylation of M protein, if any, has yet to be established (Clinton et al., 1978).

Knipe and co-workers have examined the role of N protein in virus maturation by using a temperature sensitive mutant of VSV in which N protein is rapidly degraded at the non-permissive temperature (Knipe et al., 1977a). In this case it was found that a large proportion of M protein accumulated in the cytoplasm rather than attaching to membranes. It was thus suggested that M protein could associate with membranes only after binding to nucleocapsids. In contrast, the studies reported here as well as by others (Zakowski et al., 1981) clearly show that M protein can efficiently bind to membranes *in vitro* in the absence of N protein or nucleocapsids. In addition, as stated above, mutants of M protein which display a weakened binding to nucleocapsids at the same time exhibit an increased, rather than a decreased, affinity for membranes. The differences between the *in vitro* reconstitution studies and those reported by Knipe (Knipe et al., 1977a) may reflect differences in interaction with cellular components, such as the cytoskeleton, which may be occurring *in vivo*.

The observation that M protein has affinity for both membranes and N protein has implications for both the assembly and the function of the virus. It supports the conclusion that M

protein, bound to the G protein modified plasma membrane, serves as the nucleation site for the RNP to initiate budding and, as such, would be the rate limiting step in virus assembly (Knipe et al., 1977a; Schnitzer and Lodish, 1979; Weiss and Bennett, 1980). In addition, as recently pointed out by De et al. (1982), if the interaction of M with the RNP occurs at the membrane prior to or at the time of budding, this interaction may cause the RNP to cease RNA synthesis prior to maturation of the virus. Since it appears that M protein has an influence on the transcriptive properties of the RNP and that N protein initiates encapsidation (Blumberg et al., 1983), the association of M protein with N protein in vivo would cause transcription to halt as well as initiate the condensation of the nucleocapsid. These interactions as well as other conformational changes that may be occurring would then initiate the budding event.

4.5 Construction of a Chimeric Gene Between G and M

Section 3.6 describes the construction and cloning of a hybrid gene between the VSV G and M proteins as an approach to examine the functional significance of various domains of these two molecules.

The cloning strategy was developed to insure that the correct reading frame at the G-M junction point was maintained. Based on size and restriction enzyme analysis, several positive clones were isolated. One of these clones, designated pGM03, was used for further study.

When this clone was introduced into mammalian cells by DNA

mediated gene transfer, 1-2 polypeptides were detected which were precipitable by specific antibodies directed against authentic VSV M protein. These protein species were not detected in nontransfected cells nor in cells transfected with the vector alone. Thus, based on their apparent molecular weights and specific reaction with anti-M antibody, it was concluded that these polypeptides were transcribed and translated directly from the G-M hybrid gene.

In the absence of any definitive protein characterization, the nature of these protein species can only be speculated upon. The 35-36 kilodalton species (Fig. 3.6.12 and 3.6.13) has an apparent molecular weight indicative of a glycosylated G-M hybrid protein. However, due to the small molecular weight differences between an uncleaved signal sequence of 16 amino acids and a typical Asn linked oligosaccharide chain, it is difficult to determine the precise nature of the species based on size alone. In addition, it is not known what effect a carbohydrate chain would have on the apparent molecular weight of the product in this case. If the product is glycosylated, then it must have been translocated across the endoplasmic reticulum by virtue of the signal sequence of G protein. This would indicate that the signal sequence of G is sufficient to transport a normally cytoplasmic protein across membranes and that the sequence Asn-X-Ser(Thr) on its own is sufficient to trigger glycosylation. Because of the small molecular weight differences involved, amino acid sequence analysis is required to determine whether signal sequence cleavage occurs.

In the normal situation, G protein is inserted into membranes shortly after the signal peptide exits from the large ribosommal subunit (Rothman and Lodish, 1977; Toneguzzo and Ghosh, 1978), however, it is not clear at which point the signal sequence is removed. It is felt that the signal sequence is removed during or shortly after it is inserted into the endoplasmic reticulum but is is not known whether other sequences downstream from the signal sequence are required for efficient cleavage. In the case of the G-M hybrid, the hybrid gene codes for the 82 NH_2-terminal amino acids of mature G protein (see Fig. 3.6.1) so if sequences near the amino terminus of G are required for efficient signal sequence cleavage, it is likely that the chimeric gene contains them. In addition, the cleavage of the signal sequence may not be necessary for efficient membrane insertion. For instance, ovalbumin appears to have a permanent signal sequence which is not proteolitically processed (Lingappa et al., 1978b).

The G specific sequences contained in the hybrid gene would, in fact, be of the ideal length to initiate membrane insertion and signal sequence cleavage. The G specific sequence includes codons specific for the 16 amino acids of the signal sequence and the 82 amino acids downstream from this point (total = 98 amino acids). It has been estimated that the nascent polypeptide chain of a secretory or membrane protein woud have to attain a minimal length of 70-90 amino acids for the polypeptide to penetrate the membrane and have the signal sequence removed (Rothman and Lodish, 1977; Sabatini et

al., 1982). This would include 30-40 amino acids remaining in the ribosome plus 20 residues required to cross the membrane and the 16-30 residues of the signal element.

As stated above, the size of the pGM03 specific product would indicate that it is glycosylated. This question can be addressed more accurately with the use of glycosylation inhibitors such as tunicamycin which inhibit formation of the dolichol intermediate.

The 32,000D species, immunoreactive with anti-M antibody, was also specifically detected only in cells which were transfected with pGM03 (see Fig. 3.6.13). This species is consistant with a G-M fusion protein which is not glycosylated and has had the signal sequence removed. It is, however, difficult to discriminate between polypeptides containing a signal sequence and those lacking it without further characterization.

A high molecular weight species was also detected with anti-M antibody in non-transfected cells as well as in cells transfected with pCVSVe or pGM03. This product was, however, much more prominent in the pGM03 transfected cells (see Fig. 3.6.13). This may be an artifact due to nonspecific binding or it may be a product induced specifically by transfection with pGM03. Other species cross reactive with anti-M antibody were also detected in transfected as well as non-transfected cells by the immunoblotting procedure. These species probably represent nonspecific binding detectable because of the very high sensitivity of the immunoblotting technique. Alternatively, they may represent normal cellular proteins which

contain antigenic sites related to M protein.

These important questions are difficult to address in the in vivo system because of the low level of expression of the presumed G-M product. Some of the reasons for low expression have been outlined in Section 3.6. Mellon et al. (1981) have demonstrated that while a copy of the globin gene was transcribed to very high levels in COS-1 cells, no translation product was identified. It is only recently that suitable expression vectors have become available which are able to express cloned genes to levels sufficient for detection by direct immunoprecipitation.

It is also possible that the product of the G-M gene is unstable or is rapidly degraded by intracellular enzymes. In addition, it is also conceivable that the G-M product is toxic to the cells, for instance by blocking receptors necessary for protein translocation. Using a bacterial expression vector containing cDNA corresponding to G mRNA to examine the synthesis of G protein in E. coli, it was demonstrated that the hydrophobic signal sequence of G is lethal to E. coli (Rose and Shafferman, 1981).

Another important factor in the level of expression is the efficiency of the transfection. Rose and Berghan (1982), using a transfection protocol similar to the one adopted here, estimated that only 5% of the cells become transfected.

These factors could all be affecting the detection of the G-M product. Alternatively, sufficient amount of the product may be synthesized, but the affinity of the product for the anti-M antibody

may be drastically reduced because of the nature or conformation of the hybrid molecule.

In attaching the signal sequence of G onto M protein, it was felt that some of the protein product may be secreted by the cell. No M specific products were, however, detected in the extracellular medium from pGM03 transfected cells when assayed by immunoblotting (see Fig. 3.6.13). It is possible that if there was a product secreted, it was only present in trace amounts or that it was not immunoreactive with the antibody used. It is, however, not surprising that a product of the G-M gene is not secreted given the demonstrated high affinity for membranes exhibited by M protein. Thus, although a G-M hybrid protein may be translocated across the endoplasmic reticulum, it may become associated with the intracellular membranes and, therefore, not be able to be secreted. Although it is not known whether a G-M protein would demonstrate a similar membrane affinity as authentic M protein, this should not affect the translocation of this species across the endoplasmic reticulum since this process occurs co-translationally.

Alternatively, the hybrid protein may not contain the signals necessary to specify secretion. Expression of G protein (Rose and Bergman, 1982) and the influenza haemagglutinin (Gethig and Sambrook, 1982) which do not contain the hydrophobic membrane interacting region, show that the cell treats these truncated species as normal secretory proteins. This indicates that secretory proteins, as a class, do not contain additional amino acid sequences that specify

polypeptide secretion. However, both the haemagglutinin and G protein are normally transported through the intracellular membrane system and, thus, even the truncated forms may contain signals which specify this function. The biogenesis of M protein, on the other hand, does not involve either membrane translocation or intracellular membrane transport and, thus, it is expected that the M protein would lack these locational or transport signals necessary for the subsequent fate of normal secretory proteins.

The important questions regarding the synthesis and processing of this G-M hybrid protein can be properly addressed using an in vitro transcription-translation system. Recently, a whole cell extract prepared from Hela cells and containing endogenous RNA polymerase II has been developed which is able to faithfully transcribe RNA from a DNA template (Manley et al., 1980). This has been used to develop a sequential transcription-translation system using the whole cell extract for RNA synthesis and the reticulocyte translation system for protein synthesis (Cepko et al, 1981). By using these systems and carrying out the reactions in absence or presence of exogenously added microsomal membranes, it should be possible to easily examine the synthesis, membrane insertion, signal sequence cleavage, and glycosylation of the G-M hybrid gene product.

4.6 Future Prospects

Although advances in the past few years have been substantial, there are still many aspects of membrane assembly and function which need to be addressed and with which simple enveloped animal viruses

will undoubtedly play an essential role.

In the case of G protein, covalent addition of fatty acid residues is a modification whose function remains to be convincingly elucidated. Perhaps viral mutants defective in only fatty acid acylation may be useful in this regard. However, a potentially more useful approach would be the isolation of cells which are defective in the acylating activity. A mutant Chinese Hamster Ovary cell line defective in the en bloc transfer of oligosaccharide chains has been very useful in the elucidation of the mechanism of glycosylation as well as determining the biological role of covalently attached carbohydrate (Krag et al., 1977; Gibson et al., 1980).

Since many different types of membrane proteins have been recently demonstrated to contain fatty acids, it is likely that the fatty acid residues provide some important function, at least for those proteins so modified. The precise linkage site of the fatty acids would thus be important not only from the point of eludicating function but also in the sense of understanding the nature of the target sequences necessary for acylation to take place. Perhaps the advancing technologies in purifying hydrophobic peptides in sufficient amounts coupled with sensitive characterization techniques may help define the precise linkage residue. Recently, gas chromatography-mass spectroscopy was used to demonstrate that the amino terminal glycine is the acylated amino acid in the p15 protein of murine leukemia virus (Henderson et al., 1983).

Since many proteins are fatty acid acylated, the enzyme(s)

responsible for this modification must be common to many cell types. It is not known whether the fatty acid is donated directly, for instance, through an activated ester such as palmityl-CoA as postulated (Magee and Schlesinger, 1982) or whether phospholipid is the donor as appears to be the case for the E. coli murein lipoprotein (Lai et al., 1980). An in vitro system which is capable of faithfully acylating membrane proteins would be invaluable in this regard. Rothman has shown that the oligosaccharide moiety of G protein could be processed in vitro after the addition of purified Golgi membranes. Furthermore, the Golgi membranes could be subfractionated into components which contain the acylating activity (Dunphy et al., 1981). Thus, by using this system it should be possible to purify the acylating components.

The use of photoreactive fatty acids incorporated into membrane phospholipids can be used not only to identify lipid interacting domains of membrane proteins but also to precisely determine the extramembrane and intramembrane boundaries. This is now known with certainty for any membrane protein. Recently, the transmembrane domain of glycophorin A was accurately defined by incorporating glycophorin into artificial lipid vesicles composed of phospholipids containing photosensitive carbene precursors and characterizing the resulting phospholipid-protein crosslinks by protein sequencing (Ross et al., 1982). Furthermore, by attaching photosensitive fatty acids of different chains lengths to G protein, it should be possible to further define the nature of protein-

protein interaction occurring in the viral membrane.

Currently, there is a great deal of interest in understanding the mechanism involved in the transport of membrane proteins from their sites of synthesis to their final destinations. There is controversy in regards to whether or not specific cellular proteins, in close association with the targeted protein, are involved in this process. Perhaps by photolyzing G protein, containing a photoreactive fatty acid, while it is in transit through the intracellular membranes, it may be possible to crosslink cellular proteins which are in close association with G protein. One may then be able to identify the cellular protein(s) involved by immunoprecipitating the G protein containing the crosslinked species.

With the demonstration that N protein exhibits an affinity for M protein in vitro it should be possible, for instance by using bifunctional crosslinking reagents, to examine the regions of the respective molecules necessary for this interaction. This would have important implications in regards to the recognition sequences involved in virus structure and morphogenesis.

Without doubt, the use of recombinant DNA techniques offer the most powerful means of addressing some of the above questions as well as many more. The extremely rapid advances made in recent years in gene cloning, site directed mutagenesis, and gene expression has provided important new avenues of examining interesting biological questions.

The approach described in Section 3.6 can be extended to examine the function of the membrane interacting domain of G protein. In addition, the fatty acid acylation target site can be defined by mutating specific amino acids by primer directed mutagenesis. Manipulation of the G gene should help elucidate the domains of the molecule necessary for transport and also help map the putative sorting sequences. In addition, the availability of cloned genes corresponding to the N and M proteins and their manipulation will, no doubt, lead to a clearer understanding of the nature of the processes involved in virus assembly.

APPENDIX: PROTEIN SEQUENCING AND PTH AMINO ACID IDENTIFICATION.

The automation of the Edman chemistry for protein and peptide sequencing (Edman and Begg, 1967) has had an enormous impact on the study of protein structure by making the aquisition of primary structural information considerably more rapid.. However, it has only been through the technological advancements of the past few years that it has become possible to obtain exended sequence information with proteins of biological interest which are only available in trace amounts. These advances have included subtle instrument changes (Bhown et al., 1980) to major instrument redesign (Hunkapillar and Hood, 1978; Wittman-Liebold, 1980), changes in the chemistry of degradation (Niall, 1973; Brauer et al., 1975) and advances in sample isolation and residue detection (Yuan et al., 1982; Hawke et al., 1982). Together, these advances have made it possible to sequence proteins available in the subpicomole to picomole range where previously nanomole to micromole amounts were required (Walsh et al., 1981).

