

**An Investigation of the Complexes Formed between the
Hepatitis C Virus E1 and E2 Glycoproteins**

by

Janisha Patel

A Thesis Presented for the Degree of Doctor of Philosophy

in

The Faculty of Science at the University of Glasgow

**MRC Virology Unit
IBLS Division of Virology
Church Street
Glasgow G11 5JR**

December 1999

ProQuest Number: 13818645

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 13818645

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

GLASGOW
UNIVERSITY
LIBRARY

11819

(copy 1)

Contents

Acknowledgements	1
Summary	2
Abbreviations	5
One and Three Letter Amino Acid Abbreviations	8
Chapter 1 - Introduction	9
1.1. Clinical Features of HCV	9
1.1.1. Brief Historical Background	9
1.1.2. Discovery of HCV	9
1.1.3. Classification of HCV	10
1.1.4. Virus Morphology	11
1.1.5. Other Hepatitis Viruses	12
<i>1.1.5.1. HAV</i>	<i>12</i>
<i>1.1.5.2. HBV</i>	<i>13</i>
<i>1.1.5.3. HDV</i>	<i>13</i>
<i>1.1.5.4. HEV</i>	<i>14</i>
<i>1.1.5.5. HGV</i>	<i>14</i>
1.1.6. Epidemiology and Transmission of HCV	15
1.1.7. Clinical Manifestation and Natural History of HCV Infection	16
<i>1.1.7.1. Acute HCV</i>	<i>16</i>
<i>1.1.7.2. Chronic HCV</i>	<i>16</i>
<i>1.1.7.3. Chronic Liver Disease</i>	<i>17</i>
<i>1.1.7.4. Extra-Hepatic HCV Associated Diseases</i>	<i>17</i>

1.1.8. Immune Response	18
1.1.9. Diagnosis and Histopathology	19
1.1.10. Detection of HCV RNA and Antigen in the Liver	20
1.1.11. Prevention and Treatment	20
1.2. Molecular Features of HCV	23
1.2.1. Comparison of the HCV, Flavi- and Pestivirus Genomes	23
1.2.2. 5' Untranslated Region (5' UTR) and Translation of HCV RNA	25
1.2.3. 3' UTR	27
1.2.4. HCV Virus-Encoded Non-Structural Proteins	27
1.2.5. NS2	27
1.2.6. NS3	28
<i>1.2.6.1. NS3 Proteinase</i>	<i>29</i>
<i>1.2.6.2. NS3 NTPase/Helicase</i>	<i>31</i>
1.2.7. NS4A and NS4B	32
1.2.8. NS5A	32
1.2.9. NS5B	33
1.2.10. HCV RNA Replication	34
1.3. HCV Structural Proteins	36
1.3.1. Core	36
1.3.2. p7	38
1.3.3. E1 and E2 Glycoproteins	38
1.3.5. HCV Glycoprotein Complexes	40
1.3.6. Properties of Native E1E2 Complexes	41

1.3.7. Properties of Aggregated E1E2 Complexes	43
1.3.8. Requirements for E1E2 Complex Formation	43
1.3.9. Subcellular Localisation of HCV Glycoproteins	44
1.3.10. Association of HCV Glycoproteins with Viral Proteins	46
1.3.11. Association of HCV Glycoproteins with Folding Proteins	46
1.3.11.1. The Role of E2 in the Folding of E1	46
1.3.11.2. The Role of Cellular Folding Proteins	47
1.3.12. Interactions of HCV Glycoproteins with Other Cellular Proteins and Process	48
1.3.12.1. Effects on PKR	48
1.3.12.2 E2 Binding to Lactoferrin	48
1.3.12.3. Activation of GPR Protein Synthesis	49
1.3.13. Cellular Receptors	49
1.3.14. Assembly and Release of HCV Particles	50
1.4. Properties of the Endoplasmic Reticulum	51
1.4.1. Signal Peptide Recognition and Translocation	51
1.4.2. The Luminal Environment of the ER	52
1.4.3. Isomerases Present in the ER	54
1.4.4. Chaperones Present in the ER	54
1.4.5. The Role of Covalent Modifications in Protein Folding	56
1.4.6. Oligomerisation of Proteins in the ER	58
1.4.7. Transport from the ER	59
1.4.8. Maintenance of the ER through Retrieval/Retention of Resident Proteins	60

1.4.9. ER Stress Response and Protein Degradation	61
1.5. Viral Glycoproteins of the <i>Flaviviridae</i>	64
1.5.1. Glycoproteins of Pestiviruses	64
1.5.2. Glycoproteins of Flaviviruses	65
1.6. The Semliki Forest Virus Expression System	67
1.7. Aims of the Project	69
Chapter 2 – Materials and Methods	70
2.1.1. Bacterial Strains	70
2.1.2. Vectors	70
2.1.3. Kits and Enzymes for DNA/Protein Modification	70
2.1.4. Cells	70
2.1.5. Cell culture growth media	71
2.1.6. Radiochemicals	71
2.1.7. Antibodies	71
2.1.8. Chemicals	72
2.1.9. cDNA clones	72
2.1.10. Photographic reagents	73
2.1.11. Solutions	73
2.2.1. Manipulation of DNA	75
2.2.1.1. <i>Small Scale Purification of DNA (mini-preps)</i>	75
2.2.1.2. <i>Large Scale Purification of DNA (maxi-preps)</i>	75
2.2.3. Oligonucleotide Synthesis and Purification	76
2.2.4. Quantitation of Plasmid DNA and Oligonucleotides	76

2.2.5. Restriction Enzyme Digestion of DNA	77
2.2.6. Electrophoretic Separation and Isolation of Digested DNA	77
2.2.6.1. <i>Agarose Gel Electrophoresis</i>	77
2.2.6.2. <i>Purification of DNA from Agarose Gels</i>	77
2.2.6.3. <i>Purification of Linearised DNA and PCR Fragments</i>	78
2.2.6.4. <i>Phenol/Chloroform Extraction</i>	78
2.2.6.5. <i>Ethanol Precipitation</i>	78
2.2.7. Ligation Reactions	78
2.2.8. Preparation of Competent <i>E coli</i> Cells	79
2.2.8.1. <i>Electrocompetent DH5α</i>	79
2.2.8.2. <i>Competent GM48</i>	79
2.2.9. Transformation of Competent <i>E coli</i> Cells	79
2.2.9.1. <i>Electroporation of DH5α and TOP10F' Cells</i>	79
2.2.9.2. <i>Transformation of GM48 Bacteria</i>	80
2.2.10. Automated DNA Sequencing	80
2.2.11. PCR Amplification of DNA	80
2.2.12. <i>In Vitro</i> Transcription of Linearised DNA	81
2.2.13. Maintenance of BHK Cells	81
2.2.14. Electroporation of Recombinant SFV RNA	81
2.2.15. Trypan Blue Staining	82
2.2.16. Metabolic Labelling of Electroporated Cells	82
2.2.16.1. <i>Radiolabelling with ³⁵S-amino acids</i>	82
2.2.16.2. <i>³H-Mannose Labelling</i>	83
2.2.17. Preparation of Cell Extracts and Isolation of Protein from Eukaryotic Cells	83
2.2.17.1. <i>Immobilised-Metal Affinity Chromatography (IMAC)</i>	83

2.2.17.2. <i>Immunoprecipitation</i>	84
2.2.18. TCA Precipitation	84
2.2.19. Deglycosylation of Proteins Using Endo H_f and PNGase F	84
2.2.20. Coupled <i>In Vitro</i> Transcription and Translation (<i>In Vitro</i> TNT)	84
2.2.21. Protein Analysis by SDS-PAGE	85
2.2.22. Western Blot Analysis	86
2.2.22.1. <i>Electroblotting to Nitrocellulose</i>	86
2.2.22.2. <i>Immunodetection</i>	86
2.2.23. Removal of Antibodies and Immunoglobulins from Membranes	87
2.2.24. Immunofluorescence	87
2.2.25. Preparation of PFA Fixative	88
2.2.26. Computer Software	88
Chapter 3 – Establishment of the SFV Expression System	89
3.1. Introduction	89
3.2. Construction of Histidine-Tagged E1 (E1_{his})	90
3.3. <i>In Vitro</i> Transcription of SFV RNA	90
3.4. Effect of Electroporation on Survival of BHK Cells and Efficiency of Electroporation of SFV RNA	91
3.5. Solubilisation of HCV Glycoprotein E1_{his}	92
3.6. Radiolabelling of HCV Glycoproteins and Immunodetection of E2	93
3.7. Isolation of E1_{his} and E1E2 Complexes by IMAC	95
3.8. Optimisation of Ni-NTA Binding of E1_{his}	96
3.9. Immunolocalisation and ER Retention of E1_{his} and E2 in BHK Cells	97
3.10. Discussion	98

Chapter 4 - Characterisation of the Native and Aggregated Complexes Formed by E1 and E2	104
4.1. Introduction	104
4.2. Behaviour of Strain Glasgow E1_{his} and E2 under Non-Reducing Electrophoresis Conditions	104
4.3. Behaviour of E1E2 Complexes under Reducing Conditions <i>In Situ</i>	105
4.4. Effect of Removal of DTT from Cells on E1E2 Complexes	107
4.5. Disulphide Bond Formation is not Required for the Initial Interactions Between E1 and E2	108
4.6. Comparison of Glycoproteins Produced from Strains Glasgow and H77	109
4.7. Behaviour of E1 and E2 Produced by Strain H77 in the Presence and Absence of DTT <i>In Situ</i>	112
4.8. Discussion	114
Chapter 5 – Characteristics of E1 and Their Relationship to Complex Formation with E2	119
5.1. Introduction	119
5.2.1. Construction of pSFV Recombinant Vectors Expressing Glasgow and H77 E1 in the Absence of E2	119
5.2.2. Behaviour of E1 on Expression in the Absence of E2	120
5.3.1. Construction of E1 Internal Deletion Mutants	123
5.3.2. Examination of the Properties of E1 Deletion Mutants	124
5.3.3. Analysis of Complex Formation between E2 and E1 Deletion Mutants	126
5.3.4. Further Characterisation of Deletion Mutants E1_{Δ266-303} and E1_{Δ309-338}	128
5.4.1. Construction of Cysteine Substitution Mutants	130
5.4.2. Analysis of Cysteine Substitution Mutants	131
5.4.3. Analysis of Partial Cysteine Substitution Mutants	133

5.3.5. Discussion	134
Chapter 6 – Identification of a Region in E2 Required for the Folding of E1	140
6.1.1. Introduction	140
6.1.2. Construction of Internal Deletion Mutants in Strain Glasgow E2	140
6.1.3. Properties of the E2 Deletion Mutants	142
6.1.4. Complex Formation between Wild-type E1 and E2 Deletion Mutants	143
6.2.1. Construction of Plasmids Expressing Strain H77 Truncated E2 and E2 Alone	145
6.2.2. Characterisation of the Strain Glasgow and H77 E2 Truncated Mutants	145
6.3.1. Use of Foreign Glycoprotein Sequences to Study the Properties of E2	148
6.3.2.1. Construction Of Chimeric Proteins - Cloning Of Full Length HSV-1 Glycoprotein D	148
6.3.2.2. Construction of E2 Chimeric Proteins	149
6.3.2.3. Construction of gD Chimeric Proteins	150
6.3.3.1. Cellular Distribution of Chimeric Proteins - Indirect Immunofluorescence	151
6.3.3.2. Endo H _f and PNGase F Sensitivity of Chimeric Proteins	153
6.3.4. Effect of Foreign Ecto- and Transmembrane Domains on E1E2 Complex Formation	155
6.4. Discussion	159
Chapter 7 – General Discussion	163
7.1. The Arrangement of E1 and E2 on the HCV Polyprotein	163
7.2. Formation of E1E2 Aggregates – An Intrinsic Property of the Glycoproteins or an <i>In Vitro</i> Artefact?	164
7.3. Retention of E1 and E2 Glycoproteins within the ER	168

7.4. Role of Covalent Modifications on E1 Folding	170
7.5. Model for the Role of E1 and E2 in HCV Particle Assembly	171
Chapter 8 - References	173
Appendix 1 – Strain Glasgow Core, E1 and E2 Nucleic Acid and Amino Acid Sequence	207
Appendix 2 - Amino Acid Sequence Alignment of Strain Glasgow and H77 Core, E1 and E2	211
Appendix 3 – Strain H77 Core, E1, and E2 Nucleic Acid and Amino Acid Sequence	213
Appendix 4 - HSV-1 Glycoprotein D Nucleic Acid and Amino Acid Sequence	217
Appendix 5 – Characteristics of E1 and E2 Specific Antibodies Generated by Dr A Patel	219
Appendix 6 – Restriction Enzymes Employed to Generate DNA Clones	220

Acknowledgements

I thank Professor Duncan McGeoch for providing the opportunity to conduct this study within the Division of Virology.

I express my deepest gratitude to my supervisor Dr John McLauchlan for his inexhaustible patience, encouragement and helpful advice throughout the project. I am also grateful to him for critical reading of this manuscript.

I wish to thank Dr Arvind H Patel for kindly allowing the use of immunological reagents and for helpful discussions during the course of the project.

I thank Dr Sarah Butcher for helpful advice during the early stages of the project and Dr Liz McCrudden for useful discussions during the writing of the thesis.

Invaluable discussions with friends in the laboratory Jonny, Ania, Graham, Reg, Martin, and Annette, as well as with others in the Division were much appreciated.

My thanks to staff within the Division who have eased the development of this project, in particular Lesley Taylor, for the automated sequencing service, David McNab and Alex Orr, for synthesis of oligonucleotides, and the wash room and media staff.

I thank my friends outside the Division for their support. In particular, thanks to my flatmates, past and present, for making flat 69D a home.

I am grateful to my family, my mother, Dipen my brother and especially my father for proof reading this manuscript.

Finally, this work is dedicated to Kailesh Patel who has encouraged and guided me from an early age.

Summary

Hepatitis C virus (HCV) encodes two glycoproteins, termed E1 and E2, which are presumed to constitute the proteinaceous components of the virion envelope. From a number of studies, E1 and E2 associate to form a complex however, the nature of the types of interaction between the proteins is the subject of controversy. Two types of complex, termed the native and aggregated forms, have been identified. In native complexes, E1 and E2 associate by non-covalent interactions while disulphide bonds form intermolecular links between the glycoproteins in aggregated complexes.

The aim of my project was to characterise the complex formed by E1 and E2 from a local genotype 1a strain, called strain Glasgow. Studies were later expanded to include analysis of the glycoproteins from strain H77, another genotype 1a strain that is infectious in chimpanzees. Due to the lack of efficient virus replication in tissue culture, the Semliki Forest virus vector was used for expression of the glycoproteins. This involved the introduction of *in vitro* transcribed RNA molecules into mammalian cells by electroporation and conditions to optimise the system were ascertained.

In the absence of sufficient immunological reagents during the early phases of the project, a histidine tag was inserted at the N-terminus of E1 from strain Glasgow (E1_{his}) to provide a means for purification by metal chelate chromatography. This method co-eluted E2 with E1_{his} which was later verified by the ability of polyclonal sera against E2 to co-precipitate E1 with E2. Thus, a heteromeric complex formed between E1 and E2 from strain Glasgow, a finding which is in agreement with studies on the glycoproteins from other strains. Using the available reagents for identifying E1 and E2 and by analysis under reducing and non-reducing electrophoretic conditions, most of the complex was in the form of disulphide-linked aggregates with little detectable native complex.

The requirement for disulphide bonds to enable E1 and E2 association was examined by treatment of cells synthesising the proteins with the reducing agent dithiothreitol (DTT). Addition of DTT to cells did not prevent cleavage and processing of the glycoproteins. Significantly, the intermolecular disulphide links in aggregates were disrupted by DTT but E1 and E2 continued to associate, indicating that both covalent and non-covalent intermolecular interactions occur in the aggregated complex. E1 and E2, synthesised in the

presence of DTT, also formed a complex and therefore disulphide bond formation is not a prerequisite for interactions between the proteins.

Comparisons between strain Glasgow and strain H77 showed that native E1E2 complexes were more readily detected from strain H77. It is proposed that this may be a consequence of alterations in the number of glycosylation sites in E2 between the strains as well as changes in amino acid sequences in E1 and E2. Moreover, E1 from strain Glasgow had reduced mobility as compared to strain H77 E1 protein which could not be accounted by changes in the relative sizes of the nascent proteins or glycosylation patterns. Two deletion mutants in E1 from both strains also displayed higher apparent molecular weights than predicted. The behaviour of these mutant proteins may result from the deletion of hydrophobic segments in E1. However, the apparent molecular weight difference between E1 from strain Glasgow and strain H77 could not be explained.

In an attempt to identify regions of E1 and E2 in strain Glasgow that may account for the high level of aggregation by these proteins, deletion mutants were created in E1 and E2; for comparative purposes, an identical set of deletion mutants were created also for E1 from strain H77. Immunoprecipitation studies revealed that multiple regions were involved in interactions between E1 and E2. No single deletion was able to completely abolish E1E2 interactions although native complex formation in strain H77 was decreased on removal of sequences from E1. Further analysis of complexes and the role of disulphide bond formation was undertaken by constructing a series of mutants in which the cysteine residues in E1 were replaced with either serine or alanine. With each of these mutants, E1 and E2 continued to associate but the amounts of native, non-aggregated E2 were reduced to undetectable levels. Hence, native complexes of E1 and E2 were highly sensitive to changes in either protein such that alteration even at a single cysteine residue was sufficient to reduce the abundance of native complex.

Examination of E1 expressed in the presence and absence of E2 provided a set of criteria for assessing the folding status of the protein. In the absence of E2, E1 remained in the reduced state and formed homo-oligomeric molecules that could be identified under reducing electrophoretic conditions. The formation and stability of these homo-oligomers were not entirely dependent on disulphide bond formation. Moreover, there was an increase in the quantities of glycosylated forms of E1 in the absence of E2. Using these criteria, the region in E2 that is necessary for folding of E1 was analysed by constructing

chimeric proteins between E2 and segments of foreign glycoproteins. It was found that a peptide sequence of 58 residues, which spans the cleavage site between E2 and p7 and thus contains the transmembrane domain of E2, was sufficient to permit E1 oxidation and prevent oligomerisation. By contrast, the ectodomain of E2 was unable promote E1 folding. This suggests that the transmembrane domain of E2 plays a key role in E1 folding and hence in native complex formation.

Abbreviations

µg	microgram(s)
µl	microlitre(s)
β-ME	β-mercaptoethanol
³⁵ S	sulphur isotope (35)
A	adenine
aa	amino acid
AIDS	acquired immunodeficiency syndrome
ALT	alanine transaminase
AMP	ampicillin
Anti-E1	antibodies to E1
Anti-E2	antibodies to E2
Anti-His	antibody to histidine tag
APS	ammonium persulphate
ATP	adenosine 5'-triphosphate
BFA	brefeldin A
BHK	baby hamster kidney cells
bp	base pair(s)
BPB	bromophenol blue
BSA	bovine serum albumin
BVDV	bovine viral diarrhoea virus
C	cytosine
C-	carboxy-terminus
cDNA	complementary DNA
CHX	cycloheximide
Ci	Curie(s)
COP	coat protein complex
CPG	controlled pore glass
CSFV	classical swine fever virus
CTL	cytotoxic T-lymphocyte
CTP	cytidine 5'-triphosphate
DEN	dengue (virus)
dH ₂ O	deionised molecular biology grade water
DMT	dimethoxytrityl
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dTTP	deoxythymidine 5'-triphosphate
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
E1	envelope glycoprotein 1
E2	envelope glycoprotein 2
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetra-acetic acid (disodium salt)
ELISA	enzyme-linked immunosorbent assay(s)
Endo H _f	endo-β-N-acetylglucosaminidase H _f
ER	endoplasmic reticulum
ERGIC	ER-Golgi intermediate compartment
EtBr	ethidium bromide
FITC	fluorescein isothiocyanate
g	gravitational force

G	guanine
gD	HSV-1 glycoprotein D
Glc	glucose
GlcNAc	N-acetylglucosamine
GMEM	Glasgow minimal eagles medium
gp	glycoprotein
GPI	glycosylphosphatidylinositol
gr	gram(s)
GSSG	glutathione disulphide
GT	UDP-Glc:glycoprotein glucosyltransferase
GTP	guanosine 5'-triphosphate
h	hour(s)
HA	haemagglutinin
HAV	hepatitis A virus
HBV	hepatitis B virus
HCC	hepatocellular carcinoma
HCl	hydrochloric acid
HCMV	human cytomegalovirus
HCV	hepatitis C virus
HDV	hepatitis D virus
HEV	hepatitis E virus
HGV	hepatitis G virus
HIV	human immunodeficiency virus
HRP	horseradish peroxidase
HSV-1	herpes simplex virus type 1
HVR	hyper variable region
IFN- α	alpha interferon
Ig	immunoglobulin
IMAC	immobilised-metal affinity chromatography
IP	immunoprecipitation
IRES	internal ribosome entry site
ISDR	interferon-sensitivity determining region
ISH	<i>in situ</i> hybridisation
JEV	Japanese encephalitis virus
kb	kilobase(s)
kDa	kilodalton(s)
l	litre(s)
M	Molarity
mA	milliamps
MAB	monoclonal antibody
man	mannose
mg	milligrams
min	minute(s)
ml	millilitre(s)
mM	millimolar
mmol	millimole(s)
mRNA	messenger RNA
MW	molecular weight
N-	amino-terminus
NANB	non-A, non-B (hepatitis)
NBCS	new born calf serum
NEM	N-ethylmaleimide

Ni	nickel
nm	nanometer(s)
NS	non-structural
nt	nucleotide
NTA	nitrilo triacetic acid
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PDI	protein disulphide-isomerase
PFA	paraformaldehyde
PKR	RNA-dependent protein kinase
PMSF	phenylmethylsulphonyl fluoride
PNGase F	peptide:N-glycanase F
PTB	polypyrimidine tract-binding protein
RER	rough ER
RIBA	recombinant immunoblot assay
RNA	ribonucleic acid
RNase A	ribonuclease A
RNasin	ribonuclease inhibitor
RTM	room temperature
RT-PCR	reverse transcriptase-PCR
RVE1	rubella virus E1 (glycoprotein)
SDS	sodium dodecyl sulphate
sec	second(s)
SER	smooth ER
SFV	Semliki Forest virus
SGB	stacking gel buffer
SRP	signal recognition particle
T	thymine
TBE	tick-borne encephalitis virus
TCA	trichloroacetic acid
TED	tris(carboxymethyl) ethylene diamine
TEMED	N,N,N',N'-tetramethylethylenediamine
TGN	<i>trans</i> Golgi network
TM (tm)	transmembrane (domain)
Tris	tris (hydroxymethyl) aminomethane
tRNA	transfer RNA
Tween 20	polyoxyethylene sorbitan monolaurate
U	uracil
UDP	uridine 5'-diphosphate
UTP	uridine 5'-triphosphate
UTR	untranslated region
UV	ultraviolet
VSV G	vesicular stomatitis virus G (protein)
YFV	yellow fever virus

One and Three Letter Amino Acid Abbreviations

Amino acid	Three letter code	One letter code	Codons
Alanine	Ala	A	GCU GCC GCA GCG
Arginine	Arg	R	AGA AGG CGU CGC CGA CGG
Asparagine	Asn	N	AAU AAC
Aspartic acid	Asp	D	GAU GAC
Cysteine	Cys	C	UGU UGC
Glutamine	Gln	Q	CAA CAG
Glutamic acid	Glu	E	GAA GAG
Glycine	Gly	G	GGU GGC GGA GGG
Histidine	His	H	CAU CAC
Isoleucine	Ile	I	AUU AUC AUA
Leucine	Leu	L	UUA UUG CUU CUC CUA CUG
Lysine	Lys	K	AAA AAG
Methionine	Met	M	AUG
Phenylalanine	Phe	F	UUU UUC
Proline	Pro	P	CCU CCC CCA CCG
Serine	Ser	S	AUG AGC UCU UCC UCA UCG
Threonine	Thr	T	ACU ACC ACA ACG
Tryptophan	Trp	W	UGG
Tyrosine	Try	Y	UAU UAC
Valine	Val	V	GUU GUC GUA GUG

Chapter 1 - Introduction

1.1. Clinical Features of HCV

1.1.1. Brief Historical Background

During the 1970s and 1980s high incidences of transfusion-associated hepatitis were diagnosed that could not be ascribed to any known viral agent for which serological tests had been developed at that time. The exclusion of hepatitis A or B virus infections led to the assignment of the unknown agent as non-A, non-B (NANB) hepatitis (Prince *et al.*, 1974). In 1989, modern techniques of molecular cloning aided the discovery of a novel RNA virus that accompanied NANB hepatitis and the agent was termed hepatitis C virus (HCV; Choo *et al.*, 1989). HCV has now been shown to be the cause of most cases of NANB hepatitis. In addition, a much larger number of infected individuals have been observed without overt liver disease (DiBisceglie, 1998).

1.1.2. Discovery of HCV

Serological tests developed during the 1970s for infection by hepatitis A virus (HAV) and hepatitis B virus (HBV) effectively eliminated the role of these viruses in causing transfusion associated hepatitis (Houghton, 1996). A substantial number of hepatitis cases with clinical and biochemical features of viral hepatitis not resulting from HAV or HBV continued to be diagnosed. The term non-A, non-B hepatitis was coined for these cases. The agent associated with this disease was observed initially to cause a mild illness in blood transfusion recipients, but chronic hepatitis often ensued with progression to cirrhosis and liver failure in approximately 20% of cases (Houghton, 1996). Intravenous administration of filtered human sera from patients led to the transmission of NANB hepatitis to chimpanzees. The early studies in chimpanzees confirmed the role of an enveloped virus of approximately 60 nm as the causative agent (Houghton, 1996).

Despite research conducted for over a decade, attempts to identify the infectious agent associated with NANB hepatitis had failed using conventional virological and immunological approaches that had been employed to characterise HAV and HBV. This may have resulted from insufficient concentrations of viral antibodies as well as viral

antigen in NANB hepatitis infections. Several groups focused on the application of recombinant DNA cloning technology as an alternative. Choo *et al.* (1989) devised a technique to increase viral antigen concentrations by employing a cDNA library expressed in bacteria. The cDNA library was derived from chimpanzee plasma containing a relatively high infectious titre [chimpanzee infectious doses per ml (CID ml⁻¹)], using reverse transcriptase or DNA polymerase and random primers directed towards both DNA and RNA. Following introduction into a bacterial expression vector, the library was screened for rare clones expressing viral antigen with serum from a chronic NANB hepatitis patient as a presumed source of viral antibodies. Hybridisation to human and chimpanzee DNA tested the origin of possible positive clones by Southern blot analysis. Such attempts failed to identify any novel nucleotide sequences, which led to the use of total RNA extracted from infected chimpanzee liver. A hybridisation signal was detected specifically to the cloned cDNA. The same signal was not observed when ribonuclease treatment was applied to purified RNA or when RNA derived from uninfected liver was employed. Hence, it appeared that the identified clones were from an exogenous RNA molecule associated with NANB hepatitis infection. Since only one of the strands of the cDNA clones hybridised, it was concluded that the RNA from infectious plasma was single-stranded.

The size of the RNA species homologous to the cDNA clones was elucidated by separating RNA obtained from infectious chimpanzee liver by electrophoresis under denaturing conditions, followed by hybridisation with one of the cDNA clones. The detected RNA fragments were estimated to represent a molecule of at least 10,000 nucleotides. Furthermore, the binding of oligo(dT)-cellulose indicated the presence of either a 3' terminal polyadenylated sequence or an internal A-rich tract. A single large cDNA clone indicated that one continuous, translational open reading frame (ORF) was encoded by the RNA (Choo *et al.*, 1989). The agent was called hepatitis C virus and thereafter various serological assays were developed (Kuo *et al.*, 1989). Such assays have demonstrated that HCV represents the predominant cause of transfusion-associated NANB hepatitis around the world.

1.1.3. Classification of HCV

The successful isolation of HCV cDNA was followed by cloning of almost the entire HCV genome by several groups (Kato *et al.*, 1990; Choo *et al.*, 1991; Takamizawa *et al.*, 1991).

The characteristics of the HCV genome were found to be remarkably similar to viruses belonging to the *Flaviviridae* family. Sequence analysis of the single-stranded RNA genome revealed that the large ORF encoded a polyprotein comparable in size to those of flavi- and pestiviruses. By comparison with these viruses, the HCV polyprotein was assumed to be proteolytically digested to give the individual virus products. Although the overall amino acid sequence homology of the putative HCV proteins with the *Flaviviridae* family of viruses is low, alignment of HCV and *Flaviviridae* genomes provided evidence for the existence of regions of sequence homology. This indicated a similar genetic organisation. For example, regions of the primary sequence of a putative HCV protein were homologous with helicases encoded by animal pestiviruses, plant potyviruses, and human flaviviruses (Miller & Purcell, 1990; Choo *et al.*, 1991). Hydrophobicity profiles of the N-terminal region of the HCV polyprotein were more similar to the *Pestivirus* genus than the genus *Flavivirus* (Choo *et al.*, 1991). However, despite these similarities, pestiviruses and HCV did differ significantly in primary amino acid sequence. As a result, a third genus of the *Flaviviridae* family, the *Hepaciviruses*, was proposed for classification of HCV (Robertson *et al.*, 1998).

cDNA sequence data has revealed the presence of nucleotide sequence variations within the HCV genome (Kato *et al.*, 1989). Based upon the available sequence data, all currently known isolates of HCV have been placed into one of six phylogenetically distinct groups (Fig 1.1). These groups are described as clades 1 to 6. Clades 1, 2, 4, and 5 correspond to genotypes 1, 2, 4, and 5 while clade 3 comprises genotype 3 and subtype 10, and clade 6 comprise genotypes 6, 7, 8, 9, and 11 (Davidson *et al.*, 1995; Robertson *et al.*, 1998). Genotypes 1, 2, and 3 are widely distributed throughout the world, while type 4 is predominant in the Middle East. HCV types 5 and 6 appear to be restricted to South Africa and Hong Kong respectively (Houghton, 1996).

1.1.4. Virus Morphology

HCV is the first virus whose identification was based almost entirely on examination of its genomic sequence. There is little reported characterisation by classical virological methods, mainly due to the inability to isolate substantial quantities of virus from infected individuals. However, indirect immunogold electron microscopy studies have been carried out on plasma samples with particularly high HCV RNA titres using polyclonal and monoclonal antibodies specific to the putative HCV envelope proteins (Kaito *et al.*, 1994).

This revealed spherical virus-like particles, 55 to 65 nm in diameter with spike-like projections, which were present in 1.14 to 1.16 g/ml fractions after sucrose density gradient centrifugation. Monoclonal and polyclonal antibodies to the HCV envelope protein, E2 (Section 1.3.3), specifically identified these virus-like particles. The size and morphological features of the HCV virion particles are similar to those of flaviviruses (Yuasa *et al.*, 1991; Kaito *et al.*, 1994).

1.1.5. Other Hepatitis Viruses

Epidemics of jaundice were recognised as early as the Middle Ages, but the notion that such outbreaks were the result of hepatitis caused by infectious agents were only considered in the present century. To date, five different human hepatitis viruses have been identified and characterised in detail; HAV, HBV, HCV, hepatitis D virus (HDV) and hepatitis E virus (HEV). A sixth potential hepatitis-associated agent, termed hepatitis G virus (HGV), has been discovered recently, but subsequent research has failed to demonstrate a link with hepatitis (Dejean *et al.*, 1999). Although all of these viruses cause a similar systemic disease affecting the liver, they do not all belong to the same virus family. All have an RNA-based genome, with the exception of HBV, which has a DNA genome (Reid & Dienstag, 1997).

1.1.5.1. HAV

HAV is an enterically transmitted virus causing an acute and self-limiting infection (Feinstone & Gust, 1997). The infection is frequently asymptomatic particularly in infants and young children. Chronic infection has not been reported although fulminant hepatitis and death are occasional consequences of infection (more than hepatitis B). The virus is a member of the *Picornaviridae* family, which includes the enteroviruses and rhinoviruses. Due to unique features, HAV has been assigned to the *Hepatavirus* genus, of which it is the only member. It is a non-enveloped virus of ~27-28 nm in diameter, and contains a positive-sense, single-stranded linear RNA genome approaching 7,500 nucleotides in length. The genome encodes a single polyprotein flanked by extensive 5' and 3' untranslated regions, which are involved in viral replication. Unlike cellular messenger RNA molecules that possess a cap, the 5' end of the HAV genome has a small, covalently linked viral protein known as VPg, a feature common to picornaviruses. An internal ribosomal entry site (IRES) is also located within the 5' untranslated region. The

polyprotein may be divided into three domains: P1 (capsid proteins), P2 and P3 (non-structural proteins). Each domain is ultimately cleaved to produce up to four polypeptides (Feinstone & Gust, 1997).

1.1.5.2. HBV

HBV is a blood-borne agent transmitted primarily by sexual intercourse, maternal-child transmissions and percutaneous routes. The virus is sustained through chronic infection in human hosts, rather than environmental reservoirs such as water or food as is the case for HAV and HEV. Chronic HBV infection is a common disease with an estimated global prevalence of over 300 million carriers with viral antigen markers (e.g. HBsAg), approximating to 5% of the world's population (Maynard, 1990). Patients with chronic HBV infection are at risk of developing long-term complications, including cirrhosis and hepatocellular carcinoma (HCC). HBV is the prototype member of the family *Hepadnaviridae*, which also contains other members found in wild animals (Wright *et al.*, 1997). Its virions are double-shelled particles of ~40-42nm with an outer lipoprotein envelope containing multiple related envelope glycoproteins. The encapsidated genome is a relaxed circular, partially duplexed DNA molecule of approximately 3 kb. The limited coding information within the small genome is used with extraordinary efficiency to such an extent that every nucleotide is translated. There are four overlapping ORFs in the genome, which encode envelope proteins, core protein, viral polymerase and a small regulatory protein (Wright *et al.*, 1997). Initiation sites for the translation of the seven known viral gene products occur from alternate AUG codons located within the ORF, thus maximising the number of viral proteins made (Wright *et al.*, 1997).

1.1.5.3. HDV

Disease progression in hepatitis B carriers is markedly accelerated upon superinfection with HDV. This agent is the only prototype of the *Deltaviridae* family that relies on helper functions provided by HBV to establish *in vivo* infections. Such a subviral agent is unique amongst animal viruses so far described. Transmission occurs by the parenteral route. The HDV virion is a chimera assembled with viral genome and antigen enveloped by the surface antigen of HBV. The genome consists of a 1.7 kb closed circular, rod-shaped, negative-sense, single-stranded RNA molecule capable of internal hybridisation, self-cleavage, and self-ligation (Smedile *et al.*, 1997).

1.1.5.4. HEV

In 1955, a large outbreak of epidemic hepatitis in New Delhi, India, was later confirmed to have been caused by HEV and not HAV, since it is also a faecal-orally-transmitted agent. HEV was first cloned by molecular methods in 1990 (Reyes *et al.*, 1990) and was identified as a positive-sense, single-stranded RNA virus distinct from all other known positive-sense RNA viruses of humans. The genome is contained within a non-enveloped icosahedral particle that ranges between ~27-32 nm in diameter, and encodes at least three partially overlapping ORFs. The virus is grouped with the *Caliciviridae*, a family of small, round, structured viruses, including agents associated with non-bacterial gastroenteritis in humans and animals.

1.1.5.5. HGV

During attempts to develop an animal model for hepatitis A, marmosets (*Saguinas spp.*) were administered a series of inocula derived from patients with suspected infectious hepatitis (Shimotohno & Feinstone, 1997). Serum from a surgeon (GB) with NANB hepatitis, was unique in inducing disease in marmosets. It was later demonstrated that GB virus was a distinct agent from HCV, by its failure to establish disease in chimpanzees in a manner akin to HCV. Using a PCR-based cDNA subtraction technology called representational difference analysis (RDA), the genome of this infectious agent was isolated from the marmoset serum pool. The genome had the organisational structure of *Flaviviridae* organisms and seemed to be generally related to HCV and the pestiviruses (Scott *et al.*, 1995). The genome of this agent, GBV-C, encodes a single ORF coding for a polyprotein of ~3 000 amino acids. Although GBV-C co-exists frequently with HCV, its role in affecting the rate of disease progression in HCV-infected individuals remains to be established. At present, specific diagnosis can be made by antibody assay and by amplification of the viral genome by PCR. In the absence of serologic assays, the clinical and epidemiological aspects of the agent are yet to be defined (Dejean *et al.*, 1999).

1.1.6. Epidemiology and Transmission of HCV

Although HCV infection is found throughout the world, there are significant geographic variations in the prevalence of antibodies to HCV in serum. It is estimated that about 170 million persons world-wide may be infected with HCV (Lavanchy *et al.*, 1999). The prevalence is estimated at approximately 1.5% in the developed countries, but relatively higher infection rates are found in parts of Eastern Europe and Africa (Alter, 1995; DiBisceglie, 1998). The highest prevalence rate occurs in Egypt, approaching 15% of the general population (Saeed *et al.*, 1991).

The discovery of HCV was made in the context of blood transfusion and, before the development of appropriate screening tests, this was the major route of HCV transmission in the West. However, the risk of HCV infection from tested donated blood is now estimated to be as low as 1 in 103,000 recipients (Schreiber *et al.*, 1996). HCV was also a common cause of liver disease in individuals exposed to unscreened blood products, such as haemophiliacs receiving blood clotting factors (Alter, 1993; Lavanchy *et al.*, 1999).

Presently, intravenous drug abuse is the predominant cause of transmission in the West (DiBisceglie, 1998). However, there has been a decline in rates of infection since 1989, presumably as a result of safer needle practices (Alter, 1993). Other individuals at increased risk include renal dialysis patients either with or without associated blood transfusion, and organ transplant patients. Spread of HCV may also have occurred through the practice of traditional healing techniques involving puncture of the skin (DiBisceglie, 1998). Health-care employment with occupational exposure to blood and haemodialysis are additional risk factors (Alter *et al.*, 1990).

Household or sexual transmission present lower levels of risk for HCV infection. Various studies indicate that transmission between spouses and other family members may have been due to routes of infection involving inapparent percutaneous exposure (Kao *et al.*, 1993; Rice *et al.*, 1993).

Transmission may happen *in utero* (Weiner *et al.*, 1993), at an estimated rate of about 5% from affected mothers to offspring (Ohto *et al.*, 1994). Transmission is likely to be related to the level of viremia in the mother because women transmitting HCV had significantly higher levels of virus (as assayed using RT-PCR) at the time of delivery. However, the

rate of infection was observed to be substantially increased in mothers co-infected with HCV and human immunodeficiency virus (HIV) possibly due to an increase in HCV titre as a result of immunosuppression by HIV co-infection (Thaler *et al.*, 1991). The majority of HCV infections are identified in individuals from low socio-economic backgrounds where infection is often associated with some form of high-risk behaviour. However, the exact cause of infection remains undetermined in a low percentage of infected persons (DiBisceglie, 1998).

1.1.7. Clinical Manifestation and Natural History of HCV Infection

HCV infection can progress through three broad stages of clinical manifestations in the liver: acute HCV infection, chronic HCV infection, and chronic liver disease including cirrhosis, which can lead to HCC. In addition, HCV may be associated also with certain non-hepatic diseases that may result either directly from viral replication or the immune response to chronic viral infection (Table 1.1).

1.1.7.1. Acute HCV

Most patients with acute infection are symptom-free and only 20-30% of infected persons develop jaundice (Lavanchy *et al.*, 1999). As a result, there are few reliable studies on the natural history of acute infection. However, prospective studies of post-transfusion HCV indicate that the incubation period averages 7 to 8 weeks as determined by analysis of alanine transaminase (ALT) levels (DiBisceglie *et al.*, 1991). In a typical HCV infection, levels of ALTs, released from damaged hepatocytes, fluctuate markedly during this phase of infection. Fulminant liver failure associated with HCV was observed in Japan, however its occurrence appears to be rare elsewhere geographically (Yoshida *et al.*, 1994).

1.1.7.2. Chronic HCV

In about 80% of individuals, HCV infection becomes chronic as demonstrated by the persistence of viral RNA in serum (DiBisceglie, 1998). Chronic HCV infection is a leading cause of severe liver disease (chronic hepatitis) which can necessitate liver transplantation due to liver failure (Esteban, 1993). In chronically HCV-infected patients there is a slow course of disease development with the vast majority of patients remaining well after approximately 10 years post infection. These individuals display normal liver

enzyme levels and relatively normal liver histology (Lavanchy *et al.*, 1999). Other factors such as associated alcohol consumption, co-infection with HBV or immune deficiency due to co-infection with HIV exacerbate the effects of infection (DiBisceglie, 1998). Different viral factors have been suggested to influence the persistence of infection such as the degree of viral genetic diversity (quasispecies) found in an individual patient, mode of infection (level of viral inocula), and HCV genotype (especially type 1; Booth, 1998; Gomez *et al.*, 1999).

1.1.7.3. Chronic Liver Disease

Disease progression during chronic HCV infection is typically slow, but patients with chronic hepatitis who have presumably elevated ALT enzymes have an increased likelihood of progressing to more serious disease (DiBisceglie, 1998). ALT enzymes are released by damaged hepatocytes and thus their levels are an indirect, although imperfect, indicator of insult to the liver. Cirrhosis of the liver occurs in less than 20% of the patients with HCV and is usually observed in the second or third decade after infection. The initial appearance of hepatic fibrosis (during tissue repair) seems to be an important predictor of the later development of cirrhosis (Yano *et al.*, 1996).

HCC may develop in as many as 1 to 4% of patients with established cirrhosis (DiBisceglie, 1997) and is a late complication of HCV. In some countries, dual infection with HCV and HBV appears to potentiate the development of HCC (Idilman *et al.*, 1998). Although epidemiological, clinical, and molecular data demonstrated a close association between chronic HBV infection and HCC development, the contribution of chronic HCV infection to the development of this malignancy is less clear (Idilman *et al.*, 1998). In the case of HCV, development of HCC may be related to the presence of cirrhosis (Idilman *et al.*, 1998).