Automated sequence determination is based on the phenylisothiocyanate method developed by Edman in the 1950's (Edman, 1970). The first reaction involves the coupling of phenylisothiocyanate to the free amino terminal αamino group to form a phenylthiocarbamyl derivative. The derivatized amino terminal amino acid is then cleaved from the peptide chain under anhydrous acidic conditions in the form of a 2 anilino-5-thiazolinone (ATZ) derivative. The ATZ derivative is then extracted, converted to the stable phenylthiohydantoin isomer, and

identified. The shortened peptide or protein is then subjected to further cycles of degradation.

The sequencing work reported in the section as well as in Section 3.2 was performed on a Beckman 890C sequencer equipped with a cold trap modification. The degradations were carried out using the 0.1 M Quadrol program employing a combined benzene/ethyl acetate wash (Brauer et al., 1975). This program reduces protein wash out and Quadrol retention that is a problem with the conventional 1 M Quadrol program.

A prerequisite for obtaining sequence information is the accurate identification of the released PTH amino acids. The most common methods of identification include thin layer chromatography, gas chromatography, amino acid back hydrolysis, and HPLC (Croft, 1980). In the past few years, the method of choice for PTH analysis has become HPLC. This method offers a rapid, nondestructive, reproducible, quantitative, and extremely sensitive means of detecting all PTH derivatives in a single analysis. Sensitivity is in the subpicomole range and is at times only limited by UV adsorbing impurities arising from the Edman degradation. A variety of reverse phase HPLC systems dedicated to PTH amino acid analysis have been described which utilize a number of reverse phase supports combined with isocratic or gradient elution (Zimmerman et al., 1977; Margolies and Brauer, 1978; Johnson et al., 1979; Henderson et al., 1980; Somack, 1980; Tarr, 1981, Black and Coon, 1982; Hawke et al., 1982).

The HPLC system used for the studies reported here is described in Methods. A typical separation of a mixture of PTH amino

acids is shown in Figure A-1. As demonstrated, good resolution is obtained with all PTH derivatives with the exception of the methionine/valine, and the isoleucine/lysine pairs. These derivatives are very difficult to separate (Somack, 1980) and only very recently have systems been developed to adequately resolve them (Hawke et al., 1982). The two variables in the gradient system used were pH and acetate concentration. Increasing the pH while leaving the acetate concentration constant decreased retention times of the basic and acidic amino acids while decreasing acetate concentration at constant pH decreased the retention times of only the basic amino acids. Thus, conditions could be easily standardized to compensate for experimental or column variation. It was found that the resolution of the isoleucine/lysine peaks decreased with column use to the extent that they co-eluted. In these cases, conditions could not be found to separate them.

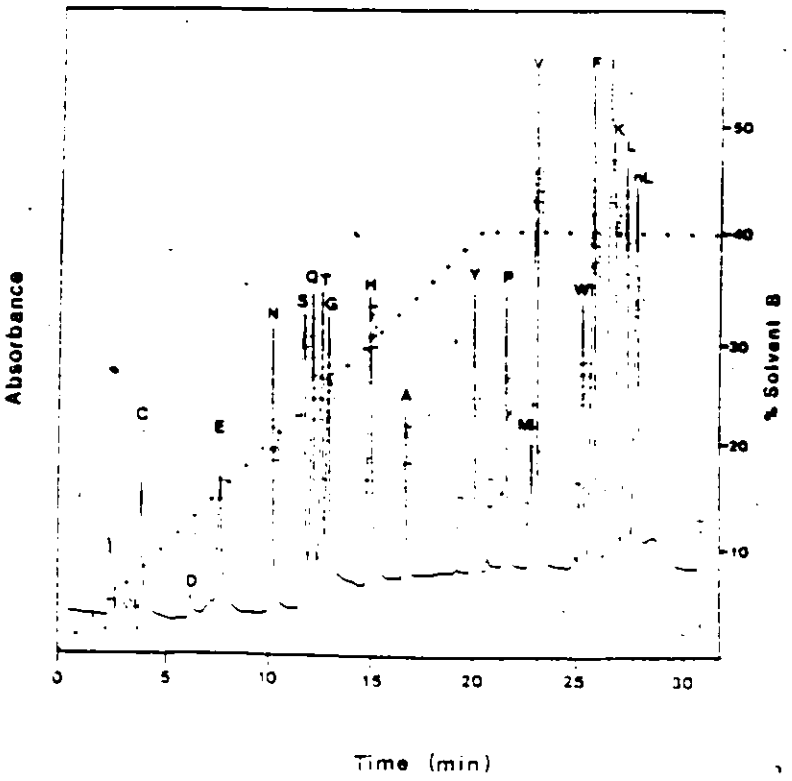
In order to examine the efficiency and accuracy of the Beckman sequencer, proteins of known sequence were degraded and selective cycles were examined by HPLC. A sham cycle (complete cycle in the absence of PITC) was performed for each protein sequenced to precondition the film and increase the initial yield.

Figure A-2 shows the amino acids identified from various cycles of 140 nmoles (2.4 mg) of sperm whale myoglobin. The PTH derivative of norleucine was included in each sample to serve as an internal standard for quantitative purposes and to correct for injection variability. In this particular run, the sequence was determined for

Figure A-1: Separation of PTH amino acids by HPLC.

The top figure shows a typical separation of a mixture of PTH amino acids, containing 1-2 nanomoles of each derivative, by HPLC as described in the Methods section. The gradient profile is indicated by the dotted line. Detection was at 254 nm using 0.16 AUFS.

The table at the bottom was compiled from results obtained from ten separate injections at different times over the life span of the column. The results for PTH-arginine are not included since detection of this derivative was variable.



PTH amino acid	mean t_r	variance	PTH amino acid	mean t_r	variance
Cys	3.94	0.04	Tyr	20.10	0.13
Asp	5.59	0.77	Pro	21.34	0.15
Glu	7.86	1.28	Met	23.05	0.25
Asn	10.77	0.23	Val	23.17	0.18
Ser	12.03	0.29	Trp	25.73	0.30
Gln	12.60	0.16	Phe	26.02	0.33
Thr	12.98	0.17	Ile	26.87	0.34
Gly	13.33	0.19	Lys	26.89	0.33
His	15.39	0.59	Leu	27.70	0.43
Ala	16.99	0.12	nLeu	28.18	0.76

30 residues from the amino terminus (Fig. A-2) and was in complete agreement with the published sequence (Edmonson, 1965). The peak eluting at 24 minutes in cycle 14 represents valine carry over from the previous cycle. This is a common occurence in later cycles and is due to the incomplete extraction of the ATZ amino acid with butyl chloride and its subsequent appearance in the following cycle.

The extent of useful sequence information that one can obtain is dependent on the repetitive yield which in turn determines the absolute yield of a residue at a given cycle. For instance, 100 cycles can be identified with a repetitive yield of 99% while only 10 cycles can be determined with a repetitive yield of 90% (Croft, 1980). On average, most investigators obtain a repetitive yield of 93-95% allowing for the identification of 30-40 residues with an adequate amount of material. The other limiting factor is the rise in background noise due to gradual peptide cleavage under the acidic conditions of the reaction.

The repetitive yield calculations for the myoglobin run shown in Figure A-2 are presented in Figure A-3. The repetitive yield was found as follows; Val (93%), Leu (95%), and Trp (85%). The yield for Trp is generally low because of acid degradation of the indole ring. Thus, tryptophan is progressively lost as the sequencing progresses. (Niall, 1973). The repetitive yields were improved by a modification of the cold trap whereby the drain plug was replaced with a permanent seal. This produced a better high vacuum which resulted in reduced extractive losses, less carry over, and repetitive yields

Figure A-2: HPLC analysis of PTH amino acids obtained from the sequential degradation of sperm whale myoglobin.

The results shown are from the sequence analysis of 140 nmole of protein. Only selected cycles are shown which are indicated by the number at the top left hand corner of each panel. The amino acid assignment is indicated by the one letter code over the peak. The peak eluting at approximately 29 minutes in each panel is the PTH-norleucine internal standard. The top left hand panel is the elution profile of the standard PTH amino acids from which the assignments were made. Typically 5-15% of the PTH derivative from each cycle was analyzed. Absorbance was at 254 nm using 0.32 AUFS.

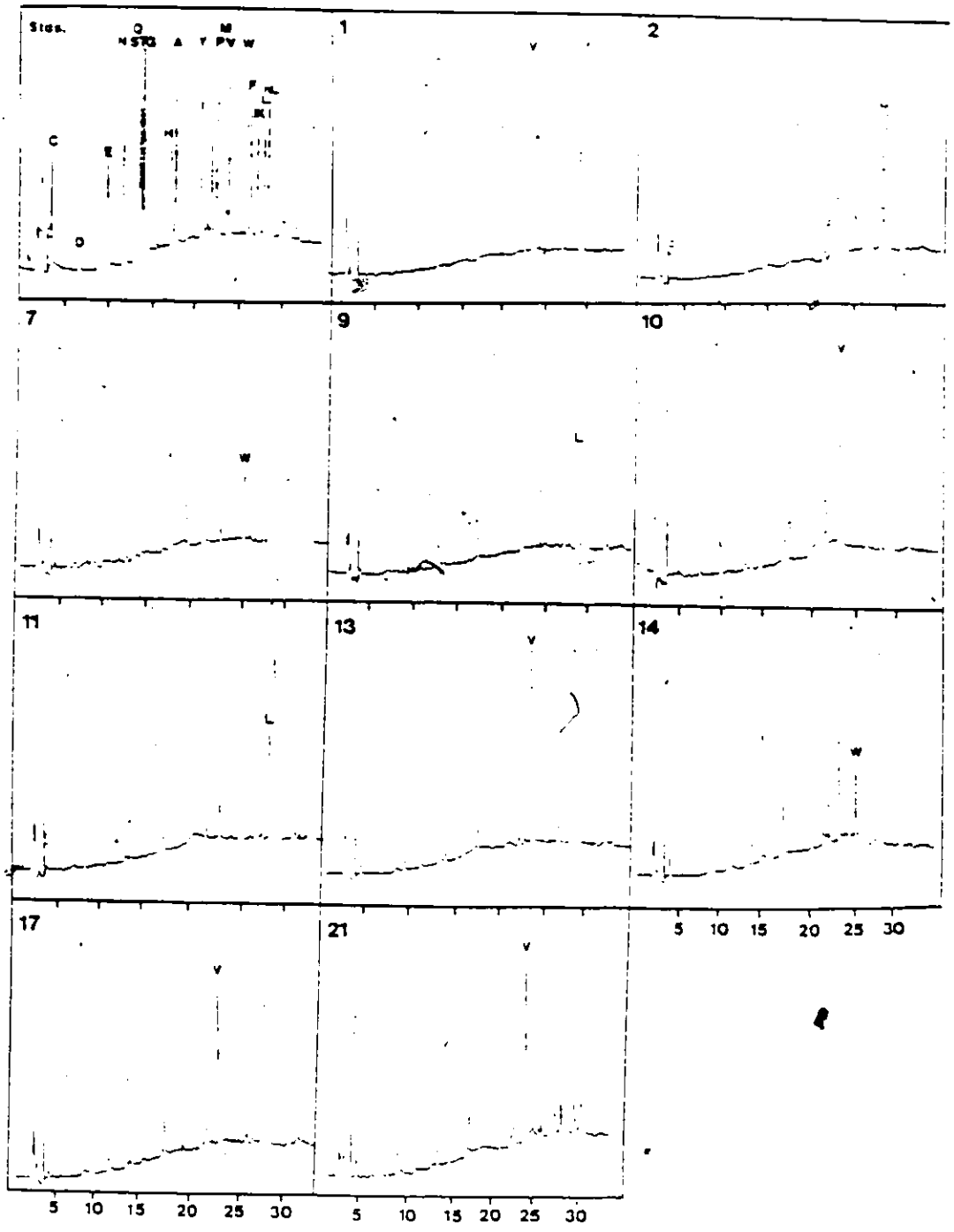
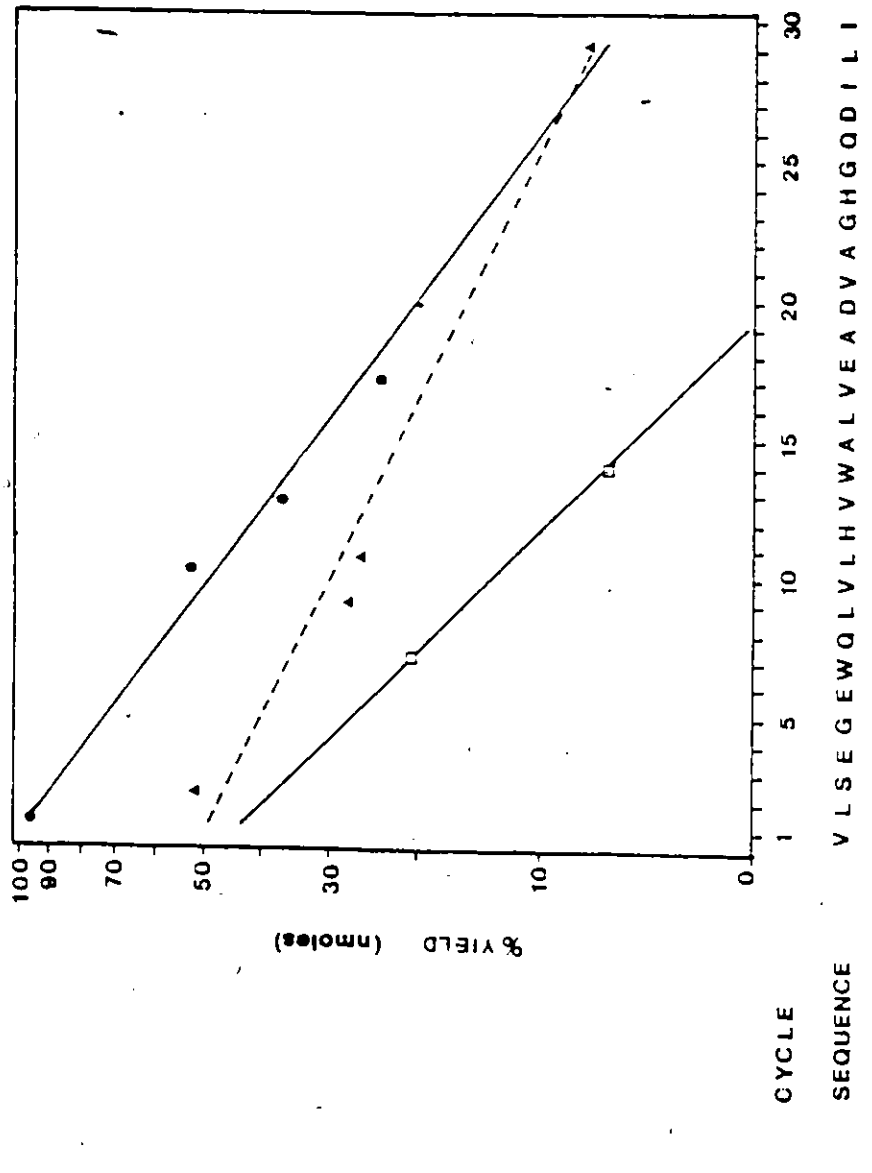

Stas.
1
2
7
9
10
11
13
14
17
21
5 10 15 20 25 30
5 10 15 20 25 30
Time (min)

Figure A-3: Repetitive yield calculation on myoglobin sequence.

This figure shows the yield of various PTH amino acids from the sequence of myoglobin as deduced from Fig. A-2 and Table A-1 as a function of percentage recovered versus cycle number. The repetitive yields were calculated from the slope of the line or by using the following formula;

$$R.Y. = \left(\frac{A}{B}\right)^{\frac{1}{n-1}}$$

where R.Y. is the repetitive yield, A is the absolute yield of residue A, B is the absolute yield of residue B, and n is the cycle number of residue B. The repetitive yields for Val, Leu, and Trp are 93, 95, and 85%, respectively. The sequence of sperm whale myoglobin shown at the bottom of the figure is from Edmonson (1965). (●) Val; (○) Leu; (■) Trp.


% YIELD (nmoles)
100
90
70
50
30
10
0
1
5
10
15
20
25
30
CYCLE
SEQUENCE
V L S E G E W Q L V L H V W A L V E A D V A G H G Q D I L I

routinely obtained in the range of 95-98%.

Figure A-4 represents the degradation of 120 nmole (1.7 mg) of hen egg lysozyme. The sequence obtained matches the published sequence unambiguously (Canfield, 1963, 1965). The repetitive yields in this case were Lys (93%), Leu (97%), and Gly (97%). The absolute yields for this run and the myoglobin run are shown in Table A-1.

The PTH conversions in this case were done with 1 N HCl followed by extraction with ethyl acetate. The PTH derivatives of arginine, histidine, and cysteic acid are not extracted into the organic phase using this conversion procedure since the side chains of the amino acids become ionized. In this case the aqueous phase is examined. For example, in cycle 5 of lysozyme shown in Fig. A-4, analysis of the aqueous phase following conversion showed the presence of arginine. There was also a substantial amount of glycine carry over from cycle 4. Since the glycine was found in the aqueous phase, this indicated that extraction with ethyl acetate was not complete.

There are various methods currently in use to convert ATZ amino acids to the corresponding PTH derivatives. These include aqueous HCl for 10 minutes at 80°C; 25% TFA for 10 minutes at 80°C; and 1 N methanolic HCl for 10 minutes at 55°C (Horn and Bonner, 1977). All three of these methods were examined for conversion efficiency and absolute yield. Thus, 350 nmole (5 mg) of lysozyme was sequenced and the residue from each cycle was dried down, suspended in a small amount of methanol, and split into three equal portions. The aliquots were dried and converted by one of the above procedures (methanolic HCl

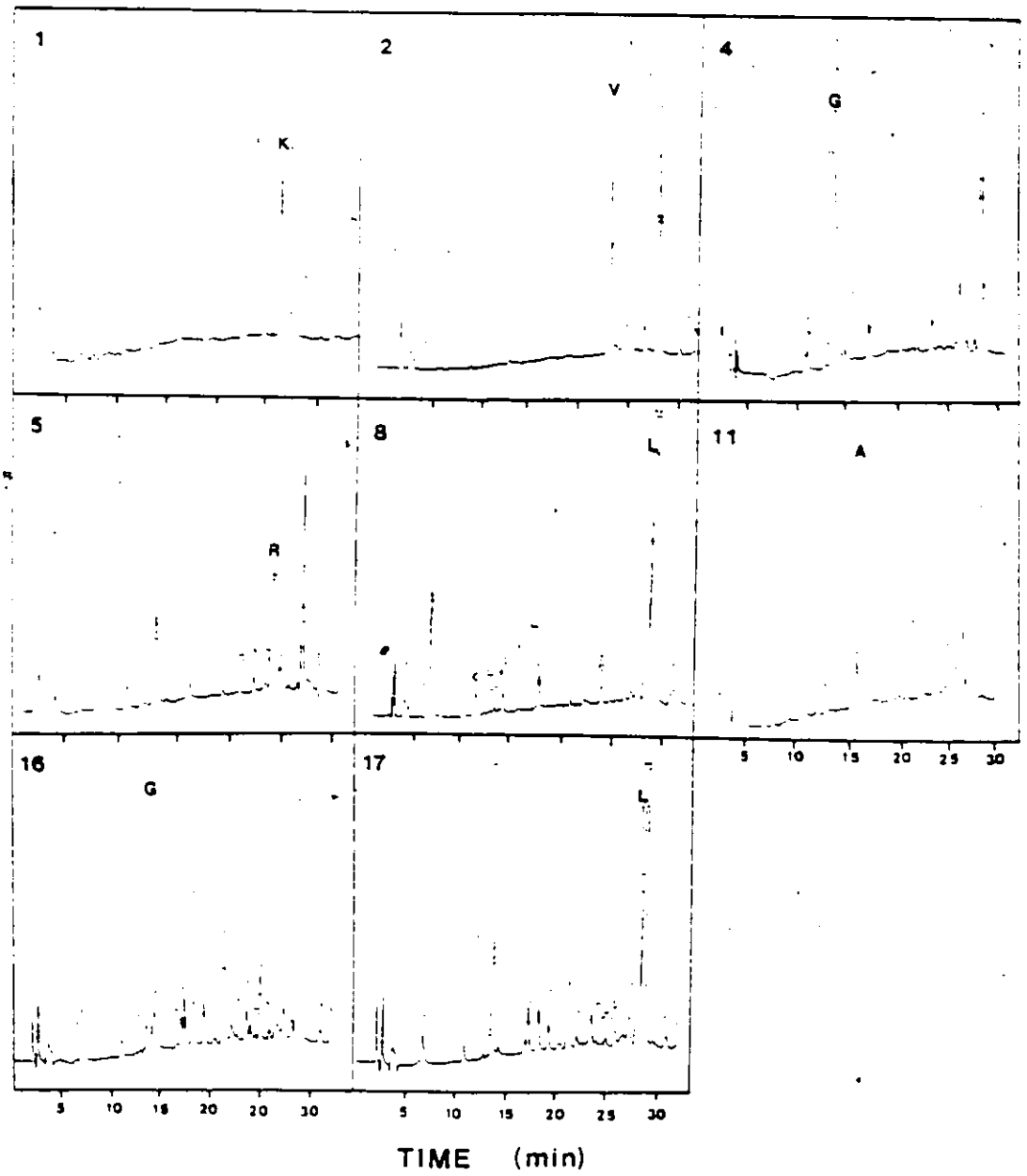


Figure A- 4: HPLC analysis of PTH amino acids obtained from the sequential degradation of hen egg white lysozyme.

The results shown are from the sequence of 120 nmole of protein. The cycles analyzed are indicated by the number at the top left hand corner of each panel and amino acid assignments are indicated by the one letter code. PTH-norleucine was used as the internal standard in each case. Absorbance was at 254 nm using 0.16 AUFS.

Table A-1: PTH amino acid analysis from the sequential degradation of myoglobin and lysozyme.

Myoglobin			Lysozyme		
Cycle No.	PTH amino acid identified	Yield (nmole)	Cycle No.	PTH amino acid identified	Yield (nmole)
1	Val	130	1	Lys	60
2	Leu	64	2	Val	59
7	Trp	28	4	Gly	37
9	Leu	36	8	Leu	29
10	Val	70	13	Lys	24
11	Leu	35	16	Gly	26
13	Val	50	17	Leu	23
14	Trp	8.4			
17	Val	33			
21	Val	22			
29	Leu	10.4			

The yield of PTH amino acids is from the data presented in Figures A-2 and A-4. The calculations were based on integration of peak size from the standard PTH amino acids and from the internal PTH-norleucine standard.

was prepared fresh by making a 1 N solution of acetyl chloride in methanol (Horn and Bonner, 1977)). PTH amino acids arising from the aq. HCl conversion were recovered by extraction with ethyl acetate, while the derivatives from the TFA and methanolic HCl conversion were recovered directly by drying under a sream of nitrogen and suspending the residue in methanol.

Figure A-5 shows the HPLC analysis of the lysine residues recovered from cycles 1 and 13. The method of conversion had no effect on the repetitive yield,it being 93-94% in each case. The absolute yield in each case was, however, higher in the case of TFA conversion. Recovery in cycle 1 was 52%, 39%, and 37.5% of the starting material for TFA, methanolic HCl, and aq. HCl, respectively. In cycle 13, the recovery was found to be 25%, 22%, and 19%, respectively. In addition, the background in cycle 13 was slightly higher in the case of methanolic HCl conversion. Thus, because of higher recoveries, cleaner background, and convenience, TFA was routinely used for the conversions.

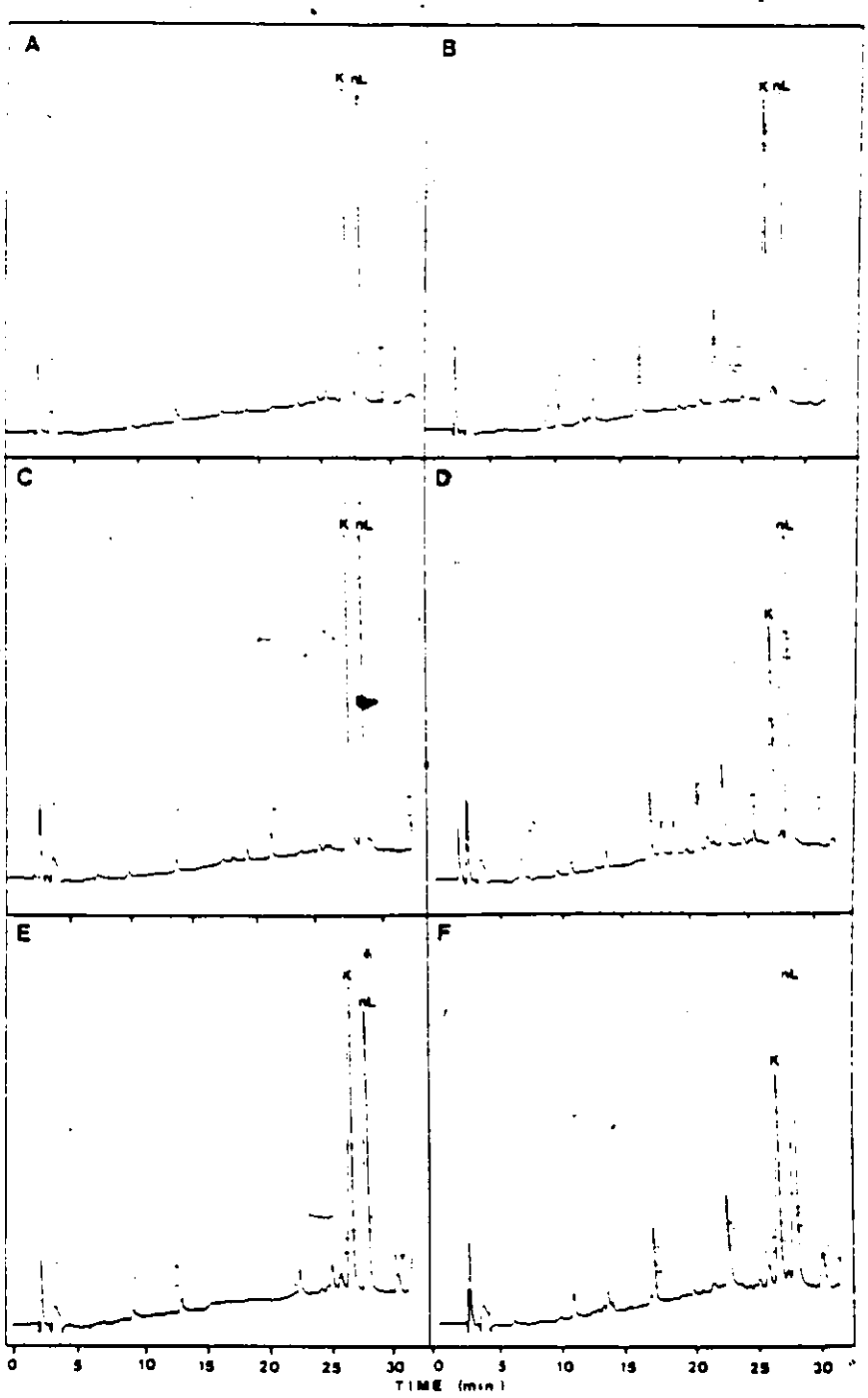
Amino terminal amino acid sequence of G protein from the Cocal and New Jersey serotypes of VSV.

Having standardized sequencing conditions and a method of PTH amino acid analysis, it was desired to examine proteins of unknown sequence.

An ongoing project in our laboratory involves a comparative structural study of various proteins present in different serotypes of

Figure A-5: Comparison of PTH-lysine recovery using different methods of conversion.