1.1.7.4. Extra-Hepatic HCV Associated Diseases

A variety of extra-hepatic clinical syndromes that are thought to be either “autoimmune” or immune complex-related diseases have been associated with HCV infection. They include cryoglobulinemia (aggregation of certain immunoglobulins at low temperatures, resulting in obstruction of small blood vessels), glomerulonephritis (a kidney disorder involving the glomeruli and antibody-antigen reaction), and arthralgias (pain in joints in the absence of

inflammation characteristic of arthritis; Johnson *et al.*, 1993). Porphyria cutanea tarda presents an example of the type of autoimmune disease strongly associated with HCV infection. The disease is characterised by photosensitivity, increased skin fragility, hyperpigmentation, and sclerodermoid plaques (Decastro *et al.*, 1993). The role of HCV as a causative factor is not clear for a number of these conditions.

1.1.8. Immune Response

Disease states associated with HCV infection may occur not only through a direct cytopathic effect of the virus but also through the immunological response to infection by the infected host (Lavanchy *et al.*, 1999). Challenge experiments on a convalescent infected chimpanzee with the same or a different HCV strain resulted in the reappearance of viremia that corresponded to the challenge virus. This observation suggests that HCV infection does not elicit protective immunity against re-infection with homologous or heterologous strains (Okamoto *et al.*, 1994b). However, the extent to which these observations can be reproduced in other infected chimpanzees is not clear. In humans, antibodies against the non-structural protein 4 (NS4) and structural core protein are detected during acute HCV infection, which makes them useful diagnostic markers (Hosein *et al.*, 1991; Chang *et al.*, 1999).

Comparison of the amino acid sequences of many HCV isolates (from all types) revealed two highly variable regions, the first of which is located at the N-terminus of E2 and is termed hypervariable region 1 (HVR-1). During chronic HCV infection, mutations in HVR-1 have been suggested to provide a mechanism for escape from the immunosurveillance system (Kato *et al.*, 1993). An asymptomatic carrier with viremia but with a low titre of anti-HVR-1 antibody is thought to develop amino acid alterations at intervals of several months (Kato *et al.*, 1994). However, immunisation with envelope glycoproteins has been associated with partial protection against an intravenous challenge with homologous HCV in five out of seven chimpanzees (Choo *et al.*, 1994). Therefore, anti-E2 antibodies are suggested to be involved in virus neutralisation, although the scope of neutralisation with such antibodies may be limited to certain isolates of HCV due to strain sequence variability (Choo *et al.*, 1994). A HCV-specific cytotoxic T-lymphocyte (CTL) response has been demonstrated in intra-hepatic inflammatory infiltrates and in circulating blood of patients with chronic hepatitis C (Kita *et al.*, 1993). This supports the possibility that liver injury by HCV infection is mediated immunologically (Kita *et al.*,

1993). However, CD8⁺ CTL activated by class I major histocompatibility complex (MHC-1) response is weak, because although conserved epitopes in the HCV non-structural protein 3 (NS3) were recognised, the T cells failed to detect the minor variants within HCV quasispecies (Weiner *et al.*, 1995). This may allow the emergence of HCV escape mutants (Weiner *et al.*, 1995). Thus, chronic infection by HCV is believed not to occur because of a lack of humoral or cellular immune responses. Rather, any response may be rendered ineffective by the high mutation rate of the virus which produces escape variants. More recent evidence suggests that core protein can modulate antigen presentation (Large *et al.*, 1999). Therefore, the virus may also have evolved other mechanisms to circumvent the immune response.

1.1.9. Diagnosis and Histopathology

Established methods to detect HCV infection include identification of antibodies against HCV antigens and detection of the viral genome by PCR. However, the detection of HCV antigens themselves in the infected host is generally inefficient. There are several sensitive enzyme-linked immunosorbent assays (ELISAs) to detect HCV antibodies to peptides from core, NS3, NS4, and NS5 proteins or to a combination of recombinant viral proteins. Despite several modifications to maximise sensitivity, these assays are still plagued by some false-positive reactivity. In PCR-negative cases, specificity can be confirmed with a supplementary assay such as the recombinant immunoblot assay (RIBA), which detects antibodies to each of the HCV proteins in a nitrocellulose strip format (DiBisceglie, 1998).

During early stages of acute infection with HCV and in immunosuppressed individuals (e.g. those with chronic renal failure and AIDS) anti-HCV antibodies may not be detected. Under these circumstances direct detection of HCV nucleic acid is performed by reverse transcription followed by PCR (RT-PCR). This has proved to be especially important for accurate monitoring of the antiviral effects in patients undergoing therapy (Alter *et al.*, 1995). In addition, liver biopsies have also become an important tool to determine chronic hepatitis by permitting estimation of the degree of inflammation (grade of hepatitis) and the amount of fibrosis present (stage of hepatitis), as well as to exclude other coincidental disease (Dhillon & Dusheiko, 1995). However, these genetic and serological diagnostic tests are not able to detect the heterogeneous HCV strains that may exist in a given patient with equal sensitivity (Lavanchy *et al.*, 1999).

1.1.10. Detection of HCV RNA and Antigen in the Liver

HCV antigen localisation using antibodies against several epitopes has been reported (Gonzalez-Peralta *et al.*, 1994). Overall, in any given tissue section only a few antigen containing cells may be detected. In hepatocytes from an infected chimpanzee, granular cytoplasmic staining was detected by monoclonal antibodies raised against HCV core antigen, and the pattern was repeated in liver biopsies from patients positive for HCV RNA and anti-HCV antibodies (Yap *et al.*, 1994). Detection of immunopositive cells showed no relationship with inflammation or necrosis (Dhillon & Dusheiko, 1995). A report described increasing immunopositivity with advanced disease, but HCV viremia was found to be lowest in these patients (Magrin *et al.*, 1994). Furthermore, immunostaining for HCV RNA showed similar cytoplasmic staining patterns were observed, with perinuclear localisation in hepatocytes (Nouri Aria *et al.*, 1993). These HCV RNA-positive cells appeared to be bile duct cells and lymphocytes (Tanaka *et al.*, 1993; Walker *et al.*, 1998). Employing the novel technique of *in situ* hybridisation (ISH) followed by PCR (ISH/PCR), the sensitivity has been significantly enhanced to detect positive cells containing less than 20 copies of the viral genome (Nuovo *et al.*, 1993). A study of strand-specific localisation-ISH, showed that the replicative intermediates of HCV were most closely associated with inflammation at the periphery of fibrosis cavities. However, there also is a lack of reproducibility reported with these techniques (Hiramatsu *et al.*, 1992; Yap *et al.*, 1994; Park *et al.*, 1996).

1.1.11. Prevention and Treatment

The incidence of transfusion-associated hepatitis C has decreased to less than 1% per recipient since the implementation of screening volunteer donors for HCV antibodies. To prevent transmission by potentially contaminated blood products, inactivation of nucleic acid by photo-reactivation chemicals has proved to be effective compared to traditional protein denaturation procedures where both viral and non-viral proteins were simultaneously destroyed (Houghton, 1996). Additional prevention strategies (e.g. needle and syringe availability) are targeted at other routes of transmission (see Section 1.1.6).

The primary goal of treatment for HCV infection is to prevent serious chronic liver disease including cirrhosis, end-stage liver disease, and HCC. The secondary goal is to reduce or eliminate the infectivity of the infected person so that the virus is not transmitted to others in the community. At present, alpha interferon (IFN- α) is the most widely accepted form of therapy for chronic HCV infection (Davis *et al.*, 1989). The optimal response is defined by persistently normal serum ALT concentrations and absence of HCV RNA from serum at the end of therapy and for at least 6 months thereafter (Poynard *et al.*, 1996). This is referred to as a sustained response, which occurs in only 20-25% of patients treated for 6 months. In addition, other factors are also associated with an increased chance of responding, including the absence of cirrhosis, lower pre-treatment HCV RNA levels, and infection with HCV genotypes other than type 1 (DiBisceglie, 1998). However, certain patient groups may continue to be difficult to manage, including those with immune deficiency due to the use of immunosuppressive drugs, HIV infection, or chronic renal failure. Furthermore, patients with features of both HCV infection and autoimmune liver disease may be at risk of developing further liver damage by the use of interferon.

Alternatives to interferon therapy have been considered because of the low proportion of patients who show sustained response, the high cost, and side effects associated with this cytokine. An adjunct to interferon, ribavirin, is an orally administered nucleoside analogue with a broad spectrum of antiviral properties. Current studies suggest that when ribavirin is used together with interferon, there is a significantly higher rate of sustained response than with interferon alone (Brillanti *et al.*, 1994). The mechanism by which ribavirin enhances the effect of interferon is not fully understood, although it may act as an immune stimulant in conjunction with other cytokines (Gish, 1999).

At present the field of HCV treatment awaits the development of new agents with specific antiviral properties targeted towards the viral protease, helicase, or RNA-dependent RNA polymerase. The crystal structure and tertiary structure of the protease and helicase have been determined recently, which could aid the search for compounds to inhibit their functions (Section 1.2.6).

Advances in molecular characterisation of HCV may help in the development of vaccine production. Analysis of HCV variation, serology and *in vitro* experiments testing HCV infectivity, suggested the requirement for the development of a broadly reactive vaccine against HCV (Kato *et al.*, 1993 and 1994, Shimizu *et al.*, 1994). However, the

development of a vaccine against HCV has faced multiple challenges. Although reliable neutralising determinants against immunogenic antigens are yet to be documented, preliminary data suggest that HVR-1 in the N-terminal part of envelope glycoprotein E2 contains B-cell and CTL epitopes (Zibert *et al.*, 1995). CTL responses against non-structural protein NS3 have also been studied (Ferrari *et al.*, 1994). The obstacles that hinder the development of an efficient HCV vaccine include: a) the lack of an *in vitro* replication assay; b) the lack of a small susceptible animal model (Inchauspe, 1999).

1.2. Molecular Features of HCV

1.2.1. Comparison of the HCV, Flavi- and Pestivirus Genomes

Members of the *Flaviviridae* family are divided into three genera, flaviviruses, pestiviruses and hepaciviruses (Rice, 1996). The genera have diverse biological properties and show no serological cross-reactivity, however, they appear to be similar in terms of virion morphology, genome organisation, and presumed RNA replication strategy. The genomes of *Flaviviridae* consist of a positive-sense, single-stranded RNA of between 9.5 to 12.5 kb, with HCV having the shortest genome (Rice, 1996). The viral RNA encodes a single polyprotein of approximately 3 460, 3 960, and 3 010 amino acids for yellow fever virus (YFV), bovine viral diarrhoea virus (BVDV), and HCV, respectively, with structural proteins at the N-terminus followed by the non-structural proteins. Generally, the structural proteins are generated by processing of the polyprotein by host proteinases while the majority of mature non-structural proteins are produced by virally encoded proteinases. Features of the genomic organisation of *Flaviviridae* are illustrated in Fig 1.2 and Table 1.2.

The genome of the prototype flavivirus, YFV is approximately 11 kb, with a type I cap at its 5' end ($m^7GpppAmp$) containing two methyl groups, followed by a conserved dinucleotide sequence AG (Rice, 1996). Although this is the predominant type of cap in all eukaryotic messenger RNAs (mRNAs), genome-length RNAs appear to be the only virus-specific mRNA molecules in flavivirus-infected cells. Genomic RNAs of insect-borne flaviviruses lack a 3' terminal poly (A) tract and terminate with the conserved dinucleotide CU. The single ORF encoded by the genome is flanked by 5' (118 nucleotides) and 3' (511 nucleotides) untranslated regions (UTRs) containing conserved elements, which play a role in replication (Brinton & Dispoto, 1988; Mandl *et al.*, 1993). Secondary structures can be predicted for the 3' terminal 90 bases of the genome of all flaviviruses, however there is little conservation of nucleotide sequences (Mandl *et al.*, 1993). For most viruses in this genus, translational initiation occurs at the first AUG in the genome, but in a small number of mosquito-borne flaviviruses [e.g. YFV and Dengue (DEN) type 1-4 viruses], a second in-frame AUG is located 12 to 14 codons downstream. Ten viral proteins are cleaved co-translationally and post-translationally from the primary translation product. Fig 1.3 illustrates the membrane topology of the flavivirus proteins;

PrM, the precursor to structural protein M, envelope protein E and non-structural protein NS1 are located in the lumen of the endoplasmic reticulum (ER), while the nucleocapsid (C) and remaining non-structural proteins, NS2 to NS5 are found in the cytoplasm (Rice, 1996).

The genomic RNA of BVDV, the prototype of the pestiviruses, has a single-stranded RNA of 12.5 kb in length (Brock *et al.*, 1992). The ORF is ~11.9 kb and is flanked by a 5' UTR of 385 bases and a 3' UTR of 229 bases (Collett *et al.*, 1988). Pestivirus genomic RNA does not contain a 3' poly (A) tract but appears to terminate with a short stretch of poly(C). Direct analysis of the 5' terminus suggested that the genomic RNA lacked a cap structure (Brock *et al.*, 1992; Deng & Brock, 1993). The failure to detect subgenomic RNAs, as for flaviviruses, has led to the belief that genome-length RNAs serve as mRNAs for translation of the viral polyprotein. The long pestivirus 5' UTR is predicted to form a highly structured RNA element that may behave as an IRES to initiate cap-independent translation of the ORF (Brown *et al.*, 1992; Deng & Brock, 1993; Poole *et al.*, 1995). The combination of host and virally encoded proteinases generates eleven distinct proteins from the polyprotein. Unlike the flaviviruses and HCV, the first pestivirus protein encoded in the ORF is a non-structural autoprotease (N^{pro}) involved in the processing between N^{pro} and core protein (Table 1.2 and Fig 1.2).

By comparison with flaviviruses and pestiviruses, HCV genomic RNA is considerably shorter at about 9.4 kb in length and is separated into 3 regions (5' UTR, ORF and 3' UTR; Okamoto *et al.*, 1994a). Given the considerable genetic divergence among isolates, the HCV 5' UTR is the most conserved RNA sequence element and is 341 to 344 bases long. The 3' UTR consists of three distinct segments, a short region variable in composition and length (28 to 42 bases), a variable length polyuracil tract [poly (U)] and a highly conserved stretch of 98 nucleotides termed the 3'X tail (Han *et al.*, 1991; Tanaka *et al.*, 1995). Conserved secondary structures, predicted at the 3' end of the ORF and the 3'X tail, may be important in the replication of HCV RNA. The length of the ORF varies slightly among isolates, encoding polyproteins of between 3,010 to 3,033 amino acid residues (Chamberlain *et al.*, 1997). Fig 1.4 illustrates details of the genome structure of HCV. There are a number of short amino acid sequence homologies with NS3 and NS5 of the flaviviruses (Miller & Purcell, 1990; Choo *et al.*, 1991). The region between residues 1230 to 1500 of the HCV polyprotein (located in NS3) has sequence similarities with a putative NTP-binding helicase encoded by flavi-, pesti-, and plant potyviruses (Hijikata *et al.*,

1993b). A region upstream of the NTP-binding helicase shares similar features with the putative (chymo)trypsin-like serine proteinase from flaviviruses and pestiviruses (Bartenschlager *et al.*, 1993; Hijikata *et al.*, 1993a). Finally, the amino acid segment 2703-2739, contains the characteristic Gly-Asp-Asp motif, which is highly conserved among viral-encoded RNA-dependent RNA polymerases (Koonin, 1991a).

In the absence of an efficient viral replication system and a small animal model the majority of the molecular characterisation HCV has been conducted using a number of heterologous expression systems. Additional information has been provided by structural crystallography studies for one viral protein, NS3 (Kim *et al.*, 1996; Love *et al.*, 1996).

1.2.2. 5' Untranslated Region (5' UTR) and Translation of HCV RNA

The 341 (to 344) nucleotides of the 5' UTR and adjacent core protein coding sequences are extremely well conserved between HCV isolates (Simmonds, 1995). The HCV 5' UTR is much longer than that of flaviviruses and no significant sequence homologies with this region of flaviviruses exist (Heinz, 1992). However, the HCV 5' UTR does closely resemble the 5' UTRs of pestiviruses BVDV and classical swine fever virus (CSFV) in terms of length and several short regions of significant homology (Houghton *et al.*, 1991). This region of the genome is predicted to be capable of forming extensive secondary structures and this has been confirmed from biochemical structural studies (Brown *et al.*, 1992). Four major structural domains have been identified, designated I to IV, that consist of stem-loop and pseudoknot structures (Fig 1.5). A number of AUG codons are located in the 5' UTR, but these are apparently not used for initiation of translation. The presence of such a long highly ordered untranslated region containing non-functional initiation codons resembles the situation in picornaviruses. The picornaviral genomes are translated by a mechanism of internal ribosome binding and initiate translation at a specific AUG codon in a cap-independent manner (Jackson *et al.*, 1990b); such RNA elements have been termed internal ribosomal entry sites (IRES). Similar to the picornaviral 5' UTR, *in vitro* expression systems have been used to identify IRES function for the HCV 5' UTR (Tsukiyamakohara *et al.*, 1992). However, unlike picornaviruses, the 3' boundary of the HCV IRES element is located within the coding sequence of the core protein and therefore is not entirely contained within untranslated sequences. Mapping studies have revealed that optimal IRES activity resides within a segment from nucleotides 40 to 370, and thus incorporates at least 29 nt from the core protein coding region (Reynolds *et al.*, 1995). The

AUG codon at position 342, utilised for initiation of translation, is located in stem-loop IV (see Fig 1.5). The genomes of some HCV genotypes are translated less efficiently than others. RNA-RNA interactions between the extreme 5' UTR and core protein coding sequences adjacent to the IRES are responsible for this regulated translation (Honda *et al.*, 1999). In tissue culture studies, the 5' UTR from genotype 2b was three times more efficient than that for genotype 6a in directing translation from a reporter construct (Collier *et al.*, 1998).

Despite the conceptual similarities between the IRES elements of HCV and picornaviruses (Jackson *et al.*, 1990b), they differ in both primary and predicted secondary structures. 5' UTRs of HCV and pestiviruses mediate initiation of translation by a different mechanism of internal ribosomal entry from that of picornaviruses (Fig 1.6; Stewart & Semler, 1997), in that cap-dependent translation is not abolished in the cell following HCV infection. In addition, this mechanism has unique features compared with those for other known mRNAs, including capped mRNAs and cellular RNAs translated via IRES elements of genes such as homeotic gene Antennapedia (Macejak & Sarnow, 1991). Briefly, a 43S preinitiation complex is formed by the assembly of eukaryotic initiation factor 3 (eIF-3), eIF-2, GTP, initiator tRNA (met-tRNA) and a 40S ribosomal subunit (Fig 1.6). Capped mRNA requires additional factors (eIF-4A, -4B, -4E, -4F, and -4G) to unwind any secondary structure adjacent to the capped 5' end to create a binding site for the 43S complex (Pestova *et al.*, 1996). The HCV IRES recruits and directly binds to components of the 43S complex, without the requirements of eIF-4A, -4B, -4E, -4F, and -4G. The structural determinants in the IRES ensure the correct spatial orientation of these binding sites for attachment of the final initiation complex at the initiation codon (Hellen & Pestova, 1999). UV cross-linking studies have identified several cellular RNA-binding proteins such as La autoantigen (Meerovitch *et al.*, 1993; Isoyama *et al.*, 1999) and polypyrimidine tract-binding protein (PTB; Borman *et al.*, 1993) that associate with the 5' UTR. These proteins have been implicated in the function of picornavirus IRES elements although they are not required for initiation from the HCV IRES (Hellen & Pestova, 1999). Nevertheless, these and other such proteins may affect the efficiency of translation *in vivo*, or ensure that only a full length genomic RNA is translated (Ito & Lai, 1999).

1.2.3. 3' UTR

As stated previously, the 3' UTR is considered to be a tripartite structure comprising the short variable sequence, a poly(U) tract and the highly conserved 3'X sequence (Fig 1.7). Computer modelling has predicted that this sequence can fold into an elaborate stem-loop structure suggestive of a critical functional component in virus replication as shown for other members of the *Flaviviridae* family (Kolykhalov *et al.*, 1996). UV cross-linking experiments and the yeast three-hybrid system have revealed the binding of PTB (Ito & Lai, 1997) and human ribosomal protein L22 (Wood & Patel, 1999) to the 3'X segment of the 3' UTR. Both or either protein may participate in the regulation of HCV RNA synthesis or translation, however such proposals await functional analysis. Interestingly, PTB binds to the 5' UTR of HCV and may also have a role in the translation of the virus (Ito & Lai, 1999).