Shown are the HPLC scans of cycle 1 (A, C, and E) and cycle 13 lysine (B, D, and F) from the sequential degradation of hen egg lysozyme. A and B, TFA conversion; C and D, methanolic HCl conversion; E and F, aqueous HCl conversion.


A
B
C
D
E
F
K
mL
0
5
10
15
20
25
30
TIME (min)

VSV. The objective is to examine the relatedness of the viruses and in particular to study structure-function relationships of the G protein. The Cocal VSV G protein has an apparant molecular weight of 74,000D while the New Jersey (Concan) G protein has a molecular weight of 64,000D. This is in comparision to the Indiana G protein which has a size of 69,000D. The exent of glycosylation in each case seems to be similar, thus, the differences in molecular weight must be due to differences in the polypeptide backbone.

The G protein from both Cocal and New Jersey was isolated from nonradioactive virus as described in Methods. In the case of Cocal G, 38 nmole (approximately 2.6 mg) was obtained and sequenced. The results from some of the HPLC data is shown in Figure A-6. From this analysis, residues number 1, 2, 4, 5, 6, 7, 11, 14, 15, 16, 17, 18, 21, 22, and 23 were identified as Lys, Phe, Ile, Val, Phe, Pro, Lys, Trp, Lys, Asn, Val, Pro, Tyr, Tyr, and Tyr, respectively. PTH amino acids over background could not be detected in cycles 3, 8, 9, 10, 12, 13, 19, and 20. The results were confirmed by amino acid analysis of a portion of each cycle after back hydrolysis (Smithies et al., 1971). The back hydrolysis and amino acid analysis was carried out by Dr. T. Hofmann at the University of Toronto. The back hydrolysis confirmed all the amino acid assignments made by HPLC and in addition identified those residues not detected by HPLC. These included serine in positions 3, 9, 19, and 20, Gln in positions 8 and 9, Gly in position 12, and Asn in position 13 (Table A-2). The amino acid analysis by back hydrolysis also indicated why certain residues were not detected by HPLC. PTH-serine

Figure A-6: HPLC analysis of PTH amino acids obtained from the sequential degradation of the Cocal VSV G protein.

Results are from the sequence of 38 nmoles of G protein purified from Cocal VSV by detergent extraction as described in Methods. Only selected cycles are shown as indicated by the number at the top left hand corner of each panel. The letter in each panel indicates the assigned amino acid. The HPLC tracings were recorded on a Beckman model BD40 chart recorder. Absorbance was at 254 nm using 0.16 AUFS.

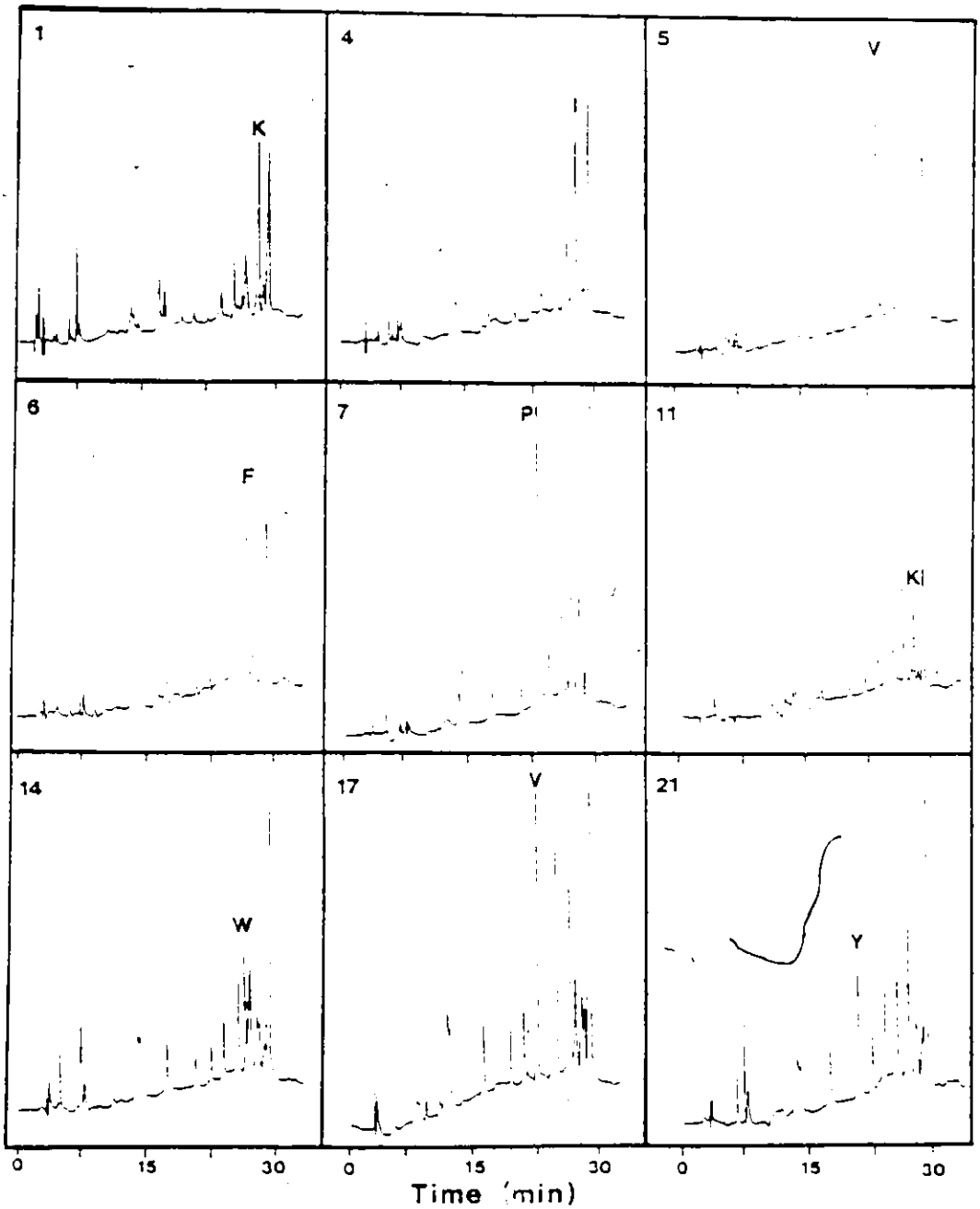

1
K
4
I
5
V
6
F
7
P
11
Kl
14
W
17
V
21
Y
0
15
30
Time (min)

Table A-2: HPLC analysis of the PTH-derivatives obtained by the sequential degradation of 38 nmoles of Cocal G protein. The amino acids indicated in parenthesis were identified by amino acid analysis. The repetitive yield was 95%.

Cycle No.	Amino Acid Identified *	Yield (nmole)
1	Lys (Lys)	20
2	Phe (Phe)	28
3	N.D. (Ser)	-
4	Ile (Ile)	16
5	Val (Val)	28
6	Phe (Phe)	20
7	Pro (Pro)	17.6
8	N.D. (Gln)	-
9	N.D. (Ser)	-
10	N.D. (Gln)	-
11	Lys (Lys)	12.2
12	N.D. (Gly)	-
13	N.D. (Asn)	-
14	Trp (Trp)	4.1
15	Lys (Lys)	9.1
16	N.D. (Asn)	-
17	Val (Val)	10
18	Pro (Pro)	10.1
19	N.D. (Ser)	-
20	N.D. (Ser)	-
21	Tyr (Tyr)	8
22	Tyr (Tyr)	6.9
23	Tyr (Tyr)	5.0

* N.D. NOT DETERMINED

is unstable since the OH on the β carbon has a tendency to undergo β elimination with the dehydro-derivative as the main degradative product (Edman, 1970). In addition, the absolute yield of serine may be reduced due to an O→N acyl migration which can occur during prolonged exposure to the cleavage medium. The PTH derivatives of Asn and Gln are also unstable since the amide group can be hydrolyzed.

The Trp residue in cycle 14 was assigned on the basis of its yield relative to the preceding and the following cycle. The absolute yield was low (4.1 nmole) as expected for a Trp residue, but a comparison of cycles 13, 14, and 15, indicated a legitimate rise for Trp in cycle 14. The consecutive Tyr residues in cycles 21, 22, and 23, were assigned on the basis of a similar criteria, however, one must look at these assignments with caution since in a string of identical amino acids, the latter ones may be due to carry over from previous cycles.

The amino acid assignments in cycles 1, 2, 4, 5, 6, 7, and 11 were also subsequently confirmed by microsequence analysis of Cocal G protein radiolabeled with those respective amino acids (Kotwal et al., in press).

A partial amino acid sequence analysis of New Jersey (Concan) G protein was also undertaken. The purification of New Jersey G protein was carried out as described for Cocal G protein, and the purity of the final product, as judged by polyacrylamide gel electrophoresis, is shown in Figure A-7. Two sequencing runs were performed with samples purified at different times (20 and 35 nmole, respectively). The HPLC

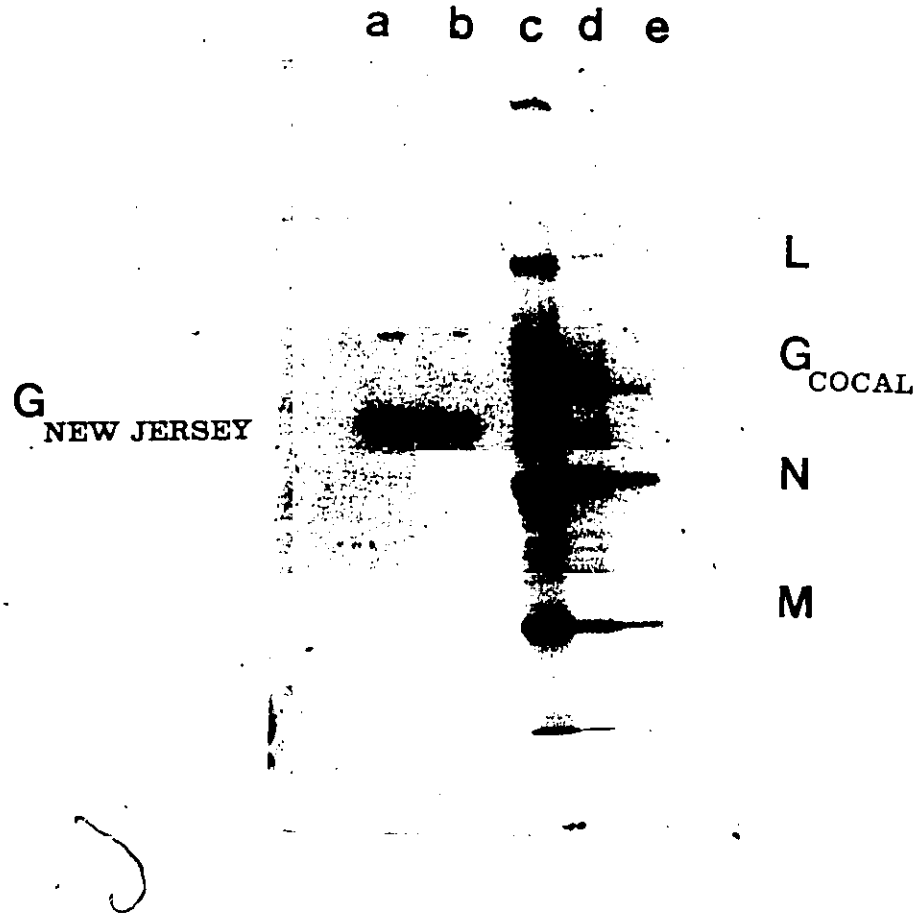


Figure A-7: Polyacrylamide gel analysis of purified New Jersey G protein.

The G protein from New Jersey (Concan) VSV was purified by detergent extraction as described in Methods. Aliquots of the purified material were analyzed on a 10% polyacrylamide gel and the proteins were visualized by staining with Coomassie Blue. Lanes a and b, 20 and 10 µg of purified New Jersey G protein, respectively; lane c, pellet remaining after solubilization of New Jersey VSV with Triton X100 and centrifugation; lanes d and e, 30 and 10 µg of Cocal VSV, respectively.

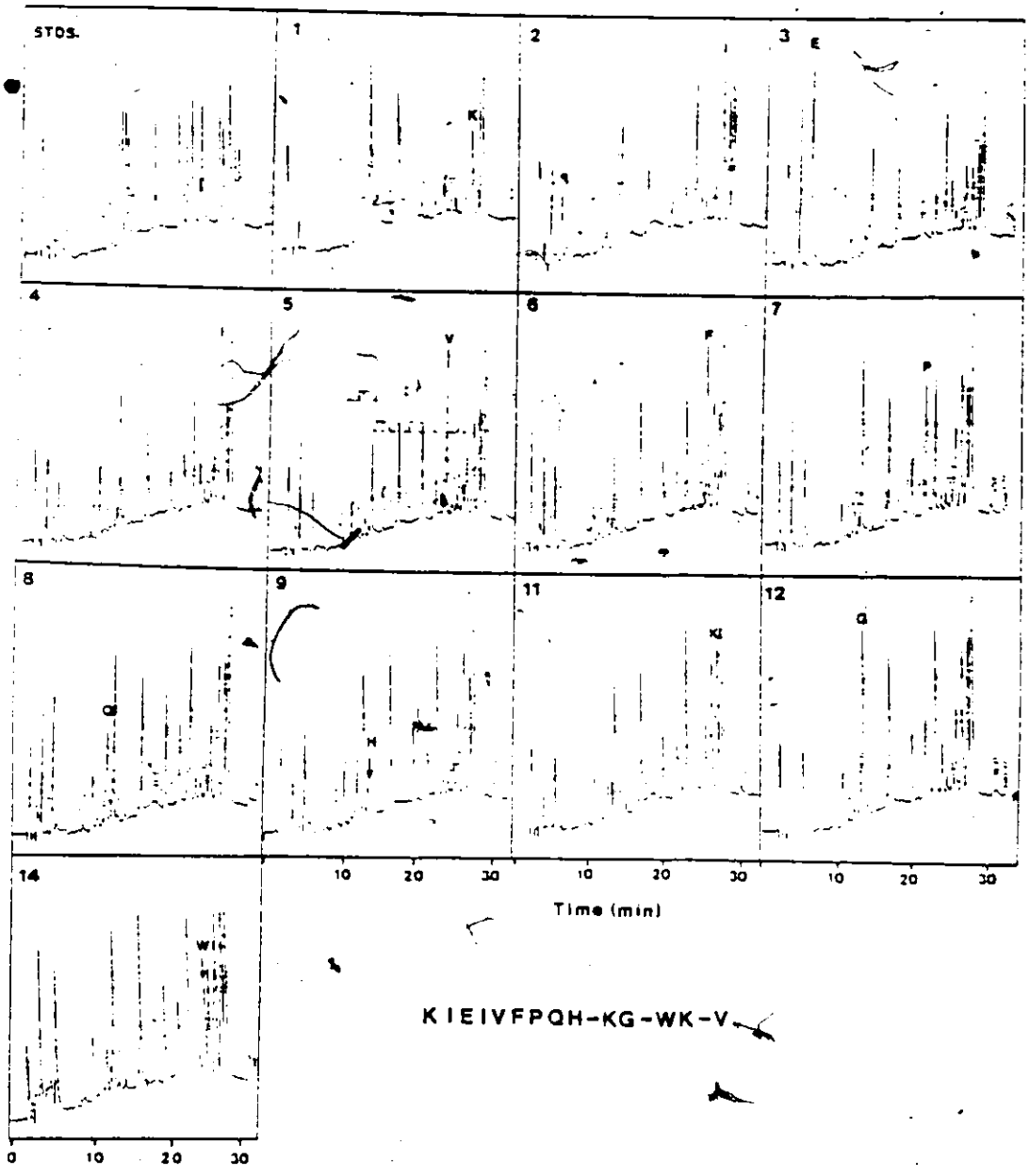
analysis (Fig. A-8) revealed a high background which was evident in two separate sequence runs. The anamolous peaks may have resulted from partial acid cleavage of the protein during exposure to HFBA resulting in asynchrony. It is probably not due to a heterogenous amino terminus since the background peaks were generally consistent in sequential cycles. A more probable explanation is the presence of UV absorbing contaminants in the reagents and solvents. The most probable source of contamination would be the HFBA and the ethyl acetate. Impurities present in ethyl acetate have been shown to react with PITC and PTH derivatives to produce many artifactual peaks on HPLC analysis and to reduce the yields of amino acids (Shively et al., 1982).