1.2.4. HCV Virus-Encoded Non-Structural Proteins

As stated earlier, the HCV genome encodes a single polyprotein. Cellular enzymes are responsible for the cleavage events which generate the structural proteins while processing of the NS2-NS5B region is mediated by two virus-encoded proteolytic activities. The NS2-NS3 junction is cleaved by a NS2/3 proteinase and the NS3/4A serine proteinase is responsible for the processing at sites between NS3/4A, NS4A/B, NS4B/5A, and NS5A/B. The polyprotein is processed into at least 10 distinct viral proteins (illustrated in Fig 1.4). The structural proteins are located in the N-terminal one-third and the non-structural and replicative enzymes located within the C-terminal two-thirds of the polyprotein. The properties of the non-structural (NS) proteins are briefly discussed in this section and are followed by a description of the structural core and envelope proteins (Section 1.3).

1.2.5. NS2

The NS2 protein extends from aa 810 to 1026 and is a transmembrane protein with its C-terminus translocated into the lumen of the ER while its N-terminus lies in the cytosol (Santolini *et al.*, 1995). Further evidence has suggested that NS2 is concentrated near the nucleus in the perinuclear space and at the edges of the cell (Kim *et al.*, 1999).

Immunoprecipitation studies have shown that NS2 can associate with the structural proteins (Matsuura *et al.*, 1994). Aside from its hydrophobic character, no significant sequence homology exists between the HCV NS2 protein and the analogous flavivirus or pestivirus proteins. Nonetheless, detailed study of the proteolytic processing of this region of the polyprotein has revealed a unique proteinase function contributed by the C-terminal portion of NS2 and the NS3 serine proteinase domain. Although cleavage at the NS2-NS3 junction involves amino acid residues 846 to 1237, which span NS2 and NS3 (Santolini *et al.*, 1995), the catalytic activity of the serine proteinase is not required (Hijikata *et al.*, 1993a). The NS2-NS3 precursor is co-translationally targeted via signal recognition particle (SRP) and SRP-receptor pathways to the ER membrane. This membrane association enhances cleavage at the NS2-NS3 junction, resulting in translocation of the C-terminus of NS2 to the ER lumen and the N-terminus being located in the cytosol, as shown by site directed mutagenesis and proteinase digestion (Santolini *et al.*, 1995). Mutagenesis studies have also identified His and Cys residues (at positions 952 and 993 in the polyprotein) that may form a catalytic dyad or, together with a Glu residue (at aa 972), a catalytic triad (Fig 1.8; Hijikata *et al.*, 1993a; Bartenschlager, 1999). However, mutations at residues 952, 972, and 993 has no effect on the processing of NS3 cleavage sites, suggesting that despite the overlap between the proteinases, their activities are independent of each other. Finally, the NS2/3 autoproteinase is proposed to be a metalloproteinase (containing a zinc ion) based on metal chelating studies. Proteolysis was shown to be tolerant to amino acid mutations in the region from P5 to P3' of the NS2-NS3 junction (aa 1022 to 1029; Reed *et al.*, 1995). However, a preference for hydrophobic amino acids at the P1 and P1' positions (aa 1026 and 1027) showed similarities to neutral metalloproteinases, such as a mammalian enzyme, endopeptidase 24.11, which is involved in processing of peptide hormones (Erdos & Skidgel, 1989; Clarke, 1997). The requirement of a cysteine residue and a metal ion for NS2/3 proteinase catalytic activity suggests employment of a unique mechanism, and thus this enzyme cannot be grouped with any of the currently known proteinase families (cysteine, metallo, serine or aspartic proteinases; Bartenschlager, 1999).

1.2.6. NS3

This HCV protein has been the subject of intensive study due to its potential as an anti-viral target. NS3 is the only HCV protein for which structural information is available. The protein is proposed to be the functional homologue of the flavivirus NS3 and the

pestivirus 125/p80 proteins. It extends from amino acids 1027 to 1658, has an apparent molecular weight of 70 kDa and is a multifunctional molecule with three enzymatic activities; a serine-type proteinase function located in the N-terminal domain and nucleoside triphosphatase (NTPase)/helicase activities in the C-terminal region (Fig 1.9; Grakoui *et al.*, 1993b; Hijikata *et al.*, 1993a). NS3 may have other non-enzymatic properties. There is evidence that it has a role in modulating protein kinase A (PKA)-mediated signal transduction and thus may influence host cell function (Borowski *et al.*, 1996). In HCV-infected liver cells, NS3 is present in the cytoplasm, but in a minority of HCV-infected cells both the cytoplasmic and the nucleus or the nucleus on its own is positive for NS3 (Errington *et al.*, 1999).

1.2.6.1. NS3 Proteinase

The minimal NS3 domain required for full proteolytic activity has been mapped to about 180 residues at the N-terminus (Bartenschlager *et al.*, 1994; Failla *et al.*, 1994). NS4A is a cofactor of the NS3 proteinase and essential for its proteolytic activity at most NS cleavage sites (Bartenschlager *et al.*, 1994; Failla *et al.*, 1994). Release of the N-terminus of NS3 by cleavage at the NS2-NS3 junction and formation of the NS3/4A serine proteinase may be an essential step in HCV replication. In pestiviruses, viral replication is able to proceed in the absence of cleavage at a similar position, but the resultant viruses are rendered non-cytopathic (Ryan *et al.*, 1998). The HCV NS3 serine proteinase appears to more closely resemble its pestivirus counterpart, rather than the flavivirus homologue, where there is a functional requirement for NS2 as a closely associated cofactor (Ryan *et al.*, 1998).

Based on comparative sequence analyses with several viral and cellular proteinases, NS3 contains a (chymo)trypsin superfamily motif, composed of a catalytic triad of His (1083), Asp (1107) and Ser (1165), in the N-terminal domain (Fig 1.9; Miller & Purcell, 1990). Substitutions affecting these residues abolish cleavage at the NS3-NS4A, NS4A-NS4B, NS4B-NS5A, and NS5A-NS5B junctions but not at the NS2-NS3 junction (Bartenschlager *et al.*, 1993). The processing between NS3 and NS4A is rapid, co-translational, and occurs in *cis*, while cleavage at the other NS3/4A-dependent sites occurs in *trans* (Bartenschlager *et al.*, 1994). NS4A cofactor is required for cleavage at the NS3-NS4A, NS4A-NS4B, and NS4B-NS5A sites and enhances cleavage at the NS5A-NS5B junction (Failla *et al.*, 1994). Sequence analysis of various HCV isolates revealed a consensus sequence for NS3/4A-dependent cleavage characterised by an acidic amino acid at the P6-position, a Cys/Thr at

P1, and a residue with a small side-chain at position P1'. The consensus is Asp/Glu-X-X-X-X-Cys/Thr↓Ser/Ala-X-X-X, where X is variable (Grakoui *et al.*, 1993a).

The NS3 proteinase domain and NS4A form a detergent-stable, non-covalent complex. Association requires 22 N-terminal residues of NS3 and a 12 residue sequence in the centre of NS4A (Fig 1.9; Bartenschlager *et al.*, 1995; Lin *et al.*, 1995). Mutagenesis studies suggested that NS4A enhances proteinase activity by inducing conformational changes to stabilise NS3 (Kim *et al.*, 1996; Steinkuhler *et al.*, 1996). The three dimensional structure of the proteinase domain of NS3 has been determined in the presence (Kim *et al.*, 1996) and absence (Love *et al.*, 1996) of a synthetically produced region of NS4A. In the presence of the NS4A peptide, the proteinase domain adopts a chymotrypsin-like fold. This consists of N- and C-terminal domains which contain 8 (one of which is contributed by NS4A) and 6 β -barrels respectively, separated by a cleft. In the absence of NS4A, the N-terminal domain adopts an open conformation that extends away from the protein. Under these conditions, the β -strands at the N-terminus contact exposed hydrophobic patches on adjacent molecules (Love *et al.*, 1996). This open conformation affects two residues of the catalytic triad (His-1083 and Asp-1107) such that the Asp residue no longer forms a hydrogen bond with His. Thus, binding of NS4A may effect conformational changes at the triad, which in turn have a bearing on proteolytic activity (Love *et al.*, 1998).

Analysis of the crystal structure of NS3 serine proteinase has indicated that the enzyme contains a zinc binding site co-ordinated by three Cys residues (Cys 1123, 1127, and 1171), and via a water molecule, by His (1175; Fig 1.8; Kim *et al.*, 1996; Love *et al.*, 1996). The role of the zinc ion appears to be strictly structural, because it can be substituted by cadmium or cobalt ions (De Francesco *et al.*, 1996). Mutagenesis of any of these residues negatively affects the serine proteinase as well as the NS2/3 autoproteinase (Hijikata *et al.*, 1993a). This suggests that both the NS2/3 proteinase and NS3 serine proteinase share the same metal binding site. Interestingly, a zinc ion co-ordinated in this way (i.e. to three amino acids and one water molecule) is generally viewed as having a catalytic role, and therefore the zinc-bound water molecule may have an essential function in the hydrolysis of the NS2-NS3 site (Wu *et al.*, 1998). Catalytically-active zinc is usually co-ordinated by nitrogen and oxygen atoms. For those enzymes where zinc has a structural role, it is ligated by cysteine thiolates, as is the case in NS3 (Berg & Shi, 1996).

It is possible that zinc co-ordinated in this way could have both catalytic (for NS2-3 cleavage) and structural (for stability of the NS3/4 complex) functions. The residues involved in the binding of zinc by NS3 are conserved among HCV genotypes although this is not the case in the flavivirus and pestivirus homologues (Ryan *et al.*, 1998), which do not possess metalloproteinase activity.

1.2.6.2. NS3 NTPase/Helicase

The two other enzymatic processes associated with NS3 are NTPase and RNA helicase activities, for which the C-terminal 465 amino acid residues are sufficient (Kim *et al.*, 1997a). From biochemical studies, the proteinase domain is required for the optimal binding of polynucleotides and RNA to the NS3 NTPase/helicase domain and thus there is interplay between both regions of NS3 (Gallinari *et al.*, 1998). Hence, combination of the proteinase domain along with regions with NTPase/helicase activities into a single polypeptide may have occurred to optimise the overall function of the protein. In the flavivirus counterpart protein, two functionally distinct domains also reside in the same polypeptide (Gorbalenya *et al.*, 1989).

In common with other RNA helicases, the HCV NS3 helicase domain contains four conserved motifs, designated A, B, C, and D (Fig 1.9). Motif A, Ala-X-X-X-X-Gly-Lys-Ser, is required for ATP binding and consequently for ATP hydrolysis and RNA unwinding (Kim *et al.*, 1997b). The second motif, Asp-Glu-Cys-His (DECH; Yamada *et al.*, 1994), suggests NS3 is a member of DEXH (Asp-Glu-X-His) subfamily belonging to the DEAD (Asp-Glu-Ala-Asp) box family of RNA helicases (Koonin, 1991b). Mutational analysis of the helicase domain showed that the conserved residues within the DEAD box motif are essential for NTPase and helicase activities (Kim *et al.*, 1997b, Wardell *et al.*, 1999). Similar to NS3, a number of DEAD box family proteins contain NTPase and RNA helicase activities (e.g. eIF-4A; Kim *et al.*, 1997b). Motifs C and D are TAT (Thr-Ala-Thr) and QRRGRTGR (Gln-Arg-Arg-Gly-Arg-Thr-Gly-Arg). The former is believed to be important in RNA unwinding and the latter is important in ATP hydrolysis and RNA unwinding (Kim *et al.*, 1997b). The NTPase activity is dependent on Mg²⁺ and enhanced by the presence of polynucleotides, poly(U) in particular (Tai *et al.*, 1996). This region of NS3 is able to unwind RNA, DNA and RNA/DNA duplexes in a 3' to 5' direction although activity is dependent on a 3' single-stranded region (Tai *et al.*, 1996). This property of NS3 makes it unique among RNA helicases including the related pestivirus

enzymes. The minimal RNA-binding size for the HCV helicase is estimated to be between seven and twenty nucleotides (Bartholomeus & Thompson, 1999). The X-ray crystal structure of this region of NS3 indicates that the NTPase and RNA binding domains are distinct (Kim *et al.*, 1998).

1.2.7. NS4A and NS4B

NS4A extends from amino acid residues 1658 to 1712 and is approximately 8 kDa in size. As described above, it functions as a co-factor to the NS3 serine proteinase. The protein also forms a non-covalent, detergent-stable complex with the NS4B-NS5A polyprotein, which may allow efficient cleavage at the NS4B-NS5A junction (Lin *et al.*, 1997). The N-terminal hydrophobic region of NS4A forms a 20 residue transmembrane helix that is followed by highly charged residues. In addition to enhancing proteolytic cleavage activity of NS3, NS4A mediates the association of NS3 with intracellular membranes, probably via the transmembrane helix. This may contribute to the activity of the proteinase *in vivo* (Hijikata *et al.*, 1993b).

NS4B has an apparent molecular weight of ~27 kDa and is generated by processing at residues 1712/1713 and 1972/1973 of the polyprotein. It is highly hydrophobic and may be membrane associated (Shimotohno and Feinstone, 1997). No functions have been ascribed to NS4B proteins of the *Flaviviridae* family including that encoded by HCV. Recent studies suggest the HCV protein may have a role in the phosphorylation of NS5A (Koch & Bartenschlager, 1999) and could be important for anchoring replication complexes to membranes. The NS4 region as a whole has two strong antigenic elements, one of which overlaps the C-terminus of NS4A and the N-terminus of NS4B. The other region is located at the C-terminus of the NS4B protein (Chang *et al.*, 1999).

1.2.8. NS5A

NS5A is generated by processing at residues 1972/1973 and 2419/2420. Expression of NS5A produces two proteins of different electrophoretic mobilities, p56 and p58. Both species are phosphoproteins and the difference in size of the two forms is a consequence of differential phosphorylation (Tanji *et al.*, 1995). Evidence suggests that both proteins are phosphorylated at serine residues in the C-terminal region of NS5A and that p58 has

additional phosphorylation sites at serine residues in the central region spanning residues 2197 to 2204 (Tanji *et al.*, 1995). Mutagenesis studies showed hyperphosphorylation to produce p58 was most efficient when NS5A was expressed as part of a continuous NS3-NS5A polyprotein (Koch & Bartenschlager, 1999). Direct interaction of NS5A with NS4A occurs via amino acids at 2135 to 2139 in NS5A (Asabe *et al.*, 1997). NS5A protein of DEN-2 is also phosphorylated, which affected its subcellular localisation and affinity for NS3 (Kapoor *et al.*, 1995). However, the physiological role of the phosphorylation of NS5A and its role in viral replication is not known as yet.

Apart from any role in virus replication, NS5A has been implicated as a factor that determines the susceptibility of the virus to treatment with interferon (Enomoto *et al.*, 1996). Clinical observations indicate that mutations in the amino acid sequence of NS5A, especially in a region between amino acids 2209 and 2248, are associated with improved responsiveness to IFN. This segment of the protein has been designated the IFN-sensitivity determining region (ISDR). Studies have shown that binding of NS5A directly with double-stranded RNA-dependent protein kinase (PKR) occurs through the ISDR (Gale *et al.*, 1997; Song *et al.*, 1999). PKR is an important antiviral activity, induced by IFN- α (Katze, 1995), which modulates apoptosis responses of the host cell to infection and is a tumour suppressor. Interaction with NS5A interferes with PKR kinase activity (Gale *et al.*, 1997) and the PKR-binding function of NS5A disrupts apoptosis by blocking the PKR-dependent host signal pathways (Gale *et al.*, 1999). It is proposed that this may be a mechanism by which HCV evades the host immune response.

1.2.9. NS5B

NS5B, generated by processing between residues 2419 and 2420 of the polyprotein, is a 65 kDa protein and is believed to be the viral RNA-dependent RNA polymerase (RdRp). The NS5B sequence is conserved not only between different strains of HCV but also in pestiviruses, flaviviruses and in other RNA viruses (Rice, 1996). Inspection of its amino acid sequence reveals the presence of four conserved motifs, termed A, B, C, and D in the central region of the NS5B polypeptide (Fig 1.10). Motif A is likely to have a role in NTP binding and catalysis while motif B may be involved in template and primer positioning. Mutations within the first motif (A) completely abolish polymerase activity, whereas changes made within the second motif (B) substantially reduce polymerase activity. Motif

C is GDD (Gly-Asp-Asp), a characteristic feature of RNA polymerase active sites and is essential for polymerase activity (Behrens *et al.*, 1996). Motif D, present in most RdRp and reverse transcriptases, also may be involved in NTP binding and catalysis. The NS5B polymerase is resistant to rifampicin, an inhibitor of prokaryotic RNA polymerases, but like other enzymes involved in the creation or breakage of phosphodiester bonds, it has an absolute requirement for divalent metal ions, Mg^{2+} or Mn^{2+} (Lohmann *et al.*, 1997). The polymerase can copy the entire HCV genome *in vitro*, as well as other non-specific templates without any additional viral or cellular factors (Lohmann *et al.*, 1997; Oh *et al.*, 1999). Additional factors would probably be necessary to confer specificity and may be involved also in the regulation of RNA synthesis. *In vitro* studies have shown that it uses the 3' end sequences of both plus and minus strands of HCV RNA as templates, implying that these sequences act in *cis* in a primer-independent manner (Oh *et al.*, 1999). These findings were confirmed in a study showing that all of the 3' UTR and a portion of the 3' coding nucleotide sequences, contain conserved stem-loop structures necessary for specific NS5B binding (Cheng *et al.*, 1999).

1.2.10. HCV RNA Replication

The viral genome of positive-sense, single-stranded RNA viruses functions also as mRNA in infected cells. During the replication process, HCV must synthesise negative-stranded RNA as a template for generating progeny virion RNAs. By analogy to related flaviviruses and pestiviruses, the synthesis of an intermediate negative strand RNA is likely to be catalysed by a cytoplasmic membrane-associated replicase complex formed by the non-structural proteins. The three viral proteins most likely to be directly involved in RNA synthesis are NS3, NS4A and NS5B (Behrens *et al.*, 1996; Tai *et al.*, 1996). Molecular studies of HCV replication have been hampered by the low efficiency of propagation with currently available cell culture systems and by the fact that the only animal model is the chimpanzee. Recently, Lohmann *et al.* (1999) established human hepatoma cell lines using bicistronic constructs. These were composed of the HCV IRES (nt 1 to 377) linked to the neomycin phosphotransferase (*neo*) gene, followed by the IRES of the encephalomyocarditis virus to direct translation of HCV sequences from NS2 or NS3 to NS5B, and the 3' UTR. These constructs generated approximately 10^8 positive-strand RNA copies per microgram of total cellular RNA, which corresponded to 100 to 500 molecules per cell. The amount of negative-strand RNA was 5- to 10-fold lower, consistent with the notion that negative-sense RNA serves as a template for the synthesis

of excess positive-sense RNA; a similar ratio is detected in flavivirus replication (Cleaves *et al.*, 1981). The NS3 and NS5A produced from replicon templates were localised predominantly in the cytoplasm thus providing evidence for the site of HCV RNA replication. Based on the localisation of double-stranded RNA (an indicator of replication) and cell-fractionation studies, flavivirus RNA synthesis is thought to occur principally on the membranes of the perinuclear ER (Westaway, 1987). The HCV structural proteins and NS2 were not required for replication of the replicon emphasising the close evolutionary relationship between HCV and pestiviruses (Behrens *et al.*, 1998). For pestiviruses, these replicons appear to replicate more efficiently ($\sim 1.5 \times 10^4$ genomes per cell) than HCV. The reduced amount of RNA detected for HCV replicons could reflect the intrinsic lower replication capacity of HCV.

As mentioned above, HCV RNA replication is proposed to take place on ER membranes, which has been based on available experimental evidence and replication strategies used by other RNA viruses. From Kunjin virus (KUN), DEN-2 and BVDV RNA-labelling studies, it is thought that virus RNA synthesis occurs through a semi-conservative and asymmetric replication cycle (Bartholomeus & Thompson, 1999). One model suggests that NS4A anchors the helicase and the polymerase complex to the ER (Fig 1.11; Bartholomeus & Thompson, 1999). However, since the experimental data has been gained using proteins expressed at high levels, further investigation is necessary. It is also believed that the proteins forming the RNA replication complex may also have a role in virus assembly and packaging of the viral RNA (Hijikata *et al.*, 1993b).

1.3. HCV Structural Proteins

1.3.1. Core

Core protein is located at the N-terminal end of the HCV polyprotein. Cleavage between core and E1 at residue 191 generates a 23 kDa product. The presumed mature core protein is produced by further cleavage at residue 174, resulting in a 21 kDa species. Although p21 is the major product from precursor polyprotein, a small proportion of p23 is detected also in tissue culture systems (Liu *et al.*, 1997). Processing at residues 174 and 191 *in vitro* is achieved only in the presence of microsomal membranes, suggesting that cleavage is directed by membrane-associated proteinases. The C-terminus of p23 is highly hydrophobic and constitutes the signal sequence that directs insertion of E1 sequences into the ER, while processing at residue 174 probably takes place on the cytoplasmic face of the ER. The majority of core is found in the cytoplasm, associated with ER membranes and lipid droplets (Barba *et al.*, 1997). There is evidence also of core protein in the nucleus that is conformationally different from the cytoplasmic form (Yasui *et al.*, 1997). A hydrophobicity profile and examination of the amino acid content of core highlight three separate domains (Fig 1.12). The first domain consists of residues 1 to 122 and is characterised by a high proportion of basic residues. Domain 2, defined by aa 123-174, is generally more hydrophobic than domain 1 and contains very few basic residues. The third region (aa 175-191) is highly hydrophobic and acts as the signal sequence for E1, but is absent from the mature core protein (Yasui *et al.*, 1998; Hope & McLauchlan, 1999). Some studies conducted using strain HCV-1 detected a smaller core-specific protein (p16), which was located in the nucleus (Lo *et al.*, 1995). The difference in localisation of p21 and p16 suggested that the two forms might have varying biological function. However, this lower molecular weight p16 species is not observed for other strains of HCV.

By comparison with the related flaviviruses and pestiviruses, core protein is presumed to form the capsid of virus particle. However, in this respect, very little is known about the protein, particularly since virus particles have not been efficiently detected in mammalian cells either expressing the structural region or the complete HCV genome. The only tissue culture system from which virus particles have been isolated are insect cells infected with recombinant baculovirus expressing the entire HCV structural region (Baumert *et al.*, 1998). The core protein detected in these particles is similar in size to the mature form

detected in other heterologous expression systems. To adopt a capsid structure, core protein is anticipated to homo-oligomerise. Regions for homotypic interaction of core have been identified in both domains 1 and 2 (Matsumoto *et al.*, 1996; Yan *et al.*, 1998). Sequences between residues 36 to 102 in domain 1 were critical for oligomerisation in the yeast-2-hybrid system, while regions involved in homotypic interactions in domain 2 remain less well defined.

Studies to understand additional functions of core protein have concentrated on its affect on host cell processes and the cellular proteins with which it interacts. It has proved to be difficult to establish association of core protein with viral proteins probably because, in heterologous expressing systems, core is highly insoluble and mainly membrane-bound (Santolini *et al.*, 1994; Matsumoto *et al.*, 1996). However, a putative interaction between core and envelope protein E1 has been observed (Lo *et al.*, 1996). In terms of other factors, core binds RNA (Santolini *et al.*, 1994) and interactions with cellular proteins have been identified. These include heterogeneous nuclear ribonucleoprotein K (hnRNP K; Hsieh *et al.*, 1998), lymphotoxin β receptor (LT β R; Matsumoto *et al.*, 1997), tumour necrosis factor receptor 1 (TNFR1; Zhu *et al.*, 1998) and a presumed DEAD box protein helicase (Owsianka & Patel, 1999; Mamiya & Worman, 1999). The majority of these interactions were identified through yeast-2-hybrid screening followed by validation by biochemical analyses. In all cases, domain 1 of core has been implicated as necessary for association (Fig 1.12).

From studies on cellular processes, it has become apparent that core can exert a variety of effects on cell metabolism. However, in the absence of a system modelling both virus replication and the pathological conditions seen in humans, correlating the observed *in vitro* affects with the situation *in vivo* has proved to be difficult. Nonetheless, various studies have shown that core can modulate apoptosis (Ruggieri *et al.*, 1997; Marusawa *et al.*, 1999), cell transformation (Ray *et al.*, 1996), transcriptional regulation (Shih *et al.*, 1993 and 1995), immune presentation (Large *et al.*, 1999), lipid metabolism (Moriya *et al.*, 1997) and the development of HCC (Moriya *et al.*, 1998). Supporting evidence is provided by data which show that core can interact directly with some factors involved in these cellular processes, e.g. LT β R and TNFR1 which play roles in the early stages of apoptosis. Although the mechanism by which core operates in these cellular processes is not thoroughly understood, it is suggested that core is functionally a pleiotropic protein and does point to its potential critical role in HCV pathogenesis.

1.3.2. p7

p7 is represented by residues 746 to 810 and is located between E2 and NS2. The cleavage events between the E2/p7 and p7/NS2 are mediated by host signal peptidase, although processing at the E2/p7 site is incomplete in a number of HCV strains (Lin *et al.*, 1994). In pestiviruses, cleavage at a similar position can also be inefficient (Elbers *et al.*, 1996). Analysis of BVDV and CSFV virions suggest that neither p7 nor E2-p7 are major structural constituents (Elbers *et al.*, 1996). Whether this feature is true for HCV p7 and E2-p7 has not yet been elucidated. p7 is mainly hydrophobic but such segments are interrupted by conserved charged residues.

1.3.3. E1 and E2 Glycoproteins

Sequence comparison studies between HCV and members of the *Flaviviridae* family suggested that putative envelope proteins followed core protein (Choo *et al.*, 1991). Studies using recombinant vaccinia viruses to express cDNA clones encompassing the HCV ORF in human hepatoma lines identified two membrane-associated species which were extensively glycosylated (Grakoui *et al.*, 1993b; Ralston *et al.*, 1993). These proteins had apparent molecular weights of approximately 31-35 kDa and 65-70 kDa, and were designated E1 and E2, respectively.

The glycoproteins are released from the polyprotein by the action of cellular proteinases and signal peptidases. The signal sequences that direct these events are located in hydrophobic domains upstream from each of the cleavage sites (Fig 1.13). Analysis of the amino termini of both E1 and E2 indicated that the polyprotein is processed by cleavage between amino acids 191/192 and 383/384 (Fig 1.4; Grakoui *et al.*, 1993b; Selby *et al.*, 1994). Cleavage at these sites generated a single E1 species. However, various E2-specific proteins have been detected which correspond to E2, E2-p7 and E2-p7-NS2 that are produced by cleavage between residues 745/746, 810/811 and 1027/1028, respectively. The E2-p7-NS2 precursor is relatively short-lived and is cleaved to produce either E2 or E2-p7 (Grakoui *et al.*, 1993b; Selby *et al.*, 1994). For some HCV strains, processing at the E2-p7 site is inefficient, leading to the production of a reasonably stable E2-p7 species (Dubuisson *et al.*, 1994; Lin *et al.*, 1994; Mizushima, *et al.*, 1994; Selby *et al.*, 1994).

Sensitivity of transiently-expressed glycoproteins to deglycosylation with endo- β -N-acetylglucosaminidase H_f (endo H_f) confirmed that E1 and E2 are modified by N-linked glycosylation (Dubuisson *et al.*, 1994). For strain H77 (genotype 1a), E1 has 5 predicted glycosylation sites while there are presumed to be 11 in E2 (Fig 1.13); in other strains and genotypes, the number of predicted glycosylation sites varies by up to 3 sites. Two studies have suggested that one of the five predicted glycosylation sites in E1 from strain H is not processed (Fournillier-Jacob *et al.*, 1996; Meunier *et al.*, 1999). From analysis of sequential mutants at each of the predicted sites, the site at residue 325 (Fig 1.13) is not glycosylated (Meunier *et al.*, 1999). The predicted glycosylation site at residue 325 has the amino acid sequence Asn-Trp-Ser-Pro (N-W-S-P). Large, hydrophobic amino acids such as Trp within the core sequence of N-glycosylation sites can generate a locally unfavourable protein conformation or block access of the cellular glycosylation machinery to the Asn residue (Bause, 1983). Moreover, a Pro residue immediately downstream of the core glycosylation sequence also is associated with blocking glycosylation. Alteration of Trp to Thr did not restore glycosylation, however an additional mutation of Pro to Gly overcame the block in processing (Meunier *et al.*, 1999). Thus, the Pro residue immediately downstream from the core glycosylation site at position 325 appears to be the principle cause of the lack of processing at this site. This agrees with other studies on the effect of Pro on N-linked glycosylation (Bause, 1983).

Expression of truncated forms of the glycoproteins, lacking their C-terminal hydrophobic regions promoted secretion of the ectodomains (Matsuura *et al.*, 1994; Selby *et al.*, 1994; Hussy *et al.*, 1996). This implied that both species are type 1 transmembrane glycoproteins, anchored by their C-terminal regions. Predictions indicate that for E2, the hydrophobic anchor domain begins at residue 713 (Fig 1.13; Mizushima *et al.*, 1994). Within E1, the situation is slightly more complicated since two hydrophobic segments exist (residues 269-286 and 342-378). Expression of E1 truncated at residue 340 by recombinant baculoviruses in insect cells required also the removal of amino acids 262-290 for secretion (Matsuura *et al.*, 1994). By contrast, Hussy *et al.* (1996) provided evidence that E1 truncated at residue 334, and which therefore retained the N-terminal hydrophobic sequences, was efficiently secreted from insect cells using the same expression system. In mammalian cells, secretion of E1 was observed on truncation at residue 311 but not at residues 330, 354 and 360 (Michalak *et al.*, 1997). The C-terminal ends of E1 and E2 are thought to provide bi-functional roles. Firstly, they act as signal sequence elements to

direct E2 and perhaps p7 respectively to the ER. Secondly, signals for ER retention also have been mapped to these regions of E1 and E2 (see Section 1.3.9).

In immune responses to HCV infection, the envelope proteins are presumed to be the primary viral antigens encountered since they exist on the surface of the virus particle. As a possible consequence of this, E1 and E2 have evolved to contain regions susceptible to sequence variation, which span approximately aa 210-223 and 315-327 in E1 and residues 384-411 and 474-481 in E2 (Fig 1.13; Kato *et al.*, 1992; Lesniewski *et al.*, 1993). The most variable region in the entire HCV genome is contained within region 384-411, which is termed HVR-1 (Weiner *et al.*, 1991). However, sequence alignments do reveal that certain amino acid positions in HVR-1 are less variable e.g. threonine at position 385, and glycine at positions 389 and 406 (Puntoriero *et al.*, 1998). Both B and T cell epitopes have been identified in HVR-1 that elicit neutralising antibodies (Rosa *et al.*, 1996). Moreover, anti-HRV-1 antibodies from human sera display a degree of cross-reactivity to different HVR-1 variants, suggesting that the conservation pattern observed could be responsible for this cross-reactivity (Puntoriero *et al.*, 1998).

1.3.5. HCV Glycoprotein Complexes

Using viral and non-viral expression systems, including transgenic mice, several groups have shown that E1 and E2 interact to form complexes (Grakoui *et al.*, 1993b; Lanford *et al.*, 1993; Ralston *et al.*, 1993; Dubuisson *et al.*, 1994; Koike *et al.*, 1995; Deleersnyder *et al.*, 1997). Two early attempts to characterise the complexes were contradictory in terms of the nature of the interaction that occurred between the two glycoproteins. Both studies employed vaccinia virus recombinants to express E1 and E2 in mammalian cells. One group suggested that the two proteins associated non-covalently (Ralston *et al.*, 1993), while the second reported that the interactions were dependent on intermolecular disulphide bonds (Grakoui *et al.*, 1993b). Subsequent investigations showed that both forms of the complex could be detected in the same cell extract following an extended incubation period using either recombinant vaccinia or adenovirus expression systems (Dubuisson *et al.*, 1994; Dubuisson & Rice, 1996; Young *et al.*, 1998). Temporally, the first type of complex detected contain heterogeneous disulphide-linked high molecular weight material, while heterodimers of E1 and E2 stabilised by non-covalent interactions are identified at later times.

Procedures employed to distinguish between the two forms of complex include non-reducing gel electrophoresis conditions and conformation-specific antibodies. Non-covalently-linked complexes can be distinguished from those that are disulphide-linked by separating purified proteins by SDS-PAGE in the absence of reducing agents such as β -mercaptoethanol (β -ME) or DTT (Hames, 1990). The covalently-linked proteins are retarded at the top of gels while the proteins in non-covalently linked complexes migrate into gels as monomers. Deleersnyder *et al.* (1997) also provided evidence for a conformation-sensitive E2-reactive monoclonal antibody (H2) which selectively recognises E1E2 heterodimers that are non-covalently linked. The proteins detected by monoclonal antibody (MAb) H2, when analysed under non-reducing conditions, migrate as monomers. More recently, another antibody (H53) from the same group displays similar properties (Cocquerel *et al.*, 1998). Other antibodies raised against E1 and E2 (e.g. A4 and A11) are less able to discriminate between the two types of complex (Dubuisson *et al.*, 1994; Deleersnyder *et al.*, 1997). This includes human monoclonal antibodies which recognise both non-covalently-linked and disulphide-linked complexes suggesting that both forms of complex could occur *in vivo* (Siemoneit *et al.*, 1995; Cardoso *et al.*, 1998; Habersetzer *et al.*, 1998).

For the purposes of this thesis, complexes of E1 and E2 that dissociate and migrate as monomers under non-reducing conditions and are therefore non-covalently-linked will be referred to as the “native”. This does not imply that such complexes are present on the virus envelope or that they play a specific role in virion morphogenesis. The second type of complex, in which E1 and E2 fail to migrate into gels in the absence of any reducing agent and are therefore linked by intermolecular disulphide bonds, will be referred to as “aggregated” complexes.

1.3.6. Properties of Native E1E2 Complexes

For the prototype strain HCV-1, immunopurified native E1E2 complexes sediment at approximately 15 S on glycerol density gradients which correspond to an apparent molecular weight of 310-320 kDa, indicative of a trimeric complex formed by the non-covalently-attached heterodimers (Ralston *et al.*, 1993). Another report indicated that agarose column-purified native E1E2 heterodimers formed dimeric and trimeric complexes by non-covalent interactions (Young *et al.*, 1998). By contrast, complexes co-precipitated

with anti-calnexin antibodies (an ER resident chaperone) and with MAb H2 revealed that oligomers sediment between molecular mass standards of 68-158 kDa (Deleersnyder *et al.*, 1997). These differences may result from the alternative methods used for purification.

E2 and E2-p7 proteins have been detected in non-covalently linked complexes, implying that the presence or absence of p7 does not hinder the formation of a native complex (Dubuisson *et al.*, 1994; Lin *et al.*, 1994). For the BK strain of HCV (genotype 1b), processing at the E2/p7 site is efficient and leads to the cleavage of the majority of E2-p7 species, suggesting that, at least for this strain, unprocessed E2-p7 would not be incorporated into HCV particles (Dubuisson *et al.*, 1994; Lin *et al.*, 1994). In the case of BVDV, the prototype pestivirus, processing in this region of the polyprotein also occurs inefficiently but E2-p7 is either not present in mature virus particles or at levels too low to be detected (Elbers *et al.*, 1996). Pulse labelling experiments, followed by immunoprecipitation using conformation-sensitive MAb H2, have revealed that the kinetics of native complex formation are slow, requiring a chase period of at least 80 min before reaching detectable levels (Deleersnyder *et al.*, 1997). For HCV glycoproteins expressed separately, conformation-sensitive MAb H2 does recognise E2 but not the E2 precursor, E2-p7-NS2 (Deleersnyder *et al.*, 1997).

Despite the high degree of sequence variability observed for E1 and E2, the glycoproteins contain highly conserved cysteine residues (8 and 20, respectively), throughout the genotypes (Robertson *et al.*, 1998). Like other cellular and viral glycoproteins, these residues are involved in the formation of intra- or intermolecular disulphide bonds, essential for some proteins to attain and stabilise their correctly folded state (Sections 1.4.2 and 1.4.3; Creighton, 1988). The kinetics of intramolecular disulphide bond formation has been monitored for E1 and E2 by analysing pulse-labelled proteins under non-reducing SDS-PAGE conditions (Dubuisson & Rice, 1996). This provides the means to distinguish between denatured, oxidised proteins, which have greater electrophoretic mobility than their denatured, reduced counterparts; oxidised proteins adopt a more compact conformation due to the presence of intramolecular disulphide bonds (Braakman *et al.*, 1992a and b; Hammond *et al.*, 1994). For a given protein, this increase in mobility over time can also reflect progressive folding towards the native state of the protein. Disulphide bond formation occurs rapidly in E2 (by ~15min) and precedes processing of precursor E2-NS2 (Dubuisson & Rice, 1996). However, the electrophoretic mobility of E1 increases

only after longer periods of chase (> 60min), suggesting that oxidation of E1 is slow. This step may be rate limiting for native complex formation (Dubuisson & Rice, 1996).

1.3.7. Properties of Aggregated E1E2 Complexes

Purified E1 and E2 proteins subjected to sucrose density gradient sedimentation and separation of gradient fractions by non-reducing SDS-PAGE revealed that aggregates may have a molecular mass greater than 158 kDa (Young *et al.*, 1998). The tendency of individually expressed E1 and E2 to form aggregates suggests that disulphide bonds may occur via homologous covalent interactions between E1-E1 and E2-E2 as well as heterologously between E1E2 (Michalak *et al.*, 1997). The non-physiologically high level of protein production achieved by viral expression systems, especially of glycoproteins which are processed co- and post-translationally in the ER, could promote abnormal aggregation. However, analysis of HCV glycoprotein assembly in cell lines where the level of expression is controlled by the tetracycline-regulated system yield similar results (Duvet *et al.*, 1998; Moradpour *et al.*, 1998). Thus, it is possible that aggregation may be an intrinsic property of HCV glycoproteins. The data derived from tissue culture systems could apply to characteristics of natural HCV infection. With regards to virus replication, inefficient formation of HCV glycoprotein complexes could down-regulate particle formation and virus replication to minimise exposure of viral antigens to the immune system. In conclusion, it is proposed that interactions between the HCV glycoproteins follow two distinct folding pathways, leading to the formation of either non-covalent E1E2 complexes or aggregates of the two proteins.

1.3.8. Requirements for E1E2 Complex Formation

From initial studies, truncated forms of E2 (truncated at aa 661, 699, 701, and 715) fail to co-precipitate co-expressed E1 (Selby *et al.*, 1994). By contrast, E2 species terminating at aa 730 or beyond are capable of co-precipitating equivalent levels of E1. Thus, sequences between aa 715-730 in E2 appear to be important for efficient E1 association as well as anchorage of E2 to membranes (Sections 1.3.3 and 1.3.9; Selby *et al.*, 1994). However, co-expression of truncated forms of both proteins (E1₃₁₁ and E2₆₆₁) from separate transcription units reveal that complexes formed between the two proteins are released into the supernatant of culture medium (Lanford *et al.*, 1993; Matsuura *et al.*, 1994; Hussy *et*

al., 1996; Michalak *et al.*, 1997). This suggests that additional sequences important for HCV glycoprotein interaction are located in their ectodomains. Analysis of truncated forms of E2 under non-reducing SDS-PAGE conditions showed that the majority of these secreted proteins form aggregates, although a small proportion of E2₆₆₁ alone migrates into a non-reducing gel as monomers (Michalak *et al.*, 1997). Furthermore, the conformation-specific antibody H2 fails to recognise all forms of secreted E2 except for truncated E2₆₆₁ protein (Michalak *et al.*, 1997).

Far-Western blotting studies and *in vitro* binding assays using N-terminal truncated and internal deletion mutant E2 proteins derived from bacterial expression systems showed that the highly variable regions in E2 are not necessary for interaction (Yi *et al.*, 1997a). The major E1-interacting site of E2 maps within the N-terminal part of E2 (aa 415 to 500 of the polyprotein). Based on the assumption that a degree of sequence conservation was required in regions involved in complex formation, Yi *et al.* (1997a) implicated a conserved region within the N-terminal region of E2, termed the Y box (aa 484-491), as important for E1 association. This region contains amino acid residues Trp-His-Tyr (aa 489-491), which, upon mutation, markedly impair E1 binding. In the same report, the N-terminal part of E1 (192-238) was implicated in E2 binding. However, it should be noted that the assays utilised would not distinguish between native and aggregated complexes.

As described in Section 1.3.3, one predicted glycosylation site (at residue 325) in E1 from strain H is not utilised (Meunier *et al.*, 1999). These authors analysed also the effect of mutation at each of the other sites (at residues 196, 209, 234 and 305) on complex formation. Using a conformation-sensitive MAb, mutations affecting positions 2 and 3 has minor effects on the assembly of the E1E2 complex. However, mutation at position 1 and predominantly at position 4 dramatically reduce the efficiency of non-covalently linked complex formation. The absence of one glycan addition site at positions 1, 2, 3 and 4 did not reduce interaction between E1 and calnexin (discussed further in Sections 1.3.11.2 and 1.4.4), which suggests that the lower level of non-covalent complex observed for mutants is not due to an alteration in their association with this chaperone.

1.3.9. Subcellular Localisation of HCV Glycoproteins

The failure to detect HCV glycoproteins on the cell surface, the lack of complex glycans, and sensitivity to endo H_f deglycosylation indicate that these proteins do not leave the ER

(Spaete *et al.*, 1992; Ralston *et al.*, 1993; Dubuisson *et al.*, 1994; Deleersnyder *et al.*, 1997). From immunostaining experiments, intracellular mature complexes co-localise with protein disulphide-isomerase (PDI), an ER resident marker (Duvet *et al.*, 1998). The absence of complex forms of glycosylation on HCV glycoproteins excludes their transit through the medial Golgi (see Fig 1.14; Duvet *et al.*, 1998). It is possible that the glycoproteins could cycle between the ER and the *cis* Golgi and not reach the medial compartment. Under such circumstances, oligosaccharides on glycoproteins transported from the ER to the *cis* Golgi would be exposed to Golgi α -mannosidase I. This enzyme processes high mannose ($\text{Man}_9\text{GlcNAc}_2$) sugar chains into $\text{Man}_5\text{GlcNAc}_2$ (Kornfeld & Kornfeld, 1985), which renders the proteins sensitive to endo D as they accumulate in the ER after several cycles through the *cis* Golgi. However, E1 and E2 remain endo D-resistant, thus suggesting that they do not cycle between the ER and the *cis* Golgi compartment (Duvet *et al.*, 1998). HPLC analysis of glycans released by PNGase F, which removes all N-linked glycans, and endo H_F treatment demonstrated that the oligosaccharide moieties released by both enzymes were similar oligomannoside species (Man_9 , Man_8 , and $\text{Man}_7\text{GlcNAc}_{1/2}$; see Fig 1.14). To examine whether the N-glycosylation sites could be modified in the presence of Golgi mannosidases and glycosyltransferases, oligosaccharides obtained from brefeldin A- (BFA-)treated cells were analysed by affinity chromatography. Twenty-four percent of glycans were found to be processed into complex types, while 75% remained of oligomannoside type, which indicates that the absence of processing of HCV glycoprotein glycans is not because of inaccessibility of Golgi oligosaccharide modifying enzymes (Duvet *et al.*, 1998).