Nevertheless, amino acid assignments were unambiguously determined for positions 1-8. Cycle 9 His was assigned on the basis of a rise in the PTH-His peak relative to cycles 8 and 10. Similarily, Trp was assigned in cycle 14. The yields from the degradation of 35 nmole (2.2 mg) of New Jersey G protein are shown in Table A-3. Subsequently, the assignments of Lys_1, $Ile_{3,4}$, Val_5, and Phe_6, were confirmed by radiosequence analysis (Kotwal et al., in press).

A comparison of the sequence from the mature glycoproteins of Cocal and New Jersey (Concan) VSV as determined from this work and the sequence of G protein from Indiana (San Juan) and New Jersey (Ogden) as determined from DNA sequence analysis of the cloned genes is shown in Figure A-9. As can be seen, there is a great deal of homology between the different serotypes examined, and in particular among

Figure A-8: HPLC analysis of PTH amino acids obtained from sequential degradation of New Jersey G protein.

Results are from the sequence of 35 nmoles of New Jersey (Concan) G protein. The cycle number is given in the top left hand corner and the letter indicates the assigned amino acid. PTH-norleucine was the internal standard. Absorbance was at 254 nm using 0.08 AUFS. The deduced amino acid sequence is given in the bottom of the figure.



K I E I V F P Q H – K G – W K – V

Table A-3: Analysis of the PTH-derivatives obtained by sequential degradation of 35 nmoles of New Jersey G protein by HPLC. The repetitive yield was 89 to 91%.

Cycle No.	PTH-Amino Acid Identified*	Yield (nmole)
1	Lys	2.1
2	Ile	19.5
3	Glu	20.4
4	Ile	15.4
5	Val	18.1
6	Phe	17.1
7	Pro	18.8
8	Gln	18.8
9	His	-
10	N.D.	-
11	Lys	3.6
12	Gly	12.4
13	N.D.	-
14	Trp	2.8
15	Lys	1.4
16	N.D.	-
17	Val	5.7
18	N.D.	-

* N.D. NOT DETERMINED

FIGURE A-9: AMINO TERMINAL SEQUENCES OF MATURE GLYCOPROTEINS FROM DIFFERENT SEROTYPES OF VSV*

		1 10 20
a)	INDIANA SAN JUAN	LysPheThrIleValPheProHisAsnGlnLysGlyAsnTrpLysAsnValProSerAsnTyrHisTyrCys-
b)	INDIANA TORONTO	LysPheThrIleValPheProTyrAsnGlnLysGlyAsnTrpLysAsnValProSerAsnTyrHisTyrCys-
c)	COCAL	LysPheSerIleValPheProGlnSerGlnLysGlyAsxTrpLysAsxValProSerSerTyrTyrTyr-
d)	NEW JERSEY CONCAN	LysIleGluIleValPheProGlnHis - LysGly - TrpLys - Val - - - - - - - -
e)	NEW JERSEY OGDEN	LysIleGluIleValPheProGlnHisThrThrGlyAspTrpLysArgValProHisGluTyrAsnTyrCys
	CONCENSUS SEQUENCE	Lys - - IleValPhePro - - - - GlyAsnTrpLys - ValPro - - Tyr - TyrCys

* This figure was taken from Kotwal et. al. (in press).

a) determined from cDNA of cloned gene (Rose and Gallione, 1981).

b) determined by amino acid sequence (Irving et. al., 1979; Kotwal et. al., in press).

c) and d) determined from this work.

e) determined from cDNA of cloned gene (Gallione and Rose, 1983).

different strains of the same serotype. Most of the differences can be accounted for by single base changes at the nucleotide level.

The highly conserved amino terminal domain of these different glycoproteins is especially intriguing in light of the fact that the signal sequences present in the precursor forms of the glycoproteins from Indiana, Cocal, and New Jersey, show very little sequence homology (Kotwal et al., in press). The amino terminal domain may thus impart some important but as yet unidentified structural or functional characteristic.

BIBLIOGRAPHY

Abraham, G. and Banerjee, A.K. (1976) Virology 71, 230

Abraham, G., Rhodes, D.P. and Banerjee, A.K. (1975) Cell 5, 51

Air, G.M., Blok, J., Hall, R.M. (1981) In Replication of Negative Strand Viruses, Ed. Bishop, D.H.L., Compans, R.W., p. 225, Elsivier North Holland: New York

Aitstiel, L.D. and Landsberger, F.R. (1981) Virology 115, 1

Alt, F.W., Bothwell, A., Knapp, M., Siden, E., Mather, E., Koshland, M. and Baltimore, D. (1980) Cell 20, 293

Bailey, G.S., Gillett, D., Hill, D.F. and Petersen, G.B. (1977) J. Biol. Chem. 252, 2218

Ball, L.A. (1977) J. Virol. 21, 411

Ball, L.A. and Wertz, G.W. (1981) Cell 26, 143

Ball, L.A. and White, C.N. (1976) Proc. Natl. Acad. Sci. USA 73, 442

Baltimore, D., Huang, A.S., Stampfer, M. (1970) Proc. Natl. Acad. Sci. USA 66, 572

Banerjee, A.D., Abraham, G. and Colonno, R.J. (1977) J. Gen. Virol. 34, 1

Banerji, J., Rusconi, S. and Schaffner, W. (1981) Cell 27, 299

Barenholz, Y., Moore, N.F. and Wagner, R.R. (1976) Biochemistry 15, 3563

Bassford, P. and Beckwith, J. (1979) Nature 277, 583

Baxt, B., Sonnabend, J.A. and Bablanian, R. (1977) J. Gen. Virol. 35, 325

Bayley, H. and Knowles J.R. (1978a) Biochemistry 17, 2420

Bayley, H. and Knowles, J.R. (1978b) Biochemistry 17, 2414

Bedouelle, H., Bassford, J., Fowler, A.V., Zabin, I., Beckwith, J. and Hofnung, M. (1980) Nature 285, 78

Bell, J.R., Hunkapiller, M.W., Hood, L.E. and Strauss, J.H. (1978) Proc. Natl. Acad. Sci. USA 75, 2722

Bergeron, J.J.M., Kotwal, G.J., Levine, G., Bilan, P., Rachubinski, R., Hamilton, M., Shore, G.C. and Ghosh, H.P. (1982) J. Cell Biol. 94, 36

Bergman, J.E., Tokuyasu, K.T. and Singer, S.J. (1981) Proc. Natl. Acad. Sci. USA 78, 1746

Besmer, P. and Baltimore, D. (1977) J. Virol. 21, 965

Beynum, G.M.A., van DeGraaf, J., Castel, A., Kraal, B. and Bosch, L. (1977) Eur. J. Biochem. 72, 63

Bhown, A.S., Cornelius, T.W., Mole, J.E., Lynn, J.D., Tidwell, W.A. and Bennet, J.C. (1980) Anal. Biochem. 102, 35

Bishop, D.H.L., Aaslestad, J.G., Clark, H.F., Flamand, A., Obijeski, J.S., Repik, P. and Roy, P. (1975b) In Negative Strand Viruses Vol 1, B.W.J. Many, R.D. Bary eds. p. 259, Academic Press: London

Bishop, D.H.L. and Smith, M.S. (1977) Rhabdoviruses, In The Molecular Biology of Animal Viruses, D. Nayak ed. p. 167, Marcel Dekker Press: New York

Bishop D.H.L., Repik, P., Obijeski, J.F., Moore, N.F. and Wagner, R.R. (1975a) J. Virol. 16, 75

Black, S.D. and Coon, M.J. (1982) Anal. Biochem. 121, 281

Bligh, E.G. and Dyer, W.J. (1959) Can. J. Biochem. Physiol. 37, 911

Blobel, G. (1980) Proc. Natl. Acad. Sci. USA 77, 1496

Blobel, G. and Dobberstein, B. (1975a) J. Cell Biol. 67, 835

Blobel, G. and Dobberstein, B. (1975b) J. Cell Biol. 67, 852

Blobel, G. and Sabatini, D.D., (1970) J. Cell Biol. 45, 130

Blumberg, B.M., Giorgi, C. and Kolakofsky, D. (1983) Cell 32, 559

Blumberg, B.M., Leppert, M., Kolakofsky, D. (1981) Cell 23, 837

Boettiger, D. (1979) Prog. Med. Virol. 25, 37

Bonner, W.M. and Laskey, R.A. (1974) Eur. J. Biochem 46, 83

Both, G.W., Moyer, S.A. and Banerjee, A.K. (1975a) Proc. Natl. Acad. Sci. USA 72, 274

Both, G., Moyer, S.A. and Banerjee, A.K. (1975b) J. Virol. 18, 1015

Brauer, A.W., Margolies, M.N. and Haber, E. (1975) Biochem. 14, 3029

Brown, F. Small, C.J. and Horzinek, M.C. (1974) J. Gen. Virol. 22, 255

Brunner, J. (1981) Trends in Biochem. Sci. 6, 44

Brunner, J., Senn, H. and Richards, F.M. (1980) J. Biol. Chem. 255, 3313

Bucher, D.J., Kharitonenkow, L.G., Zakomiridin, V.B., Grigoriev, S.M., Klimenko, S.M. and Davis, J.F. (1980) J. Virol. 36, 586

Burge, B.W. and Huang, A.S. (1970) J. Virol. 6, 176

Burge, B.W. and Huang, A.S. (1979) Virology 95, 445

Bussereau, F., Cartwright, B., Doel, T.R. and Brown, F. (1975) J. Gen. Virol. 29, 189

Canefield, R.E. (1963) J. Biol. Chem. 238, 2698

Canefield, R.E. (1965) J. Biol. Chem. 240, 1997

Capone J., Toneguzzo, F. and Ghosh, H.P. (1982) J. Biol. Chem. 257, 16

Carmicheal, G.C., Schaffhausen, B.S., Dorsky, D.I., Oliver, D.B. and Benjamin, T.L. (1982) Proc. Natl. Acad. Sci. USA 79, 3579

Carr, S.A., Biemann, K., Shoji, S., Parmalee, D.C. and Titani, K. (1982) Proc. Natl. Acad. Sci. USA 79, 6128

Carroll, R.B. and Smith, A.E. (1976) Proc. Natl. Acad. Sci. USA 73, 2254

Carroll, A.R. and Wagner, R.R. (1978) J. Virol. 25, 675

Carroll, A.R. and Wagner, R.R. (1979) J. Virol. 29, 134

Cartwright, B. (1977) J. Gen. Virol. 34, 249

Cartwright, B., Smale, C.J. and Brown, F. (1969) J. Gen. Virol. 5, 1

Cartwright, B., Talbot, P. and Brown, R. (1969) J. Gen. Virol. 7, 267

Cartwright, G.S., Smith, L.M., Heinzelman, E.W., Ruebush, M.J., Parce, J.W. and McConnel, H.M. (1982) Proc. Natl. Acad. Sci. USA 79, 1506

Cepko, C.L., Hansen, U., Handa, H. and Sharp, P. (1981) Mol. Cell. Biol. 1, 919

Chanda, P.K. and Banerjee, A.K. (1981) J. Biol. Chem. 256, 11393

Chatis, P.A. and Morrison, T.G. (1979) J. Virol. 29, 957

Chatis, P.A. and Morrison, T.G. (1981) J. Virol. 37, 307

Cheng, H.L., Blattner, F.R., Fitzmaucile, L., Mushinski, J.F. and Tucker, P.W. (1982) Nature 296, 410

Chinchar, V.G., Amesse, L.S. and Portner, A. (1982), Biochem. Biophys. Res. Commun. 105, 1296

Chow, J.M., Schnitzlein, W.M. and Reichmann, M.E. (1977) Virology 77, 579

Chowdhry, V. and Westheimer, F.H. (1979) Ann. Rev. Biochem. 48, 293

Chu, G. and Sharp, P.A. (1981) Gene. 13, 197

Cleveland, D.W., Fischer, S.G., Kirschner, M.W. and Laemmli, U.K. (1977) J. Biol. Chem. 252, 1102

Clinton, G.M., Burge, B.W. and Huang, A.S. (1978a) J. Virol, 27, 340

Clinton, G.M., Burge, B.W. and Huang, A.S. (1979) Virology 99, 84

Clinton, G.M., Guerina, N.G., Guo, H. and Huang, A.S. (1982) J. Biol. Chem. 257, 3315

Clinton, G.M. and Huang, A.S. (1981) Virology 108, 510

Clinton, G.M., Little, S.P., Hagen, F.S. and Huang, A.S. (1978b) Cell 15, 1455

Colonno, R.J. and Banerjee, A.K. (1976) Cell 8, 197

Colonno, R.J. and Banerjee, A.K. (1977) Virology 77, 260

Colonno, R.J. and Banerjee, A.K. (1978) Cell 15, 93

Combard, A. and Printz-Ane, C. (1979) Biochem. Biophys. Res. Commun. 88, 117

Combard, A. Printz-Ane, C., Martinet, C. and Printz, P. (1977) J. Viro 21, 913

Courtneidge, S.A., Levinson, A.D. and Bishop, J.M. (1980) Proc. Natl. Acad. Sci. USA 77, 3783

Crewther, W.G. and Inglis, A.S. (1975), Anal. Biochem. 68, 572

Croft, L. R. (1980) Handbook of Protein Sequence Analysis, John Wiley and Sons, N.Y.