Recently, using chimeric proteins in which E2 domains are exchanged with corresponding domains of a protein normally transported to the plasma membrane (CD4), an ER-retention signal has been mapped to the C-terminal 29 amino acids of E2 (Cocquerel *et al.*, 1998). Similarly, the fusion of the C-terminal 31 residues of E1 to the ectodomains of CD4 and CD8 result in localisation of these chimeric proteins to the ER (Cocquerel *et al.*, 1999). The ability of the C-terminal domains of E1 and E2 to act as ER retention signals with heterologous proteins has been confirmed by others (Flint & McKeating, 1999). Although both E1 and E2 contain signals for ER retention, they do not show any sequence similarity apart from stretches of hydrophobic residues. Hereafter, the stretches of hydrophobic residues at the C-terminal ends of E1 and E2 will be referred to as transmembrane domains. Several specific signals have been identified for the retention/retrieval of ER proteins and these are described in Section 1.4.8.

1.3.10. Association of HCV Glycoproteins with Viral Proteins

From immunoprecipitation studies, E1 could co-precipitate with core and detection of this interaction did not require linkage of the two proteins in *cis* (Lo *et al.*, 1996). However, association of core with E2 was not observed. NS2 and NS4A also were shown to associate with E2 through residues 699-730 (Selby *et al.*, 1994). This was confirmed in a separate study, where either anti-E1 or anti-E2 antibodies co-precipitated NS2 and NS3, although the reciprocal experiments using anti-NS2/NS3 antibody yielded negative results (Matsuura *et al.*, 1994). At present, no viral proteins have been shown to interact with native E1E2 complexes, but for viral assembly interaction by the envelope with other viral species may occur.

1.3.11. Association of HCV Glycoproteins with Folding Proteins

1.3.11.1. The Role of E2 in the Folding of E1

Studies evaluating the effect of E1 and E2 on the folding of each other found that the folding kinetics of E2 are similar in the presence and absence of E1 based on the recognition of E2 using conformation-sensitive MAb H2 (Deleersnyder *et al.*, 1997; Michalak *et al.*, 1997). This result led to the conclusion that the presence of E1 is not necessary for the folding of E2. Using intramolecular disulphide bond formation as an assay for E1 folding, the oxidised form of the protein appears slowly when E2 is co-expressed (Dubuisson & Rice, 1996). However, in the absence of E2, no oxidised form of E1 is detected, indicating that E2 affects the folding of E1 (Michalak *et al.*, 1997). It has been proposed that E2 may play a chaperone-like role in E1 folding by interaction directly with E1 (Michalak *et al.*, 1997), and this role may complement the functions performed by ER-resident chaperones. E2 folds rapidly and its ability to achieve a disulphide-dependent conformation may be necessary to confer this chaperone-like function. The structural PrM glycoprotein of the Japanese encephalitis virus (JEV) provides a similar function in the folding of the envelope E protein (Konishi & Mason, 1993). In this case, co-synthesis of PrM is necessary for proper folding, membrane association and assembly of E protein.

1.3.11.2. The Role of Cellular Folding Proteins

In the ER, a significantly large proportion of the soluble proteins represent molecular chaperones such as immunoglobulin heavy-chain binding protein (BiP/GRP78), glucose regulated protein (GRP94), and calreticulin as well as membrane-associated calnexin (Reddy & Corley, 1998). They function by binding to nascent polypeptides in order to prevent aggregation and to correct the folding of proteins trapped in misfolded structures. Calnexin and calreticulin are lectin-like chaperones that interact transiently with glycoproteins during their folding and maturation (Tatu & Helenius, 1997). An additional function of chaperones is to retain incompletely and incorrectly folded proteins in the ER and thus act in concert with quality control mechanisms to prevent transport of misfolded structures to other organelles (e.g. the Golgi apparatus) involved in protein maturation and secretion (Section 1.4).

Calnexin, calreticulin and BiP interact with E1 and E2, whereas no interaction has been detected between GRP94 and HCV glycoproteins (Dubuisson & Rice, 1996; Choukhi *et al.*, 1998). However, calnexin associates mainly with native complexes from data based on pulse-chase experiments. By contrast, calreticulin and BiP were found to preferentially associate with aggregated complexes (Choukhi *et al.*, 1998), as is observed for many misfolded proteins, e.g. vesicular stomatitis virus G protein (VSV G; Doms *et al.*, 1987). Both E1 and E2 expressed alone are co-precipitated with calnexin, calreticulin and BiP (Dubuisson & Rice, 1996; Deleersnyder *et al.*, 1997; Choukhi *et al.*, 1998; Liberman *et al.*, 1999). The association of HCV glycoproteins with calnexin and calreticulin is faster than that with BiP, and the kinetics of interaction with calnexin and calreticulin are similar. The maximum amounts of E2 that co-precipitate with calnexin and calreticulin reach about 75% of the total recombinant E2 while ~70% is achieved for binding of E2 to BiP (Choukhi *et al.*, 1998). For E1, maximum amounts co-precipitated with calnexin, calreticulin and BiP are ~40, 20, and 10%, respectively. Over-expression of calnexin, calreticulin or BiP with HCV glycoproteins does not improve the proportion of non-covalent HCV glycoprotein complexes assembled (Choukhi *et al.*, 1998).

Apart from chaperones, the ER also contains a high level of other folding proteins such as PDI (see Section 1.4.3; Doms *et al.*, 1993). PDI foldase catalyses the shuffling of inappropriately formed disulphide bonds (discussed further in Section 1.4.3; Doms *et al.*, 1993; Reddy & Corley, 1998). Due to the high conservation of cysteine residues in HCV

glycoproteins, PDI-like foldases are likely to be involved in the formation of covalent disulphide bonds in E1 and E2. However, no direct evidence for this has been observed.

1.3.12. Interactions of HCV Glycoproteins with Other Cellular Proteins and Process

1.3.12.1. Effects on PKR

In vitro assays have shown that HCV E2 associates with the cellular regulatory factor PKR (Section 1.2.8; Taylor *et al.*, 1999). PKR is activated by interferon and mediates inhibition of protein synthesis by phosphorylation of the translation initiation factor eIF2 α . It is also capable of autophosphorylation. Thus, PKR is an intrinsic component of the antiviral response to infection. A segment has been identified in E2 between residues 276 and 287 which has homology to the sites in eIF2 α and PKR that are phosphorylated by PKR. However, there is no evidence for phosphorylation of this site in E2. Thus it is proposed that E2 may act as a pseudosubstrate for PKR and thereby block the inhibitory effects of PKR on protein synthesis and cell growth. The proposed homology between E2 and eIF2 α /PKR exists in genotypes 1a/1b (which display interferon resistance) but not in genotypes 2a/2b/3a (which have greater sensitivity to interferon). Inhibition of PKR leads to promotion of cell growth and inactivation of apoptosis signalling pathways; such a mechanism may contribute to HCV-associated HCC (Taylor *et al.*, 1999).

1.3.12.2 E2 Binding to Lactoferrin

Using both Far-Western blot analysis and *in vitro* binding assays, E1 and E2 bind to lactoferrin (Yi *et al.*, 1997). There was evidence also that a truncated form of E2 secreted from cells could associate with lactoferrin. This cellular protein is a member of the transferrin family of iron-binding glycoproteins and is present mainly in breast milk as well as other exocrine secretions. It was proposed that association of lactoferrin with extracellular HCV may interfere with the infectivity of the virus. Since lactoferrin is present in high concentrations in breast milk and in plasma during pregnancy, this could explain the low level of transmission of HCV from mother to baby.

1.3.12.3. Activation of GRP Protein Synthesis

In a recent study, expression of E2 but not E1 activated transcription from the promoters of two chaperones, GRP78 (BiP) and GRP94, and induced elevated levels of intracellular GRP78 protein (Lieberman *et al.*, 1999). This activation was dependent on targeting of E2 to the ER. Interestingly, E2 was bound stably to GRP78 but this was not the case for E1. It is proposed that accumulation of E2 in the ER may induce a stress response in the ER which acts to stimulate levels of certain chaperone proteins (discussed in Section 1.4.9).

1.3.13. Cellular Receptors

Using flow cytometry, a truncated form of E2 (truncated at residue 715), expressed in and isolated from mammalian but not insect or yeast cells, could bind to MOLT-4 cells with high affinity ($K_d \approx 10^{-8}$ M; Shimizu *et al.*, 1992; Rosa *et al.*, 1996). These cells are a human cell line reported to allow low-level HCV replication *in vitro*. Subsequently, a cDNA library prepared from this cell line was screened by transient transfection in a mouse-derived cell line using truncated E2 protein as a probe (Pileri *et al.*, 1998). A cDNA clone containing the coding sequence of a cellular receptor, CD81, was identified that conferred E2-binding capacity to the mouse cells. CD81 is a widely expressed 25 kDa molecule belonging to the tetraspanin superfamily that span the cell surface membrane four times and form two extracellular loops. The interaction between E2 and CD81 has been confirmed using several biochemical methods including immunoprecipitation, immunoblotting and fluorescence-activated cell sorting (FACS; Pileri *et al.*, 1998; Flint *et al.*, 1999a and b). Binding of virus particles, derived from infectious human sera, to CD81 was prevented also in the presence of anti-E2 antibodies (Pileri *et al.*, 1998). E2 associates with the extracellular loop 2 (EC2) region of CD81, although multiple regions in E2 are found to be necessary for interaction with EC2 (Flint *et al.*, 1999b; Pileri *et al.*, 1998). Binding of CD81 with E2 induces aggregation of lymphoid cells in addition to inhibiting B-cell proliferation, suggesting that E2 associated with CD81 could modulate cell function (Flint *et al.*, 1999a).

In light of the presence of CD81 on a wide variety of cell types, an alternative study has presented evidence for HCV binding to COS-7 cells transiently expressing the human LDL receptor (LDLR; Monazahian *et al.*, 1999). RNA-carrying material or particles isolated

from different groups of patients are heterogeneous with respect to their buoyant density and sedimentation. This is probably due to the binding of these particles to β -lipoproteins (low density lipoprotein [LDL], very low density lipoprotein [VLDL]) and immunoglobulins (IgG/IgM; Prince *et al.*, 1996). It is proposed that HCV and LDL may compete directly for the cellular LDLR and that LDL in sera of patients may regulate the binding of HCV to this potential HCV receptor (Monazahian *et al.*, 1999). Alternatively, LDL, already bound to HCV particles in serum, may act as a ligand for attachment of the virus to cells expressing LDLR on their surface.

1.3.14. Assembly and Release of HCV Particles

Due to the absence of an efficient *in vitro* virus propagation system, the detection of low levels of viral particles in infected individuals and the lack of a convenient animal model, the factors necessary for HCV particle assembly have not been clearly defined. By comparison with the related flaviviruses, hepatitis C virion morphogenesis may occur by budding into intracellular vesicles. The lack of complex glycans, the ER localisation of expressed HCV glycoproteins, and the absence of these proteins on the cell surface would support this model. To date, efficient particle formation and release has not been observed in mammalian transient expression assays, suggesting that an appropriate intracellular environment has not been identified or that certain viral factors do not function correctly. The properties of the HCV particles may reflect those of pestiviruses, for which virion release is inefficient. In the case of CSFV, a substantial fraction of the virus remains cell associated and the virions are extremely fragile (Laude, 1977). If this were the case for HCV, intracellular sequestration of HCV glycoproteins and virion formation could play a role in the establishment and maintenance of chronic infections by avoiding immune surveillance and preventing lysis of virus-infected cells via antibody and complement. Nonetheless, using recombinant baculovirus containing the cDNA of HCV structural proteins, HCV-like particles have been purified from insect cells (Baumert *et al.*, 1998). The cDNA was designed to encode the capsid protein (core) as well as the envelope proteins (E1 and E2), although the sequence for p7 was not included. From this study, the structural proteins assembled into enveloped virus-like particles of between 40 and 60 nm in diameter in large cytoplasmic cisternae. This organelle was presumed to have evolved from the ER. In addition, HCV RNAs were demonstrated to selectively associate with these particles rather than non-HCV transcripts.

1.4. Properties of the Endoplasmic Reticulum

The ER is a major site of cellular protein synthesis and is subdivided into three distinct but continuous subcompartments; the rough ER (RER), smooth ER (SER) and the nuclear envelope. The RER can be distinguished from the SER by the ribosomes that coat regions of the RER giving it a characteristic speckled appearance and hence its name (Lee & Chen, 1988). Microsomes generated by homogenisation of cells have demonstrated that the RER contains many proteins not present in SER, which may help bind ribosomes and assist in producing the flattened shape of this region of the ER (Gilmore, 1991). Proteins destined for the ER are directed to the organelle during their synthesis. The processes of translation and translocation across the ER membrane are coincident, thus avoiding contact of the nascent peptide chain with the cytosol (Rapoport, 1992). Importantly, the newly synthesised proteins are monitored for correct folding before they proceed along the secretory pathway. Besides protein production, the ER is also the major site for lipid biosynthesis (Bishop & Bell, 1988).

The SER is the site for production of vesicles carrying newly synthesised proteins and lipids to the Golgi apparatus. Some specialised cells have extended SER for additional functions such as synthesis of steroid hormones (Mori & Christensen, 1980). For example, hepatocytes, the major cell type in the liver are a major site for the production of lipoprotein particles. These particles carry lipids via the bloodstream to other organs in the body (Pease & Leiper, 1996). The enzymes that synthesise the lipid components of lipoproteins are located in the membrane of the SER.

1.4.1. Signal Peptide Recognition and Translocation

Targeting of ribosomes to the RER is initiated following the synthesis of the first hydrophobic segment of a nascent polypeptide (Blobel & Dobberstein, 1975). *In vitro* translation of proteins which would normally be secreted from the cell are slightly larger when synthesised in the absence of microsomes as compared to those made in the presence of microsomes derived from RER; the observed size difference is due to the presence of a N-terminal leader peptide. This serves as a signal peptide for directing membrane and secreted proteins to the ER before removal from the peptide chain by cleavage by signal peptidases in the ER membrane. Cleavage to remove the signal peptide sequence occurs

during protein synthesis. Signal sequences are not highly conserved in their primary sequence but are defined by their hydrophobic nature (High & Dobberstein, 1992). Moreover they are not necessarily present at the N-terminus of nascent chains. Indeed, for transmembrane proteins containing an internal signal peptide, more positively-charged amino acids immediately preceding the hydrophobic core can determine the orientation of the protein with respect to the membrane (High & Dobberstein, 1992).

The signal peptide is recognised by the SRP as it emerges from the ribosome, which mediates the targeting of a ribosome/nascent chain/SRP complex to the ER membrane in a GTP-dependent manner (Meyer *et al.*, 1982). This complex docks to the ER membrane via the SRP receptor, an integral membrane protein (Fig 1.15). The nascent chain in the complex is transferred into an aqueous translocation channel, the translocon, which is rapidly sealed from the cytosolic environment by a tight ribosome-membrane junction. Critical components of the translocon include the heterotrimeric Sec61 complex and translocating-chain associated membrane protein (TRAM; Hedge & Lingappa, 1997). In addition to these, signal peptidase (a complex of 5 proteins) and oligosaccharyl transferase (a 3-protein complex), important for the maturation of nascent chains, interact with the nascent translocating polypeptide on the luminal side of the ER (Hedge & Lingappa, 1997). Other luminal proteins, (e.g. BiP, GRP94, calnexin, PDI) associate also with nascent chain substrates. These proteins act as molecular chaperones to promote proper folding and assembly. Although they are crucial for proper protein maturation, at present their roles in the translocation process are not fully understood (Hedge & Lingappa, 1997).

1.4.2. The Luminal Environment of the ER

Proteins, translocated into the lumen of the ER, begin to fold before translation is terminated, although folding to a final conformational state may continue for varying periods of time for different proteins. The folding rate is dependent on the complexity of the protein, the physiological state of the cell and temperature but other factors such as cell type and expression levels also are important. Post-translational folding can be monitored by the appearance of conformation-dependent antigenic epitopes, acquisition of enzymatic or receptor functions, and physical features such as increased protease resistance (Doms *et al.*, 1993).

For transmembrane proteins, folding occurs in three topologically and biochemically distinct environments; the ER lumen, the ER membrane, and the cytosol (Doms *et al.*, 1993). The transmembrane domain adopts an α -helical configuration within the hydrophobic interior of the bilayer, while the cytoplasmic domain follows the parameters that apply for folding of cytoplasmic proteins. The ectodomain folds in the ER lumen and can be co-translationally modified by acylation, glycosylation, proline isomerisation, and disulphide bond formation. Although the information required for a protein to attain its final three-dimensional structure resides in its primary sequence, the folding process is critically dependent on the presence of chaperones and folding enzymes. These proteins increase the efficiency of folding of any nascent chain by preventing irreversible aggregation and catalyse rate-limiting steps such as disulphide bond formation (Freedman *et al.*, 1994) and proline isomerisation (Lang *et al.*, 1987).

Other factors in the ER also facilitate optimal conditions for folding and oxidation of the translocated proteins. Among the properties that distinguish the ER from the cytosol are the high concentration of calcium and the reduction-oxidation (redox) potential that is sufficiently oxidising to allow disulphide bond formation. Calcium ions (Ca^{2+}) are sequestered in the ER from the cytosol and storage in the ER lumen is facilitated by a high concentration of Ca^{2+} -binding proteins (e.g. calreticulin and PDI; Helenius *et al.*, 1997) which require this ion for their function. The disulphide redox potential of the ER is buffered by glutathione, which exists in thiol (GSH) and disulphide (GSSG) forms. In the cytoplasm of most cells, glutathione is present at a total concentration of 1-10mM, and only 1% of this is in the GSSG form. The environment within the ER is considerably more oxidising than the cytoplasm. This is achieved by maintaining the oxidised form of glutathione at ratios approximating to 1.1 to 3.1 for GSH:GSSG (Hwang *et al.*, 1992) through a glutathione transporter in the ER membrane (Ferrari & Söling, 1999). The formation of one disulphide bond in a protein requires two sequential thiol-disulphide exchanges involving the mixed-disulphide intermediate (Fig 1.16; Creighton, 1988). Therefore, disulphide links in proteins only occur if the protein conformation brings pairs of Cys residues into suitable proximity. Despite the oxidising environment, shuffling of disulphide bonds is permitted in the ER and disulphide links only form between Cys residues that are favoured by protein conformation (Creighton, 1988).

1.4.3. Isomerases Present in the ER

Folding enzymes and molecular chaperones are the most abundant proteins in the ER and, at nearly millimolar concentrations, the majority of them greatly outnumber the substrate proteins (Tatu & Helenius, 1997). Among the folding enzymes, protein disulphide isomerase (PDI) and members of the PDI family are the most important and best understood. PDI carries a KDEL ER retention signal (Section 1.4.8), and contains a C-X-X-C domain that is homologous to thioredoxin, a ubiquitous small protein active in cytoplasmic redox reactions (Freedman *et al.*, 1994). The function of PDI is dependent on the disulphide redox potential of the ER, reflected in the ratio of reduced to oxidised glutathione (see previous section). The active site of PDI is estimated to be about 50% reduced under these conditions. Apart from glutathione, other oxidising equivalents are believed to enter the ER via electron transport machinery. For example, Ero 1p, a yeast protein, can directly oxidise PDI, avoiding the need for glutathione (Frand & Kaiser, 1998 and 1999; Pollard *et al.*, 1998). Thus, mechanisms may exist for the direct oxidation of GSH, of nascent proteins or to re-oxidise the PDI family proteins that introduce disulphide bonds into nascent chains (Rietsch & Beckwith, 1998). The main role of PDI is to catalyse disulphide-bond formation in nascent proteins through glutathione-protein mixed-disulphide species, by acting as an isomerase independent of re-oxidation of GSSG (Ferrari & Söling, 1999). The redox/isomerisation activities of PDI enable isomerisation reactions to proceed at several thousand times the uncatalysed rate (Fig 1.17).