Dahlberg, J.E. (1974) Virology 58, 250

David, A.E. (1977) Virology 76, 98

David, A.E. (1978) J. Gen. Virol. 39, 149

Davis, N.L. and Wertz F.W. (1982), J. Virol. 41, 821

De, P.B., Thornton, G.B., Luk, D. and Banerjee, A.K. (1982) Proc. Natl. Acad. Sci. USA 79, 7137

Dickerson, R.E. and Geis, I. (1969) The Structure and Action of Proteins, Harper and Row: London

Dickson, R.B., Willingham, M.C. and Pastan, I. (1981) J. Cell Biol. 89, 29

Doel, T.R. and Brown, F. (1978) J. Gen. Virol. 38, 351

Dubovi, E.J. and Wagner, R.R. (1977) J. Virol. 22, 500

Dulbecco, R. and Vogt, M. (1954) J. Exptl. Med. 99, 167

Dunigan, B.D. and Lucas-Lennard, J.M. (1983) J. Virol. 45, 618

Dunphy, W.G., Fries, E., Urban, L.J. and Rothman, J.E. (1981) Proc. Natl, Acad. Sci. USA 78, 7453

Edman, P. (1970) Protein Sequence Determination, S.B. Needleman ed. p. 211, Springer-Verlag, N.Y.

Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80

Edmonson, A.B. (1965) Nature 205, 883

Eger, R., Compans, R.W. and Rifkin, D. (1975) Virology 66, 610

Emerson, S.U. and Wagner, R.R. (1972) J. Virol. 10, 297

Emerson, S.U. and Wagner, R.R. (1973) J. Virol. 12, 1325

Emerson, S.U. and Yu, Y.H. (1975) J. Virol. 15, 1345

Engelman, D.M. and Steitz, T.A. (1981) Cell 23, 411

Etchison, J.R. and Holland, J.J. (1974a) Virology 60, 217

Etchison, J.R. and Holland, J.J. (1974b) Proc. Natl. Acad. Sci. USA 71, 4011

Etchison, J.R., Roberston, J.S. and Summers, D.F. (1977) Virology 78, 375

Etchison, J.R., Roberston, J.E. and Summer, D.F. (1981) J. Gen. Virol. 57, 43

Evans, D., Pringle, C.R. and Szil'agyi, J.F. (1979) J. Virol. 31, 325

Fan, D.P. and Sefton, B.M. (1978) Cell 15, 985

Feuer, B.I., Uzgiris, E.E., Deblois, R.W., Cluxton, D.H. and Lenard, J. (1978) Virology 90, 156

Fong, B.S., Hunt, R.C. and Brown, J.C. (1976) J. Virol. 20, 658

Fries, E. and Rothman, J.E. (1980) Proc. Natl. Acad. Sci. USA 77, 3870

Gallione, C.J., Greene, J.R., Iverson, L.E. and Rose, J.K. (1981) J. Virol. 39, 529

Gallione, C.J. and Rose, J.K. (1983) J. Virol. 46, 162

Garoff, H. (1974) Virology 62, 385

Garoff, H., Frischauf, A.M., Simons, K, Lehrach, H., and Delius, H. (1980) Nature 288, 236

Garber, E.A., Kruefer, J.G., Hanafusa, H. and Goldberg, A.R. (1983) Nature 302, 161

Gething, J., White, J.M. and Waterfield, M.D. (1978) Proc. Natl. Acad. Sci. USA 75, 2737

Ghosh, H.P. (1980) Rev. Infec. Dis. 2, 26

Ghosh, K. and Ghosh, H.P. (1972) J. Biol. Chem. 247, 3369

Ghosh, K. and Ghosh, H.P. (1982) Nuc. Acid Res. 10, 6341

Ghosh, H.P., Toneguzzo. F. and Wells, S. (1973) Biochem. Biophys. Res. Commun. 54, 228

Gibson, R.S., Kornfeld, S. and Schlesinger, S. (1981) J. Biol. Chem. 256, 456

Gibson, R., Leavitt, R. Kornfeld, S. and Schlesinger, S. (1978) Cell 13, 671

Gibson, R., Schlesinger, S. and Kornfeld, S. (1979) J. Biol. Chem. 254, 3600

Gibson, R., Schlesinger, S. and Kornfeld, S. (1980) Trends in Biochem. Sci. 5, 290

Gilmore, R., Blobel, G. and Walter, P. (1982a) J. Cell Biol. 95, 463

Gilmore, R., Walter, P. and Blobel, G. (1982b) J. Cell Biol. 95, 470

Gluzman, Y. (1981) Cell 23, 175

Goldfine, H., Harley, J.B. and Wyke, J.A. (1978) Biochem. Biophys. Acta. 512, 229

Gottlieb, C., Kornfeld, S. and Schlesinger, S. (1979) J. Virol. 29, 344

Graham, F.L. and van der Eb, A.J. (1973) Virology 52, 456

Green, R.T., Meiss, H.K., and Rodriguez-Boulan, E. (1981) J. Cell Biol. 89, 230

Greenberg, G.R., Chakrabarti, P. and Khorana, H.G. (1976) Proc. Natl. Acad. Sci. USA 73, 86

Gregoriades, A. (1980) J. Virol. 36, 470

Gregoriades, A. and Frangione, B. (1981) J. Virol. 40, 323

Griffiths, G., Quinn, P. and Warren, G. (1983) J. Cell Biol. 96, 835

Grinna, L.S. and Robbins, P.W. (1979) Fed. Proc. 38, 291

Grubman, M.J., Moyer, S.A., Banerjee, A.K. and Ehrenfeld, E. (1975) Biochem. Biophys. Res. Commun. 62, 531

Grubman, M.J. and Shafritz, D.A. (1977) Virology 81, 1

Grubman, M.J., Weinstein, J.A. and Shafritz, D.A. (1977) J. Cell Biol. 74, 43

Gupta, C.M., Costello, C.E. and Khorana, H.G. (1979) Proc. Natl. Acad. Sci. USA 76, 3139

Gupta, C.M., Radhakrishan, R., Gerber, G.E., Olsen, W.L., Quay, J.C. and Khorana, H.G. (1979) Proc. Natl. Acad. Sci. USA 76, 2595

Gupta, C.M., Radhakrishan, R. and Khorana, H.G. (1977) Proc. Natl. Acad. Sci. USA 74, 4315

Habener, J.T., Rosenblatt, M., Kemper, B., Kronenberg, H.M., Rich, A. and Potts, J.J. (1978) Proc. Natl. Acad. Sci. USA 75, 2616

Hale, A.H., Witte, O.N., Baltimore, D. and Eisen, H.N. (1978) Proc. Natl. Acad. Sci. USA 75, 970

Hantke, K. and Braun, V. (1973), Eur J. Biochem. 34, 284

Haspel, M.V., Pellegrino, M.A., Lampert, P.W. and Oldstone, M.B.A. (1977) J. Exp. Med. 146, 146

Hawke, D., Yuan, P.M. and Shively, J.E. (1982) Anal. Biochem. 120, 302

Hecht, T.T. and Summers, D.F. (1972) J. Virol. 10, 578

Hecht, T.T. and Summers, D.F. (1976) J. Virol. 19, 833

Heine, J.W. and Schnaitman, C.A. (1969) J. Virol. 3, 619

Heine, J.W. and Schnaitman, C.A. (1971) J. Virol. 8, 786

Helenius, A., Fries, E. and Kartenbeck, J. (1977) J. Cell Biol. 75, 866

Helenius, A. and Simons, K. (1972) J. Biol. Chem. 247, 3656

Henderson, L.E., Copeland, T.D. and Oroszlan, S. (1980) Anal. Biochem. 102, 1

Henderson, L.E., Krutzsch, H.C. and Oroszlan, S. (1983) Proc. Natl. Acad. Sci. USA 80, 339

Herman, R.C., Adler, S., Lazzarini, R.A., Colonno, R.J., Banerjee, A.K. and Westphal, H. (1978) Cell 15, 587

Herman, R.C., Shubert, M., Keene, J.D. and Lazzarini, R.A. (1980) Proc. Natl. Acad. Sci. USA 77, 4662

Hermodson, M.A., Ericsson, L.H., Titani, K., Neurath, H. and Walsh, K.A. (1972) Biochem. 11, 4493

Hill, V.M., Marness, L. and Summers, D.F. (1981) Virology 113, 109

Holland, J.J., Doyle, M., Perrault, J., Kingsbury, D.T. and Etchison, J. (1972) Biochem. Biophys. Res. Commun. 46, 634

Horn, M.J. and Bonner, A.G. (1977) In Solid Phase Methods in Protein Sequence Analysis. Inserm. Symposium No. 5., Previero, A., Colett-Previero, M.A., eds., Elsivier North Holland

Huang, A.S. (1977) Bacteriol. Rev. 41, 811

Huang, A.S., Baltimore, D., and Stampfer, M. (1970) Virology 42, 946

Huang, A., Huang, L. and Kennel, S.J. (1980) J. Biol. Chem. 255, 8015

Huang, R.T.C., Wahn, K., Klenk, H.D. and Rott, R. (1979) Virology 97, 212

Huang, R.T.C., Wahn, K., Klenk, H.D. and Rott, R. (1980) Virology 104, 294

Hunkapillar,M.W. and Hood, L.E. (1978) Biochem. 17, 2124

Hunt, D.M., Emerson, S.U. and Wagner, R.R. (1976) J. Virol. 18, 603

Hunt, L.A., Etchison, J.R. and Summers, D.F. (1978) Proc. Natl, Acad. Sci. USA 75, 754

Hunt, L.A. and Summers, D.F. (1977) J. Supramol. Struct. 7, 213

Hsu, C.H. and Kinsbury, D. (1980) Animal Virus Genetics, Fields, B.N., Aenish, J., Fox, C.F., eds. p. 613. Academic Press: New York

Hsu. C.H., Kingsbury, D.W. and Murti, K.G. (1979) J. Virol. 32, 304

Hsu, M.C., Schied, A. and Choppin, P.W. (1979) Virology 95, 476

Inblum, R.L. and Wagner, R.R. (1974) J. Virol. 13, 113

Inblum, R.L. and Wagner, R.R. (1976) J. Virol. 15, 1357

Inouye, M., and Halegoua, S. (1980) Crc. Crit. Rev. Biochem. 10, 339

Irving, R.A., Hofmann, T. and Ghosh, H.P. (1982) FEBS Lett. 104, 257

Irving, R.A., Toneguzzo, F., Rhee, S.H., Hofmann, T. and Ghosh, H.P. (1979) Proc. Natl. Acad. Sci. USA 76, 570

Iverson, C.E. and Rose, J.K. (1982) J. Virol., 356

Jackson, R.C., Walter, P. and Blobel, G. (1980) Nature 284, 174

Jacobs. J.W., Kemper, B., Niall, H.G., Habener, J.F. and Potts, J.T. (1974) Nature 249, 155

Jaye, M.C., Godchaux III, W. and Lucas-Lenard, J. (1982) Virology 116, 148

Johnson, D.C. and Schlesinger, M.J. (1980) Virology 103, 407

Johnson, D.C., Schlesinger, M.J. and Elson, E.L. (1981) Cell 23, 473

Johnson, D.C. and Spear, P.G. (1983) Cell 32, 987

Johnson, N.D., Hunkapiller, M.W. and Hood, L.E. (1979) Anal. Biochem. 100, 335

Kang, C.Y. and Lambright, P. (1977) J. Virol. 21, 1252

Kang, C.Y. and Prevec, L. (1969) J. Virol. 3, 404

Kang, C.Y. and Prevec, L. (1970) J. Virol. 6, 20

Kates, M., (1972) Techniques of Lipidology, Work, T.S. and Work, E., eds., North Holland/American Elsivier: New York

Katz, F.N. and Lodish, H.F. (1979) J. Cell Biol. 80, 416

Katz, R.N., Rothman, J.E., Lingappa, V., Blobel, G. and Lodish, H.F. (1977) Proc. Natl. Acad. Sci. USA 74, 3278

Kaufman, R.J. and Sharp, P.A. (1982) Mol. and Cell. Biol. 2, 1304

Keene, J.D., Schuber, M. and Lazzarini, R.A. (1979), J. Virol. 32, 167

Keene, J.D., Thornton, B.J. and Emerson, S.U. (1981) Proc. Natl. Acad. Sci. USA 78, 6191

Kehry, M.S. Ewald, S., Douglas, R., Sibley, C. Raschke, W., Fambrough, D. and Hood, L. (1980) Cell 21, 393