Peptidyl proline isomerases (PPI) are also considered to be folding enzymes. They catalyse the slow *cis-trans* isomerisation of proline peptide (X-P) bonds in oligopeptides and accelerate slow, rate-limiting steps in protein folding (Fischer *et al.*, 1989). However, no effect of cyclosporin A, an inhibitor of PPI, on the folding of influenza haemagglutinin (HA) is observed (Doms *et al.*, 1993). Therefore, it is not known if this enzyme is important for folding of certain viral glycoproteins.

1.4.4. Chaperones Present in the ER

The most widely studied molecular chaperones in the ER include BiP, calnexin and calreticulin, in addition to GRP94 and Erp72 (Hammond & Helenius, 1995; Helenius *et al.*, 1997). A chaperone is a protein that can assist unfolded or incorrectly folded protein to

attain the native state by providing a micro-environment in which competing aggregation reactions are reduced (Eisenberg, 1999). Chaperones are capable also of mediating the reversal of pathways that lead to incorrectly-folded structures. As proteins fold, intermediates may be transferred to different molecular chaperones. Moreover, the sites with which different chaperones interact may only be transiently exposed during the folding process. These events contribute to the quality control function of the ER. Chaperones also can substitute for each other, thus ensuring the efficient retention of misfolded proteins and a backup system for ER quality control. Sequential transfer of distinct folding intermediates to molecular chaperones is not pre-determined by the ER, and the order of events is dictated by the primary amino acid sequence of the translocated protein (Reddy & Corley, 1998).

The soluble luminal protein, BiP, was initially identified in complexes with immunoglobulin heavy chains, which are retained in the ER until they become associated with light chains (Doms *et al.*, 1993). Generally, BiP associates transiently and non-covalently via exposed hydrophobic patches in nascent proteins and dissociates during the folding process. Synthesis of this protein is induced by increased levels of secretory membrane and misfolded proteins, thus providing indirect evidence for its function (Sidrauski *et al.*, 1998; Liberman *et al.*, 1999). Like most resident proteins it contains the ER retrieval motif, KDEL, which enables BiP, associated with nascent protein, to remain in the ER during folding (Reddy & Corley, 1998).

During folding, polypeptides with N-linked oligosaccharides interact transiently and specifically with two homologous lectin-like chaperones that are unique to the ER, membrane-bound calnexin and soluble calreticulin. Both are not glycosylated proteins and bind specifically to partially trimmed, monoglucosylated forms of N-linked core glycans ($\text{Glc}_1\text{Man}_{5-9}\text{GlcNAc}_2$) present on folding intermediates. They do not bind glycans with two or more glucose residues (Fig 1.18; Helenius *et al.*, 1997; Tatu & Helenius, 1997). For these chaperones, oligosaccharide recognition is thought to be the first step prior to binding to the protein moieties of glycoproteins (Parodi, 1999). Glycoprotein folding intermediates are retained in the ER by a cycle of glucose removal and addition coupled to chaperone binding and release (Helenius *et al.*, 1997; Choukhi *et al.*, 1998). An example is provided by influenza HA (Helenius *et al.*, 1997), where cycling of the molecules on and off the chaperones is important for proper folding. Although the impact on protein maturation by calnexin and calreticulin is similar to classically described chaperones, they

have distinctive features. For example, they do not directly distinguish between folded and unfolded proteins, but instead depend on covalent modification in the carbohydrate moieties of glycoproteins, catalysed by glycosidase II and glucosyltransferase, to assert their chaperone function (Figs 1.18 and 1.19; Helenius *et al.*, 1997).

One effect believed to be induced by calnexin and calreticulin binding to folding glycoproteins is the immobilisation of specific cysteine residues, thus preventing incorrect disulphide pairing. Recently, specific interaction of a member of the PDI family, ERp57, with soluble secretory proteins bearing N-linked carbohydrates has been demonstrated (Elliott *et al.*, 1997). Although it is suggested that ERp57 is a cysteine-dependent proteinase and a thiol-dependent reductase, its precise function remains to be established. However, consistent with its thiol-dependent reductase activity, similar to PDI, ERp57 may act in concert with calnexin and calreticulin to modulate glycoprotein folding, and thereby facilitate the quality control mechanisms operating in the ER (Elliott *et al.*, 1997).

1.4.5. The Role of Covalent Modifications in Protein Folding

The common forms of covalent modification that occur in the ER are co-translational addition of N-linked glycosylation and post-translational exchange of the C-terminal transmembrane tail for a glycosylphosphatidylinositol (GPI) anchor. The latter modification involves the simultaneous cleavage of the transmembrane domain and creation of a covalent bond between the GPI anchor and the ectodomain of the protein (Brown, 1992).

N-linked glycans are added to growing polypeptide chains as 14-residue oligosaccharides ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$) by oligosaccharyltransferase, which is associated with the translocation machinery. The precursor oligosaccharides are partially assembled and attached to dolichol lipid on the cytosolic side of the ER membrane. They are subsequently “flipped” across the membrane to the lumen of the ER, where the addition of glucose residues is achieved. These precursor molecules are transferred to the side chain of asparagine residues in the consensus sequence N-X-S/T, where X is any amino acid except proline. Transfer of the oligosaccharide generally occurs co-translationally as soon as the acceptor site enters the lumen of the ER. Loss of glycosylation sites on proteins causes defects in folding, which in turn leads to aggregation, misfolding, and disulphide cross-linked complexes that are retained in the ER (Doms *et al.*, 1993). This suggests that

large, hydrophilic oligosaccharide core units are needed for nascent chains to fold properly. In the case of native proteins, their hydrophobic regions are normally buried within the protein, leaving the hydrophilic residues exposed on the surface to allow solubility in the aqueous environment of the cell. However, folding intermediates are generally more hydrophobic than fully folded proteins, and the function of carbohydrate moieties is proposed to render these intermediates more soluble and less likely to form irreversible aggregates. Hence, this modification is considered to occur prior to folding. The importance of individual N-linked carbohydrate moieties within a given protein with respect to folding varies, and there appears to be a hierarchy of glycosylation sites in terms of folding and stability (Doms *et al.*, 1993). Moreover, addition of new N-linked sites is able to compensate for the loss of others or rescue proteins, which would otherwise be retained. This has been shown to be the case for influenza virus HA and VSV G proteins (Machamer & Rose, 1988a and b; Gallagher *et al.*, 1992). Thus, the main role of carbohydrates is to provide a global function to increase solubility, rather than a localised function in the folding process.

Almost immediately following the transfer of oligosaccharides to the nascent proteins, the carbohydrate moiety is modified. Sequential removal of three glucose residues present on the core glycan is performed by glucosidase I (for the first glucose residue) and glucosidase II (for the remaining two glucose residues) to yield $\text{Man}_9\text{GlcNAc}_2$ glycans (Fig 1.19). The monoglucosylated glycan ($\text{Glc}_1\text{Man}_{7-9}\text{GlcNAc}_2$) intermediate, which is formed during the stepwise removal of glucose residues, binds to calnexin and calreticulin (see Fig 1.18; Helenius *et al.*, 1997). In addition, monoglucosylated glycans also result from re-glucosylation of fully glucose-trimmed oligosaccharides by action of the ER luminal enzyme, UDP-Glc:glycoprotein glucosyltransferase (GT; Parodi, 1999). The quality control mechanism for glycoproteins involves binding of monoglucosylated species to calnexin/calreticulin. While associated to these chaperones the final glucose residue on the glycans is removed by glucosidase II, thus liberating the glycoprotein from the chaperones. If the proteins are properly folded they are transported to the *cis* Golgi. However, improperly folded glycoproteins are recognised by GT and re-glucosylated, to allow re-association to calnexin/calreticulin (Parodi, 1999).

GT is active in millimolar concentrations of Ca^{2+} ions, and these levels are readily attained in the ER. The enzyme has the ability to glucosylate denatured (misfolded) glycoproteins, but not native proteins. This is achieved by the recognition of hydrophobic patches

exposed on misfolded proteins that are typically buried in native proteins. Additional experiments using endo H_f revealed that GT recognised not only protein domains exposed in misfolded conformations, but also the innermost GlcNAc unit of oligosaccharides (Fig 1.14). This suggests that in native glycoproteins, recognition of the innermost GlcNAc unit by GT would not occur and thereby glucosylation of proteins in this state would be prevented. Since glucosylation/de-glucosylation cycles continue until glycoproteins are properly folded, GT appears to function as a sensor of the tertiary structure of glycoproteins (Parodi, 1999). Finally, high mannose oligosaccharide moieties are further trimmed by ER mannosidase, yielding Man₈ species, but prolonged residence in the ER can result in Man₇-glycan containing glycoproteins.

1.4.6. Oligomerisation of Proteins in the ER

Apart from acquiring secondary and tertiary structures prior to exiting the ER, membrane or secretory proteins may have to achieve the correct quaternary (oligomeric) structure for their function. Oligomeric proteins must also meet quality control standards to exit the ER and some mechanisms monitoring their folding are identical to those used for the retention of incorrectly folded subunits. For some proteins, oligomerisation begins in the ER and is not completed until after exit from the ER (e.g. influenza HA; Tatu *et al.*, 1993). In other cases, oligomerisation may not be initiated until leaving the ER (e.g. the HIV-1 envelope glycoprotein; Earl *et al.*, 1991). Oligomerisation can proceed post-translationally, as for influenza HA and VSV G protein (Doms *et al.*, 1987) or co-translationally. One example is thrombospondin, which folds following oligomerisation by means of its first N-terminal domain (Prabakaran *et al.*, 1996). Correct folding is demonstrated by the appearance of new antigenic epitopes and the delayed dissociation of the molecular chaperone BiP/GRP78. Therefore, quaternary organisation can influence the tertiary structure, giving rise to the idea of “oligomerisation-assisted subunit folding”. Procollagen is another example where oligomerisation is required for completion of folding (Bulleid *et al.*, 1997). Moreover, the interdependence between folding and oligomerisation is an intrinsic property of the proteins themselves and not a consequence of information relayed by the cellular environment, since the two distinct oligomerisation pathways (i.e. within and outside the ER) occur in the same cell type. Events governing the productive folding of glycoproteins are summarised in Fig 1.20.

1.4.7. Transport from the ER

Following folding and assembly, proteins must move away from the ER to other target organelles (e.g. the Golgi apparatus) or to the plasma membrane. In the case of viral glycoproteins, they may accumulate at their target site for virus budding. If the protein transits into and through the Golgi apparatus, additional post-translational modifications such as carbohydrate processing, proteolytic cleavage, and oligomerisation may complete the process of conformational maturation. The transport from the ER to Golgi apparatus is mediated by membrane cargo vesicles, that are encased by two types of coats, coat protein complex (COP) I and COPII (Scales *et al.*, 1997). Although the roles of COPs in the secretory membrane traffic are not firmly understood, COPII is believed to mediate anterograde traffic, since these are shown to fuse with Golgi membranes, while COPI mediates retrograde movement (Section 1.4.8; Scales *et al.*, 1997). Components of COPI are necessary for the retrieval of KDEL/K(X)KXX-bearing proteins to the ER (Jackson *et al.*, 1990a) which is essential to maintain ER function. Many membrane proteins that function at the interface of the ER and Golgi complex constitutively cycle between the ER and the Golgi complex (Cole *et al.*, 1998). These include 1) the KDEL receptor (KDELRL), which retrieves soluble ER-resident proteins that have escaped into the secretory pathway, 2) vesicle N-ethylmaleimide-sensitive factor attachment protein receptors (v-SNARE) involved in ER to Golgi transport and, 3) putative cargo receptors, such as the ER-Golgi intermediate compartment (ERGIC) mannose-binding protein p53.

Cargo exported from the ER by vesicle carriers first appear in pre-Golgi intermediates then move on microtubules from the *trans* to the *cis* Golgi regions. The movement of cargo between the Golgi apparatus and beyond, as well as the endocytic compartments, is mediated by vesicles whose formation is initiated by the selection and concentration of cargo. The coat and targeting components function together at different steps of the exocytic and endocytic pathway to form vesicle sorting machines that move specific types of cargo from one destination to the next (Allan & Balch, 1999). It is proposed that retrograde and not anterograde vesicles play a prominent role by functionally modifying cargo-containing compartments generated by the ER. Ultimately, the movement of biosynthetic cargo is controlled by only two sorting steps; one at the ER to select protein that is properly folded and a second at the *trans* Golgi network (TGN), where a sorting decision directs biosynthetic cargo to the proper downstream compartment reflecting the state of the cell (Allan & Balch, 1999). Recycling between the various compartments of

the exocytic pathway provides a link with cell surface events, enabling the cell to balance and regulate outward membrane flow.

1.4.8. Maintenance of the ER through Retrieval/Retention of Resident Proteins

Many luminal ER-resident proteins contain a C-terminal tetrapeptide KDEL sequence (Pelham, 1995). This sequence is utilised to retrieve proteins from the *cis* Golgi via binding to KDEL receptors (KDELR). KDELR specifically bind to KDEL carrying ligands in a pH-dependent manner, with an optimum around pH 5.0 (Nilsson and Warren, 1994). The pH along the exocytic pathway becomes increasingly acidic towards the TGN, and this allows KDELR to bind ligands with high affinity. Ligand-receptor binding would then trigger incorporation of this complex into vesicles to be returned to the ER. The more neutral pH in the ER releases the ligand into the lumen and KDELR returns to its original location awaiting another round of ligand delivery (Nilsson & Warren, 1994). The retrograde movement of KDELR proteins requires a crucial aspartic acid residue, located in one of seven membrane-spanning domains of the receptor protein (Townsley, 1993). The position of the aspartic residue is thought to enable interactions with membrane-spanning domains of other KDELR proteins. Retrograde transport may be achieved by oligomerisation of the receptors which would form patches at the *cis* Golgi that induce budding. Oligomerisation of receptors could be triggered by binding of the ligand.

Several ER-resident membrane proteins also contain a KDEL motif in their cytoplasmic domains (e.g. BiP). However, for other resident type I membrane proteins with the N-terminus in the lumen, a signal consisting of two critical lysines has been identified that requires the residues to be positioned at -3 and -4/-5 relative to the C-terminus of the protein (K(X)KXX, where X is any amino acid). For type II proteins, where the C-terminus is in the lumen, two arginines (RR) are part of the signal, and they have to be within the first five N-terminal residues of the protein (Nilsson & Warren, 1994). Reporter proteins fused with such motifs are found in the ER, but they may acquire also Golgi modifications. Thus, these dibasic signals behave similarly to the KDEL motif and act to retrieve rather than retain proteins (Jackson *et al.*, 1990a; Nilsson & Warren, 1994).

The presence of these motifs alone do not necessarily localise other cellular proteins to the ER, indicating that they cannot be the sole determinants for correct localisation of ER-resident proteins. Interestingly, removal of the K(X)KXX motif from an endogenous ER enzyme, GT, did not result in loss of ER retention, and oligosaccharide analysis of other endogenous ER proteins (e.g. glucosidase II and hexose-6-phosphate dehydrogenase) failed to reveal any Golgi modifications (Brands *et al.*, 1985). This indicated that endogenous ER proteins do not leave the organelle in detectable amounts. Therefore, retention signals must exist in proteins endogenous to the ER that operate independently of ER retrieval signals. Insight into the mechanism for the retention of ER proteins was provided by studies based on Golgi-localised proteins, the majority of which are type II membrane proteins. Based on such evidence, the membrane-spanning domain (the signal anchor) and part of its flanking regions contain sufficient information for correct Golgi cisternae localisation (reviewed by Machamer, 1993). Similarly, the membrane-spanning domain of lamin B receptor, a resident protein of the inner membrane domain of nuclear envelope which is continuous with ER membranes, was found to be sufficient for correct localisation (Smith & Blobel, 1993). Two models have been proposed to explain how a membrane-spanning domain could mediate retention. The first model is based on retention through oligomerisation and exclusion from anterograde transport vesicles (Nilsson & Warren, 1994). The second model postulates that retention occurs through differences in membrane thickness along the exocytic pathway (Bretscher & Munro, 1993). This model relies on the observation that the ER has low cholesterol content as compared with the plasma membrane (van Meer, 1989). Studies on the transmembrane domain of a yeast protein, UBC6, have shown that it functions as an ER-retention sequence in mammalian cells on fusion to the transferrin receptor (Yang *et al.*, 1997). Lengthening of the UBC6 transmembrane domain from 17 to 21 amino acids permits movement from the ER to the Golgi complex. It is proposed that the thicker membranes in the Golgi complex play an active role in such sorting mechanisms.

1.4.9. ER Stress Response and Protein Degradation

The quality control mechanisms of the ER ensure the transport of incompletely folded, misassembled, and unassembled proteins to the secretory pathway is selectively inhibited. These defective proteins are retained in the ER, where their accumulation activates an ER-stress response and/or they are eventually degraded. As well as cellular proteins, the fact that foreign proteins such as viral glycoproteins are also subject to quality control

demonstrates that the sorting criteria is based on general rather than specific structural properties of the protein products. However, the tendency of viral glycoproteins to form large aggregates increases the likelihood of these proteins being retained in the ER. BiP forms a stable association with misfolded proteins and, via its KDEL sequence, these proteins remain in the ER.

The accumulation of incorrectly-folded proteins triggers an ER-stress response that has at least two distinct components. The first is termed the unfolded-protein response (UPR) that describes the transcriptional activation of ER resident chaperones (BiP) and foldases (PDI; Shamu *et al.*, 1994). The second component consists of a rapid repression of protein synthesis. Both responses relieve ER stress by increasing the capacity of the ER to actively fold proteins and decrease the demands made on the organelle by attenuating protein synthesis rate (Harding *et al.*, 1999).

The first gene known to be required for a functional UPR was *IRE1*, which encodes a transmembrane protein (Cox *et al.*, 1993). Ire1p resides in the ER or inner nuclear membrane that is continuous with the ER. Its N-terminal region faces the ER lumen, and the C-terminal portion faces the nucleus or cytoplasm. The N-terminal half serves as a sensor domain that detects the accumulation of unfolded proteins while the C-terminal region possesses a protein kinase function. Induction of the UPR is initiated by oligomerisation of Ire1p via the transmembrane domains resulting in trans-autophosphorylation by neighbouring Ire1p molecules (Fig 1.21). Subsequently, instead of triggering a classical kinase cascade, activated Ire1p cleaves *HAC1^u* mRNA (u for uninduced), which codes for a transcription factor, Hac1p, at two splice junctions. tRNA ligase, recruited to the activated Ire1p joins the two exons to produce the spliced form of the message, *HAC1ⁱ* mRNA (I for UPR-induced). Both forms of *HAC1* mRNA exit the nucleus and associate with polyribosomes. However, only the spliced form gives rise to protein, Hac1p^I, which enters the nucleus and upregulates genes containing the unfolded protein response element (UPRE; Fig 1.21; Sidrauski *et al.*, 1998).

The reduction in protein synthesis in response to ER stress is attained by disassembly of polysomes and increased phosphorylation of eIF2 α , which interferes with the formation of an active 43S translation-initiation complex. Two different eIF2 α kinases have been identified in mammalian cells, haem-regulated eIF2 α kinase (HRI) and the interferon-

inducible PKR. More recently, an ER-membrane bound protein, with a luminal domain that shares blocks of identity with Ire1p and a cytoplasmic portion that has a protein-kinase domain similar to that of PKR and HRI, has been named the PKR-like ER kinase (PERK; Harding *et al.*, 1999). This protein is proposed to be an alternative eIF2 α kinase that responds to ER stress.

The efficiency with which chaperones retain proteins in the ER prevents their premature degradation. However, prolonged retention of proteins in association with BiP directs these proteins towards the degradation pathway. Recent communications have shown significant amounts of irreversibly misfolded proteins are dislocated from the ER by reverse transport through the Sec61 translocon into the cytosol where they can be degraded by the ubiquitin-regulated proteasome pathway. This process has been termed ER-associated degradation (ERAD; Brodsky & McCracken, 1997). Incorrectly-folded glycoproteins translocated to the cytosol are initially deglycosylated by the action of cytosolic N-glycanase (PNGase) prior to proteolytic degradation (Suzuki *et al.*, 1998), although the intracellular location of these deglycosylation events is not known. The ER membrane protein, calnexin, also recruits cytosolic Ubc7p to the ER surface and thereby mediates its activation. Ubc7p is an ubiquitin-conjugating enzyme that catalyses the covalent attachment of ubiquitin to specific proteolytic substrates and this process appears to be a prerequisite for retrograde transport out of the ER. The involvement of the translocon and calnexin suggests that the export and degradation pathway, both of which are ATP-dependent, may be coupled in some way. Moreover, certain viral proteins, namely human cytomegalovirus (HCMV) US2 and US11, and HIV-1 Vpu, can specifically mediate the cytosolic export and subsequent degradation by the ubiquitin-proteasome machinery of ER-associated MHC class I and CD4 molecules, respectively (Margottin *et al.*, 1998). It is postulated that other cellular protein components interact with Sec61 to form a retro-translocon programmed for reverse protein translocation, to operate the ER degradation process (Plemper & Wolf, 1999). Indeed, in yeast, proteins called Der and Hrd are necessary for ER-associated degradation via the ubiquitin-proteasome pathway. Since similar complex degradation pathways exist in higher eukaryotes, it is likely that mammalian counterparts of Der and Hrd are present which have similar functions.