Keller, P.M., Uzgiris, E.E., Cluxton, D.H. and Lenard. J. (1978) Virology 87, 66

Kelly, J.M., Emerson, S.U. and Wagner, R.R. (1972) J. Virol. 10, 1231

Kingsford, L. and Emerson, S.U. (1980a) J. Virol. 33, 1097

Kingsford, L., Emerson, S.U. and Kelley, J.M. (1980b) J. Virol. 36, 309

Kinney, R.M. and Trent, D.W. (1982) Virology, 121, 345

Klapper, D.G., Wilde, G.E. and Capra, J.D. (1978) Anal. Biochem. 85, 126

Klenk, H.D., Wollert, W., Rott, R. and Scholtissek, G. (1974) Virology 57, 28

Klockmann, U. and Leppert, W. (1983) FEBS Lett. 151, 257

Knipe, D.M., Baltimore, D. and Lodish, H.F. (1977a) J. Virol. 21, 1149

Knipe, D.M., Lodish, H.F. and Baltimore, D. (1977b) J. Virol. 21, 1121

Knipe, D.M., Baltimore, D. and Lodish, H.F. (1977c) J. Virol. 21, 1128

Knipe, D., Lodish, H.F. and Baltimore, D. (1977d) J. Virol. 21, 1140

Knipe, D., Rose, J.K. and Lodish, H.F. (1975) J. Virol. 15, 1004

Knudson, D.L. (1973) J. Gen. Virol. 20, 105

Kornfeld, S., Li, E. and Tabas, I. (1978) J. Biol. Chem. 253, 7771

Koshland, D. and Botstein, D. (1980) Cell 20, 749

Krag, S.S., Cifone, M., Robbins. P.W. and Baker, R.M. (1977) J. Biol. Chem. 252, 3561

Krebs, E.G. and Beavo, J.A. (1979) Ann. Rev. Biochem. 48, 923

Kruppa, J. (1979) J. Biochem. 181, 295

Kurilla, M.G., Pinnica-Worms, H. and Keene, J.D. (1982) Proc. Natl. Acad. Sci. USA 79, 5240

Kvist, S., Wiman, K. Claesson, L., Peterson, P. and Dobberstein, B. (1982) Cell 29, 61

Laemmli, U.K. (1970) Nature 227, 680

Lai, J.S., Philbrick, W.N. and Wu, H.C. (1980) J. Biol. Chem. 255, 5384

Lai, J.S., Sarvas, M., Brammar, W.J., Neugebauer, K. and Wu, H.C. (1981) Proc. Natl. Acad. Sci. USA 78, 3506

Landsberger, F.R. and Aitstiel, L.D. (1980) Ann. NY Acad. Sci. 348, 419

Landsberger, F.R. and Compans, R.W. (1976) Biochem. 18, 2356

Laskey, R.A. and Mills, A.D. (1975) Eur. J. Biochem. 56, 335

Laver, W.G. (1971) Virology 45, 275

Lazzarini, R.A., Chien, I., Yang, Y. and Keene, J.D. (1982) J. Gen. Vir. 58, 429

Lazzarini, R.A., Keene, J.D. and Schubert, M. (1981) Cell 26, 145

Leamnson, R.N. and Halpern, M.S. (1976) J. Virol. 18, 956

Leavitt, R., Schlesinger, S. and Kornfeld, S. (1977) J. Biol. Chem. 252, 9018

Leblanc, P., Capone, J. and Gerber, G.E. (1982) J. Biol. Chem. 257, 14586

Legault, D., Pakayesu, D. and Prevec, L. (1977) J. Gen. Virol. 351, 53

Lenard, J. (1978) Ann. Rev. Biophys. Bioeng. 7, 139

Lenard, J. and Miller, D.K. (1982) Cell 28, 5

Lepage, M. (1964) J. Lipid Res. 5, 587

Leppert, M., Rittenhouse, L., Perrault, J., Summers, D.F. and Kolakofsky, D. (1979) Cell 18, 735

Lesnaw, J.A., Dickson, L.R. and Curry, R.H. (1979) J. Virol. 31, 8

Lesnaw, J.A. and Reichman, M.E. (1975) Virology 63, 492

Li. E., Tabas, I. and Kornfeld, S. (1978) J. Biol. Chem. 253, 7762

Lin, J.C., Knanzawa, H., Ozols, J. and Wu, H.C. (1978) Proc. Natl. Acad. Sci. USA 75, 4891

Lingappa, V.R., Katz, F.N., Lodish, H.F. and Blobel, G. (1978a) J. Biol. Chem. 253, 8667

Lingappa, V.R., Shields, D., Woo, S.L.C. and Blobel, G. (1978b) J. Cell. Biol. 79, 567

Little, L.M., Zavada, J., Der, C.J. and Huang, A.S. (1983) Science 220, 1069

Little, S.P. and Huang, A.S. (1977) Virology 81, 37

Little, S.P. and Huang, A.S. (1978) J. Virol. 27, 330

Lodish, H.F. (1980) Bioch. Soc. Symp. 45, 105

Lodish, H.F., Katz, F.N., Rothman, J.E. and Knipe, D.M. (1978) Birth Defects 14, 155

Lodish, H.F. and Kong, N. (1983) Virology 125, 335

Lodish, H.G. and Porter, M. (1980a) J. Virol. 33, 52

Lodish, H.F. and Porter, M. (1980b) J. Virol. 36, 719

Lodish, H.F. and Porter, M. (1980c) Cell 19, 161

Lodish, H.F. and Porter, M. (1981) J. Virol. 38, 504

Lodish, H.F. and Weiss, R.A. (1979) J. Virol. 30, 177

Lodish, H.F., Zilberstein, A. and Porter, M. (1981) Methods in Cell Biol. 23, 5

Lowry, O.H., Rosenbrough, N.J., Farr, A.L. and Randall, R.J. (1951) J. Biol. Chem. 193, 265

Lynch, K.R., Pennica, D., Ennis, H.L. and Cohen, P.S. (1979) Virology 98, 251

Lynch, K.R., Pennica, D., Ennis, H. and Cohen, P.S. (1981) Virology 108, 277

MacLennan, D.H., Yip, C.C., Iles, G.H. and Seeman, P. (1973) Cold Spring Harbor Symp. Quant. Biol. 33, 469

Magee, A. and Schlesinger, M.J. (1982) Biochem. Biophys, Acta. 649, 279

Mancarella, D.A. and Lenard, J. (1981) Biochemistry 20, 6872

Manley, J.L., Fire, A., Cano, A., Sharp, P.A. and Gefter, M. (1980) Proc. Natl. Acad. Sci. USA 77, 3855

Marcus, P.I., Sekellick, M.J., Johnson, L. and Lazzarini, R.A. (1977) Virology 82, 242

Margolies, M.N. and Brauer (1978) J. Chromatography 148, 429

Marnell, L.L. and Wertz, G.W. (1979) Virology 98, 88

Marnell, L.L. and Wertz, G.W. (1982) J. Gen. Virol. 60, 165

Marsh, M., Balzau, E. and Helenius, A. (1983) Cell 32, 931

Martinet, C., Combard, A., Printz-Ane, C. and Printz, P. (1979) J. Virol. 29, 123

Marvaldi, J., Lucas-Lenard, J., Sekellick, M. and Marcus, P.I. (1977) Virology 79, 267

Marvaldi, J., Sekellick, M.J., Marcus, P.I. and Lucas-Lenard, J. (1978) Virology 84, 127

Matlin, K.S., Reggio, H., Helenius, A. and Simons, K. (1982) J. Mol. Biol. 156, 609

Mellon, M.G. and Emerson, S.U. (1978) J. Virol. 27, 560

Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981) Cell 27, 279

Meyer, D.I. (1982) Trends in Biochem. Sci. 7, 320

Meyer, D.I. and Dobberstein, B. (1980) J. Cell Biol. 87, 503

Meyer, D.I., Krause, I. and Dobberstein, B. (1982a) Nature 297 647

Meyer, D.I., Louvard, I., Dobberstein, B. (1982b) J. Cell Biol. 92, 579

Miller, K.D., Feuer, B., Vanderoef, R. and Lenard, J. (1980) J. Cell Biol. 84, 421

Miller, D.K. and Lenard, J. (1980) J. Cell Biol. 84, 430

Min Jou W. Verhoeyen, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W., Threifall, G., Barber, C., Carey, M. and Emtage, S. (1980) Cell 19, 683

Montelars, R.C. and Ruecker, R.R. (1977) Arch. Biochem. Biophys. 178, 555

Moore, N.T., Patzer, E:F., Shaw, J.M., Thompson, T.F. and Wagner, R.R. (1978) J. Virol. 27, 320

Moreno, F., Fowler, A., Hall, M., Silhavy, T.F., Zabin, I. and Schwartz, M. (1980) Nature 286, 356

Morrison, T.G. and Lodish, H.F. (1975) J. Biol. Chem. 250, 6955

Morrison, T.G. and McQuain, C. (1977) J. Virol. 21, 451

Morrison, T.G. and McQuain, C.O. (1978) J. Virol. 26, 115

Morrison, T.G., McQuain, C.O. and Simpson, D. (1978) J. Virol. 28, 368

Morrison, T., Stampfer, M., Baltimore, D. and Lodish, H.F. (1974) J. Virol. 13, 62

Moyer, S.A. and Banerjee, A.K. (1975) Cell 4, 37

Moyer, S.A. and Summers, D.F. (1974a) Cell 2, 63

Moyer, S.A. and Summers, D.F. (1974b) J. Virol. 13, 455

Mudd, J.A. (1974) Virology 62, 573

Mudd, J.A. and Summers, D.F. (1970a) Virol. 42, 328

Mudd, J.A. and Summers, D.F. (1970b) Virol. 42, 958

Mudd, J.A. and Swanson, R.E. (1978) Virology 88, 263

McAllister, P.E. and Wagner, R.R. (1976) J. Virol. 18, 550

McGeoch, D.J. (1979) Cell 17, 673

McGeoch, D.J. (1981) J. Gen. Virol. 55, 1

McGowan, J.J., Emerson, S.U. and Wagner, R.R. (1982) Cell 28, 325

McSharry, J.J. (1977) Virology 83, 482

McSharry, J.J. and Choppin, P.W. (1978) Virology 84, 172

McSharry, J.J., Compans, R.W. and Choppin, P.W. (1971) J. Virol. 8, 722

McSharry, J.J., Ledda, C.A., Freiman, J.F. and Choppin, P.W. (1978) Virology 84, 183

McSharry, J.J. and Wagner, R.R. (1971) J. Virol. 1, 59

Naeve, C.W. and Summers, D.F. (1980) J. Virol. 34, 764

Naito, S. and Ishihama, A. (1976) J. Biol. Chem. 251, 4307

Nakai, T. and Howatson, A.F. (1968) Virology 35, 268

Narita, K. (1970) Protein Sequence Determination, S.B. Needleman, ed., p. 25, Spring Verlag, N.Y.

Neilson, J.B.K., Caulfield, M.P. and Lampen, J.O. (1981) Proc. Natl. Acad. Sci. USA 78, 3511

Newcomb, W.W. and Brown J.C. (1981) J. Virol. 39, 295

Newcomb, W.W., Tobin, G.J., McGowan, J.J. and Brown, J.C. (1982) J. Virol. 41, 1055

Niall, H.D. (1973) Automated Edman Degradation. The protein Sequenator, In Meth. Enzymology Hirs. C.H.W., Timasheff, S.N. ed. 27, 942

Nues, D.L. and Koch, G. (1976) J. Virol. 19, 572

Omary, M.B. and Trowbridge, I.S. (1981) J. Biol. Chem. 10, 4715

Orenstein, J., Johnson, L., Shelton, F. and Lazzarini, R.A. (1976) Virology 71, 291

Pal, R., Barenholz, Y. and Wagner, R.R. (1981) J. Biol. Chem. 256, 10209

Patton, J.T., Davis, N.L. and Wertz, G.W. (1983) J. Virol. 45, 155

Patzer, E.J., Moore, N., Barenholz, Y., Shaw, J.M. and Wagner, R.R. (1978) J. Biol. Chem. 253, 4544

Patzer, R.J., Wagner, R.R. and Dubovi, E.J. (1979) CRC Crit. Rec. Biochem. 6, 165

Pearse, B.M.F. and Bretscher, M. (1981) Ann. Rev. Biochem. 50, 85

Pelham, R.B. and Jackson, R.J. (1976) Eur. J. Biochem. 67, 247

Pennica, D., Lynch, K.R., Cohen, P.S. and Ennis, H.L. (1979) Virology 94, 484

Pepinsky, R.B. and Vogt, V.M. (1979) J. Mol. Biol. 131, 819

Perlman, S.M. and Huang, A.S. (1973) J. Virol. 12, 1395

Peters, K. and Richards, R.M. (1977) Ann. Rev. Biochem. 46, 523

Petri, W.A., Pal, R., Barenholz, Y. and Wagner, R.R. (1981) J. Biol. Chem. 256, 2625

Petri, W.A. and Wagner, R.R. (1979) J. Biol. Chem. 254, 4313

Petri, W.A. and Wagner, R.R. (1980) Virology, 107, 543

Petric, M. and Prevec, L. (1970) Virology 41, 615

Pinney, D.F. and Emerson, S.U. (1982) J. Virol. 42, 889

Piwnica-Worms, H. and Keene, J.D. (1983) Virology 125, 206

Ploegh, H.L., Orr, N.T. and Strominger, J.L. (1980) Proc. Natl. Acad. Sci. USA 77, 6081

Porter, A.G., Barber, C., Carey, N.H. Hallewell, R.A., Threlfall, G. and Emtage, J.S. (1979) Nature 282, 471

Prehn, S., Nurnberg, P. and Rapoport, T. (1981) FEBS Lett. 123, 79

Prehn, S.A., Tsamaloukas, A. and Rapoport, T.A. (1980) Eur. J. Biochem. 107, 185

Preston, C.M. and McGeough, D.J. (1981) J. Virol. 38, 593

Quay, S.C., Radhakrishnan, R. and Khorana, H.G. (1981) J. Biol. Chem. 256, 4444

Quinn, P., Griffiths, G. and Warren, G. (1983) J. Cell. Biol. 96, 851

Reading, C.L., Penhoet, E.E. and Ballou, C.E. (1978) J. Biol. Chem. 253, 5600

Reidler, J.A., Keller, P.M., Elson, E.L. and Lenard, J. (1981) Biochem. 20, 1345

Repik, P. and Bishop, D.H.L. (1973) J. Virol. 12, 969

Repik, P.N., Lamond, F. and Bishop, D.H.L. (1976) J. Virol. 20, 157

Rice, C.M., Bell, J.R., Hunkapillar, M.W., Strauss, E.G. and Strauss, J.H. (1982) J. Mol. Biol. 154, 355

Rice, C.M. and Strauss, J.H. (1981) Proc. Natl. Acad. Sci. USA 78, 2062

Robertson, M.A., Etchison, J.R., Robertson, J.S., Summers, D.F. and Stanley, P. (1978) Cell 13, 515

Robb, R.J., Terhorst, C. and Strominger, J.L. (1978) J. Biol. Chem. 253, 5319

Robbins, P.W., Hubbard, S.C., Turco, S.J. and Wirth, D.F. (1977) Cell 12, 893

Rodriguez-Boulan, E.R. and Pendergast, D. (1980) Cell 20, 45

Rodriguez-Boulan, E. and Sabatini, D. (1978) Proc. Natl. Acad. Sci. USA 75, 5071

Rogers, J., Early, P., Carter, C., Calame, K. Bond, M., Hood, L. and Wall, R. (1980) Cell 20, 303

Rose, J.K. (1980) Cell 19, 415

Rose, J.K. and Bergman, J.E. (1982) Cell 30, 753

Rose, J.K. and Gallione, C.J. (1981) J. Virol. 39, 519

Rose, J.K. and Iverson, L.E. (1979) J. Virol. 32, 404

Rose, J.K. and Knipe, D. (1975) J. Virol. 15, 994

Rose, J.K. and Shafferman, A. (1981) Proc, Natl. Acad. Sci. USA 78, 6670

Rose, J.K., Trachsel, H., Leong, K. and Baltimore, D. (1978) Proc. Natl. Acad. Sci. USA 75, 2732

Ross, A.H., Radhakrishan, R., Robson, R.J. and Khorana, H.G. (1982) J. Biol. Chem. 257, 4152

Roth, M.G., Fitzpatrick, J.P. and Compans, R.W. (1979) Proc. Natl. Acad. Sci. USA 76, 6430

Rothman, J.E. (1981) Science 213, 1212

Rothman, J.E. and Fine, R.E. (1980) Proc. Natl. Acad. Sci. USA 77, 780

Rothman, J.E. and Fries, E. (1981) J. Cell. Biol. 89, 162

Rothman, J. E., Katz, F. and Lodish, H.F. (1978) Cell 15, 1447

Rothman, J.E. and Lodish, H.F. (1977) Nature 269, 775

Rothman, R.E., Pettegrew, H.B. and Fine, R.E. (1980) J. Cell Biol. 86, 182

Rothstein, A., Grinstein, S., Ship, S. and Knauf, P.A. (1977) Trends in Biochem. Sci. 2.

Rowlands, D.J. (1979) Proc. Natl. Acad. Sci. USA 76, 4793

Roy, P. and Bishop, D.H.L. (1972) J. Virol. 9, 946

Sabatini, D.D., Kreibich, G., Morimoto, T. and Adesnik, M. (1982) J. Cell Biol. 92, 1

Scheid, A. and Choppin, P.W. (1973) J. Virol. 11, 283

Schlegel, R., Dickson, R.B., Willingham, M.C. and Pastan, I.H. (1982b) Proc. Natl. Acad. Sci. USA 79, 2291

Schlegel, R., Tralka, T.S., Willingham, M.C. and Pastan, I. (1983) Cell 32, 639

Schlegel, R., Willingham, M. and Pastan, I. (1981) Biochem. Biophys. Res. Commun, 102, 992

Schlegel, R., Willingham, M.D. and Pastan, I.H. (1982a) J. Virol, 43, 871

Schlesinger, M.J. (1981) Ann. Rev. Biochem. 50, 193

Schlesinger, M.J., Magee, A.I. and Schmidt, M.F.G. (1980) J. Biol. Chem. 255, 10021

Schlesinger, M.J., Magee, A.I. and Schmidt, M.F.G. (1981) In Replication of Negative Strand Viruses, Bishop, D.H.L. Compans, R.W. eds. p. 673. Elsivier/North Holland: New York

Schlesinger, M.J. and Malfer, C. (1982) J. Biol. Chem, 251, 9887

Schloemer, R.H. and Wagner, R.R. (1974) J. Virol. 14, 270

Schloemer, R.H. and Wagner, R.R. (1975a) J. Virol. 16, 237

Schloemer, R.H. and Wagner, R.R. (1975b) J. Virol. 15, 882

Schmidt, M.F.G. (1982a) Virology, 116, 327

Schmidt, M.F.G. (1982b) Trends in Biochem. Sci. 7, 322

Schmidt, M.F. and Schlesinger, M.J. (1979) Cell 4, 813

Schnitlein, W.H. O'Banion, M.K., Poirot, M.K. and Reichman, M.E. (1983) J. Virol. 45, 206

Schnitzer, T.J., Dickson, C. and Weiss, R.A. (1979) J. Virol. 29, 185

Schnitzer, T.J. and Lodish, H.F. (1979) J. Virol, 29, 443

Schnitzer, T.J., Weiss, R.A. and Zavada, J. (1977) J. Virol. 23, 449

Schubert, M. Keene, J.D., Lazzarini, R.A. and Emerson, S.U. (1978) Cell 15, 103

Sege, K., Rask, L. and Peterson, R. (1981) Biochem. 20, 4523

Segrest, J.P., Jackson, R.L., Marchesi, V.T., Guyer, R.B. and Terry, W. (1971) Biochem. Biophys. Res. Commun, 49, 964

Sefton, B.M. (1976) J. Virol. 17, 85

Sefton, B.M. and Gaffney, B.J. (1979) Biochem. 18, 436

Sefton, B.M. and Keegstra, K. (1974) J. Virol. 14, 522

Sefton, B.M., Trowbridge, I.S., Cooper, J.A. and Scolnick, E.M. (1982) Cell 31, 465

Sharon, N. and Lis, H. (1982) Glycoproteins. In The Proteins, H. Neurath, R.H. Hill eds., Vol V, p. 1. Academic Press: New York

Shaw, J.M., Moore, N.F., Patzer, E.J., Correa-Freire, M.C., Wagner, R.R. and Thompson, T.E. (1979) Biochem. 18, 538

Silver, J. and Hood, L.E. (1974) Anal. Biochem. 60, 285

Silver, M., McFadden, G., Wilton, S. and Dales, S. (1979) Proc. Natl. Acad. Sci. USA 76, 4122

Simons, K. and Garoff, H. (1980) J. Gen. Virol. 50, 1

Simonsen, C.C., Batt-Humphries, S. and Summers, D.F. (1979a) J. Virol. 31, 124

Simonsen, C.C., Hill, V.M. and Summers, D.F. (1979b) J. Virol. 31, 494

Simpson, R.W., Hauser, R.E. and Dales, S. (1969) Virology 37, 285

Skehel, J.J. and Waterfield, M.D. (1975) Proc. Natl. Acad. Sci. USA 72, 93

Smith, W.P., Tai, P.C. and Davis, B.D. (1981) Proc. Natl. Acad. Sci. USA 78, 3501

Smithies, O., Gibson, D. Fanning, E.M., Godfiesh, R.M., Gilman, J.C. and Ballatyne, D.L. (1971) Biochem. 10, 4912

Somack, R. (1980) Anal. Biochem. 104, 464

Soule, H.R. and Butel, J.S. (1979) J. Virol. 30, 523

Sprague, J., Chondra, J.H., Arnheiter, H., and Lazzarini, R.A. (1983) J. Virol, 45, 773

Stamminger, G. and Lazzarini, R.A. (1974) Cell 3, 85

Stamminger, G. and Lazzarini, R.A. (1977) Virology 78, 202

Stampfer, M. and Baltimore, D. (1973) J. Virol. 11, 520

Stampfer, M., Baltimore, D. and Huang, A.S. (1969) J. Virol. 4, 154

Stanners, C.P., Francoeur, A.M. and Lam. T. (1977) Cell 11, 273

Stoffel, W., Anderson, R. and Stahl, J. (1975) Hoppe-Seylers Z. Physiol. Chem. 356, 1123

Stoffel, W., Schrieber, C. and Scheefers, H. (1978) Hoppe-Seylers Z. Physiol. Chem. 359, 923

Sugita, K., Arita, H., Kawanami, J. and Sato, K. (1979) J. Gen. Virol. 45, 249

Sveda, M.M., Markoff, L.J. and Lai, C.J. (1982) Cell 30, 649

Szil'agyi, J.F., Pringle, C.R. and Macpherson, T.M. (1977) J. Virol. 22, 381

Tabas, I. and Kornfeld, S. (1978) J. Biol. Chem, 253, 7779

Tabas, I., Schlesinger, S. and Kornfeld, S. (1978) J. Biol. Chem. 253, 716

Tannenbaum J., Goorha, R. and Granoff, A. (1979) Virology 95, 227

Tarantino, A.L. and Maley, F. (1977) Anal. Biochem. 77, 185

Tarr, G.E. (1981) Anal. Biochem. 111, 27

Tarr, G.E., Beecher, J.R., Bell, M. and McKean, D.J. (1978) Anal. Biochem. 84, 622

Tato, F., Beamand, J.A. and Wyke, J.A. (1978) Virology 88, 71

Taube, S.E. and Braun, P.C. (1982) J. Gen. Virol. 59, 319

Taube, S.E. and Rothfield, L.I. (1978) J. Virol. 26, 730

Testa, D., Chanda, P.K. and Banerjee, A.K. (1980) Cell 21, 267

Thimmig, R.L., Hughes, J.V., Kinders, R.J., Milenkovic, A.G. and Johnson, T.C. (1980), J. Gen. Virol. 50, 279

Thiry, L. Cogniaux-LeClerc, J., Content, J. and Tack, L. (1978) Virology 87, 384

Thomas, J.R. and Wagner, R.R. (1982) J. Virol. 44, 189

Tomita, M. and Marchesi, V. (1975) Proc. Natl. Acad. Sci. USA 72, 2964

Toneguzzo, F. (1977) Ph.D. Thesis, McMaster University

Toneguzzo, F. and Ghosh, H.P. (1975) FEBS Lett. 50, 369

Toneguzzo, F. and Ghosh, H.P. (1976) J. Virol. 17, 477

Toneguzzo, F. and Ghosh, H.P. (1977) Proc. Natl. Acad. Sci. USA 74, 1516

Toneguzzo, F. and Ghosh, H.P. (1978) Proc. Natl. Acad. Sci. USA 75, 715

Towbin, T., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350

Vassalli, P., Tedghi, R., Lisowska-Bernstein, B., Tartakoff, A. and Jaton, J.C. (1979) Proc. Natl. Aca. Sci. 76, 5515

Verhoeyen, M., Fang, R. Jou, W.M., Devos, R., Nuylebroeck, D., Samsan, E. and Fries, W. (1980) Nature 280, 771

Villa-Komaroff, L., McDowell, M., Baltimore, D. and Lodish, H.F. (1975) In Meth. In Enzymology. K. Moldave, L. Grossman eds., vol. XXX, p. 709, Academic Press: New York

Villarreal, L.P., Breindle, M. and Holland, J.J. (1976) Biochem. 15, 1663

von Heijne, G. (1981) Biochem, Soc. Symp. 46, 259

Wagner, R.R. (1975) In Comprehensive Virology, Frannkel-Conrat, H. and Wagner, R.R. eds., vol. 4, p. 1, Plenum Press: New York

Wagner, R.R., Prevec, L., Brown, F., Summers. D.T., Sokol, F. and MacLeod, R. (1972) J. Virol. 10, 1288

Wagner, R.R., Schnaitman, T.C., Snyder, R.M. and Schnaitman, C.A. (1969) J. Virol. 3, 611

Wagner, R.R., Snyder, R.M. and Yamazaki, S. (1970) J. Virol. 5, 548

Walsh, K.A., Ericsson, L.H., Parmelee, D.C. and Titani, K. (1981) Ann. Rev. Biochem. 50, 261

Walter, P., Jackson, R.C., Marcus, M.M., Lingappa, V.R. and Blobel, G. (1979) Proc. Natl. Acad. Sci. USA 76, 1795

Warren, G.B. and Dobberstein, B. (1978) Nature 273, 569

Weber, K. and Osborn, M. (1969) J. Biol. Chem. 244, 4406

Weck, P.K., Carroll, A.R., Shattuck, D.M. and Wagner, R.R. (1979) J. Virol. 30, 746

Wenland, J., Willingham, M.C., Gallo, M.G. and Pastan, I. (1982) Cell 28, 831

Weinstein, B.B., Marsh, J.B., Glick, M.C. and Warren, L. (1969) J. Biol. Chem. 244, 4103

Wertz, G.W. and Levine, M. (1973) J. Virol. 12, 253

Wickner, W. (1979) Ann. Rev. Biochem. 48, 23

Wickner, W. (1980) Science 210, 861

Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G. and Chasin, L. (1979) Proc. Natl. Acad. Sci. USA 76, 1373

Wiktor, J., Gyorgy, I., Schlumberger, H.D., Sokol, F. and Koprowski, H. (1973) J. Immunol. 110, 269

Wilson, H. and Lenard, J. (1981) Biochem. 20, 1349

Wilson, I.A., Skehel, J.J. and Wiley, D.C. (1981) Nature 289, 366

Winter, G. and Fields, S. (1980) Nuc. Acid Res. 8, 1965

Winter, G., Fields, S. and Brownlee, G.G. (1981) Nature 292, 72

Witt, D.J., Naeve, C.W. and Summers, D.T. (1981) J. Gen. Virol. 56, 383

Witte, O.N. and Baltimore, D. (1977) Cell 11, 505

Wittman-Liebold, B. (1980) In Polypeptides and Hormones, R.F. Beers, Jr., and E.G. Bassess, ed., p. 87, Raven: New York

Wu, F.S. and Lucas-Lenard, J. (1980) Biochemistry 19, 804

Yuan, P.M., Pande, H., Clark, B.R. and Shively, J.E. (1982) Anal. Biochem. 120, 289

Zakowski, J.J., Petri, W.A. and Wagner, R.R. (1981) Biochem. 20, 3902

Zakowski, J.J. and Wagner, R.R. (1980) J. Virol. 36, 93

Zilberstein, A., Snider, M.D., Porter, M. and Lodish, H.F. (1980) Cell 21, 417

Zimmerman, C.L., Appela, E. and Pisano, J.J. (1977) Anal. Biochem, 77, 569