1.5. Viral Glycoproteins of the *Flaviviridae*

As a member of the *Flaviviridae* family of viruses, E1 and E2 glycoproteins of HCV have a number of properties in common with the envelope proteins of related pestiviruses and flaviviruses. These include their position in the polyprotein, processing to produce individual polypeptides by host proteinases, and orientation in the lipid membrane (see Fig 1.2).

1.5.1. Glycoproteins of Pestiviruses

Pestiviruses (BVDV, CSFV, and border disease virus) encode three envelope-associated glycoproteins (E proteins), E^{ms} (gp 44), E1 (gp33), and E2 (gp55). The C-terminus of the virion core protein is hydrophobic which is proposed to serve as the signal sequence initiating translocation of E^{ms} into the ER lumen. The E^{ms} glycoprotein of CSFV contains up to nine potential N-linked glycosylation sites and exists as disulphide-linked homodimers of about 97 kDa (Thiel *et al.*, 1991; Rumenapf *et al.*, 1993). E^{ms} does not appear to contain a membrane anchor and is secreted from cells in considerable quantities (Rumenapf *et al.*, 1993). E1 and E2 are predicted to be integral membrane proteins and contain three and six N-linked glycosylation sites, respectively (Weiland *et al.*, 1990). Similar to E^{ms}, E2 forms homodimers but also associates with E1 to generate a heterodimeric complex (Thiel *et al.*, 1991). All of the above interactions are mediated by covalent, intermolecular disulphide bonds. In virions, it is considered that E1 and E2 are anchored in the lipid bilayer while the mechanism by which E^{ms} associates is unclear. It is possible that it interacts either non-covalently with E1E2 dimers or by hydrophobic interactions with the lipid bilayer (Rumenapf *et al.*, 1993). Antibodies that neutralise virus infectivity are directed against epitopes located on E2 and E^{ms} (Weiland *et al.*, 1990 and 1992), but effectively no anti-E1 antibodies are detected, suggesting that it may be buried in the viral envelope. Purified preparations of either E^{ms} or E2 could inhibit infection by CSFV but apparently by separate mechanisms (Holst & Moormann, 1997). Moreover, CSFV E2 could block infection of cells by BVDV suggesting that the viruses share common receptors.

HCV does not appear to encode a protein equivalent to E^{ms} which has been identified as an RNase (Schneider *et al.*, 1993). The RNase activity of E^{ms} has been mapped to two short

regions of amino acids, both containing histidine residues essential for catalysis. E^{ms} strongly inhibits the protein synthesis of lymphocytes leading to apoptosis of such cell types (Bruschke *et al.*, 1997). Since pestivirus infections are characterised by leukopenia (reduction in the number of leucocytes) and immunosuppression, the above data suggests that E^{ms} plays an important role in the pathogenesis of pestiviruses. However, a contrasting affect of the protein has been reported in non-lymphocytic cells (Hulst *et al.*, 1998). This study reported that RNase-negative CSFV induced apoptosis in swine kidney cells, and the affect was neutralised by rescue of the inactivated E^{ms} gene. The authors speculated that the RNase activity may depress levels of cellular and viral RNA in infected cells, thus enhancing viral persistence.

1.5.2. Glycoproteins of Flaviviruses

Flaviviruses [e.g. DEN, YFV and JEV] encode two envelope proteins designated PrM (precursor to M) and E (50 kDa). E protein is a multi-functional protein that is involved in cell receptor binding and virus entry via fusion with a host cell membrane at permissible pH (Heinz, 1986). From X-ray crystal studies, the structure of a soluble fragment from the E ectodomain of tick-borne encephalitis (TBE) virus, generated by trypsin digestion, has been determined (Rey *et al.*, 1995). This has revealed that the ectodomain consists mostly of β -sheet and loop structures but long segments of α -helix are absent. By contrast to most other known viral fusion proteins that are trimeric spikes in both native and fusogenic states (e.g. influenza HA protein; Bullough *et al.*, 1994), the native form of the E ectodomain is a head-to-tail dimer oriented parallel to the viral membrane. It is composed of three separate domains which contain the antigenic determinants, the putative receptor binding site and the fusion peptide. All three domains are involved in interactions between monomers of the E ectodomain. The C-terminal region of E protein, referred to as the stem-anchor region, is important for interaction with PrM during virus assembly and the structural changes associated with membrane fusion at low pH (Allison *et al.*, 1999). In the full-length, native state, the C-terminus of each monomer of the ectodomain would be followed by the stem-anchor region of the protein at distal ends of the homodimer. The stem region contains two regions (H1^{pred} and H2^{pred}) predicted to be α -helical while the anchor domain consists of two transmembrane segments (TM1 and TM2). H1^{pred} has a role in the trimerisation of homodimers at low pH, a process that is irreversible and necessary

for fusion activity (Allison *et al.*, 1995 and 1999). H2^{pred} and TM1 are contiguous and involved in interactions with PrM (Allison *et al.*, 1999).

PrM is approximately 22 kDa and is a glycosylated precursor to M (7-8 kDa). PrM forms a heteromeric complex with E in immature virions (Wengler & Wengler, 1989), an interaction that is proposed to prevent acid-induced conformational changes in the E protein during virion maturation in the TGN and secretory vesicles (Heinz *et al.*, 1994; Stiasny *et al.*, 1996; Pryor *et al.*, 1998). By contrast to the covalent interactions between pestivirus E1 and E2 glycoproteins, PrM and E associate by non-covalent links. E-PrM complexes for several flaviviruses (JEV, YFV, DEN and TBE viruses) are immunogenic and protective against challenge (Mason *et al.*, 1991; Pincus *et al.*, 1992; Fonseca *et al.*, 1994; Heinz *et al.*, 1995). Assembly of flaviviruses occurs in the ER and immature virions are transported via the secretory pathway to the TGN. For most flaviviruses, just before or at the time of release from the cell, cleavage of PrM on immature virions to produce mature M is mediated by host furin activities that are located in the vesicles between the TGN and plasma membrane (Rice, 1996; Stadler *et al.*, 1997). One exception is DEN virus in which cleavage of PrM to M is inefficient, resulting in the release of virions containing the precursor protein (Wang *et al.*, 1999). However, infectivity of DEN virus does not appear to depend on quantitative cleavage of PrM to M, as virus particles containing PrM and no M protein retain about one-eighth the infectivity of normal particles (Randolph *et al.*, 1990).

Unlike HCV, flaviviruses encode a non-structural membrane glycoprotein of 45 kDa called NS1. The bulk of the protein remains within the cell in a homodimeric form where it attaches to membranes (Smith and Wright, 1985; Winkler *et al.*, 1989; Fan and Mason, 1990). A significant amount of the protein also is secreted and small quantities are present on the cell surface. In the secreted form, hexameric species of NS1 have been detected (Flamand *et al.*, 1999). Recent evidence suggests that the protein has a role in RNA replication such that mutations in NS1 induce reduced accumulation of viral RNA (Muylaert *et al.*, 1996 and 1997). Complexes of NS1 dimers and E protein also have been detected however these may represent non-specific aggregates (Blitvitch *et al.*, 1995).

1.6. The Semliki Forest Virus Expression System

Semliki Forest virus (SFV) has been used as a model to study the molecular biology of RNA virus multiplication since it grows well and to high titre in a broad range of tissue culture cells such as baby hamster kidney (BHK) cells (Berglund *et al.*, 1993) and chick embryo cells (Liljestrom & Garoff, 1991b). SFV is a positive-stranded RNA virus of the genus *Alphavirus* of the family *Togaviridae*. It is a relatively simple virus, encoding only nine distinct proteins. Four non-structural proteins are concerned with viral RNA synthesis, and four structural proteins form the capsid (the C protein) and the envelope (E1, E2, and E3) proteins. A small 6 kDa protein encoded by the structural region of the genome is not incorporated into virions. The structural proteins are encoded by a sub-genomic RNA species, termed 26S, while the non-structural proteins are translated from the genomic 42S RNA. The structural and non-structural proteins are generated by post-translational cleavage from precursor polyproteins (Atkins *et al.*, 1999). The original isolate of SFV, L10, is neurovirulent for mice and rats, causing lethal encephalitis by infection of the central nervous system (CNS) and depletion of myelin-forming cells (Liljestrom *et al.*, 1991).

The infectious clone of SFV, constructed from the prototype strain, is designated pSP6-SFV4; the virus produced by transcription of this infectious clone is labelled SFV4. An additional feature from infection of mammalian cells with SFV4 *in vitro* is increased cell death as a result of apoptosis. As with other viruses including alphaviruses, cell death is concurrent with infection and is thought to occur from inhibition of cell macromolecular synthesis as well as general disruption of cell metabolism and synthetic processes. Use of the infectious clone has shown that the non-structural region of the SFV genome is necessary for the induction of apoptosis since the structural region may be deleted without affecting apoptosis (Glasgow *et al.*, 1998). Moreover, deletion of part of the non-structural nsP2 gene abrogates both apoptotic induction and viral RNA synthesis. By contrast, for Sindbis virus, another model alphavirus, apoptosis induction, viral virulence and counteraction of the effects of the bcl-2 gene (a cellular function involved in cell survival) are all functions of the viral E2 envelope protein gene (Rathmell & Thompson, 1999). Apoptosis induced by SFV is not dependent on p53 (a cell product which blocks proliferation or activates apoptosis), and therefore does not result from damage to cellular DNA. Possible mechanisms of SFV-induced apoptosis include activation of PKR or caspase 3 due to leakage of cytochrome *c* from mitochondria.

The initial vector system developed from the full-length infectious clone of SFV induced high-level transient expression of cloned genes in transfected cells (Liljestrom & Garoff, 1991a). The genes of interest were cloned in place of the structural proteins of SFV, normally generated from the sub-genomic 26S RNA species. To achieve production of a heterologous gene, RNA is transcribed from the SFV vector *in vitro* and transfected into cultured cells. Viral RNA is replicated from both 42S and 26S promoters and expression occurs as in a productive infection. In addition, a helper clone was constructed by deletion of the RNA packaging signal which lies in the nsP2 gene. This RNA sequence is recognised by the capsid protein for encapsidation. The structural proteins deleted in the SFV vector are thus provided *in trans* by the helper construct. Dual transfection of helper and vector constructs result in the packaging of vector RNA into particles, but not the helper (Liljestrom & Garoff, 1991b). Particles are subsequently released from the cell and are infectious. However, since packaged vector RNA lacks the viral structural protein genes, no progeny are produced from cells infected with these particles.

1.7. Aims of the Project

The major aim of this project was to study the complexes formed by the HCV E1 and E2 glycoproteins. Initially, the glycoproteins produced by strain Glasgow, a local genotype 1a strain, were analysed. Subsequently, E1 and E2 from another strain, H77, which is infectious in chimpanzees, were used to extend the study and for comparative purposes. Since no propagation system is available for HCV, expression was achieved using the SFV vector, pSFV1. At the outset of the project, little was known about the role of disulphide bonds in the formation and stabilisation of complexes. To address this question, disulphide bonds were disrupted both biochemically and by site-directed mutation of cysteine residues. A set of criteria also were established to distinguish the different types of complex (native and aggregated) formed by E1 and E2. These were employed to determine the role of sequences in E1 and E2 in the formation of native and aggregated complexes.

HCV associated disease	Percentage	Time period of detection	Symptoms and disease
Acute HCV infection	~20%	7-8 weeks	15% clear the virus Jaundice in some cases
Chronic HCV infection	~80%	Over a period of 10 years post infection	Persistent infection with HCV, characterised by detection of HCV RNA at least 6 months after acute HCV infection
Cirrhosis	~20%	20-30 years post infection	Liver inflammation in patients with chronic HCV infection, characterised by abnormal levels of liver ALTs
HCC	~4%	Beyond 30 years post infection	Development of HCC and liver failure
Extra-hepatic	NK	NK	e.g. cryoglobulinemia, arthralgias, glomerulonephritis

Table 1.1 Natural history of HCV infection. (NK – not known; adapted from Lavanchy *et al.*, 1999)

Genus	Genome size	Genome Features	Translation strategy	Virion proteins	Secreted glycoproteins	Common non-structural proteins/functions	Distinct non-structural proteins/functions	Viral biology
<i>Flavivirus</i>	~11 kb	5' cap; short 5' UTR; ~3 400 aa ORF; long 3' UTR	Cap-dependent	C - basic, poorly conserved. M and E - transmembrane glycoproteins; PrM premature M protein	NS1	Serine proteinase - NS2B and NS3. NTPase/RNA helicase - NS3. RdRp - NS5	RNA triphosphatase motif in NS3 and methyl transferase motif in NS5	Transmitted by mosquitoes or ticks; acute limited infection
<i>Pestivirus</i>	~12.5 kb	Long 5' UTR; ~4 000 aa ORF; short 3' UTR	IRES	C - basic, highly conserved. E1 and E2 - transmembrane glycoproteins. E ^{ms} - virion associated RNase.	E ^{ms} - secreted and virion associated	Serine proteinase - p125/p80. NTPase/RNA helicase - p125/p80. RdRp - p75	Additional viral autoproteinasase, N ^{pro} .	No known insect vector; can cause chronic infection
<i>Hepacivirus</i>	~9.4 kb	Long 5' UTR; ~3 000 aa ORF; long 3' UTR	IRES	C - basic, highly conserved. E1 and E2 (and E2-p7) - transmembrane glycoproteins.	None identified	Serine proteinase - NS3 and NS4A. NTPase/RNA helicase - NS3. RdRp - NS5B	Additional viral metallo (Zu) proteinase, NS2-3	Blood borne; commonly causes chronic infection.

Table 1.2. Comparison of features of flaviviruses, pestiviruses, and hepatitis C virus. (RdRp - RNA dependent RNA polymerase, containing a GDD motif; adapted from Rice, 1996)

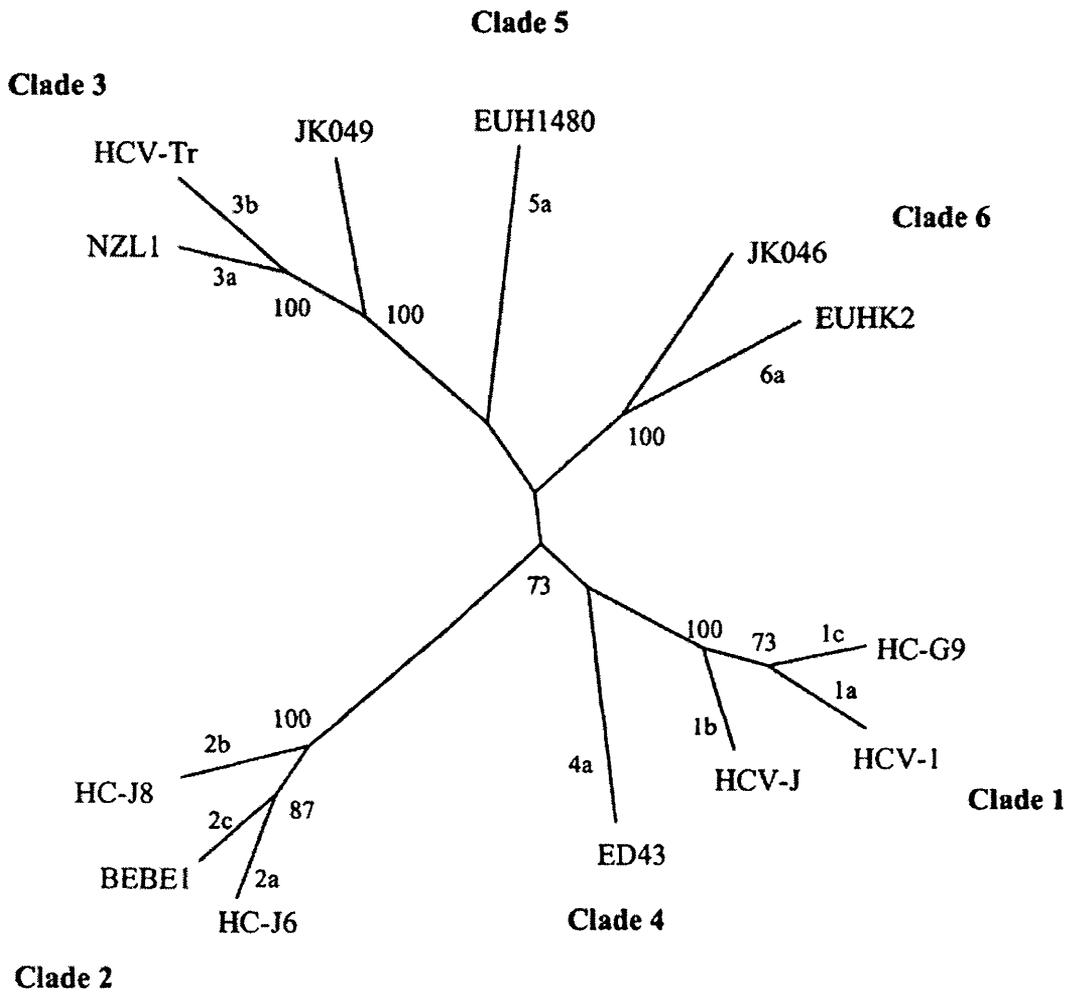


Fig 1.1. Phylogenetic analysis of whole representative HCV sequences (Taken from Robertson *et al.*, 1998)

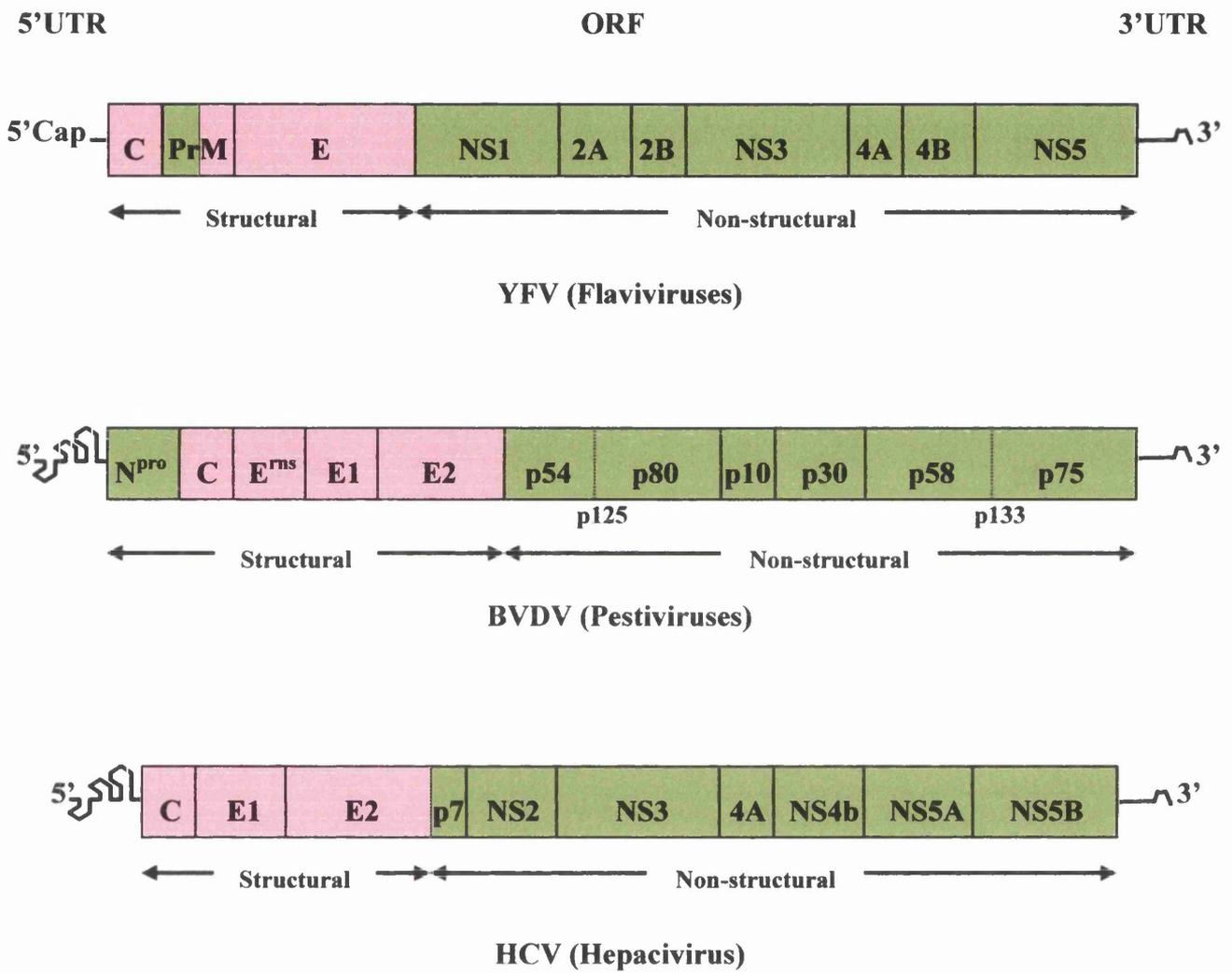


Fig 1.2. Genomic organisation of the *Flaviviridae* family of viruses.

ER Lumen

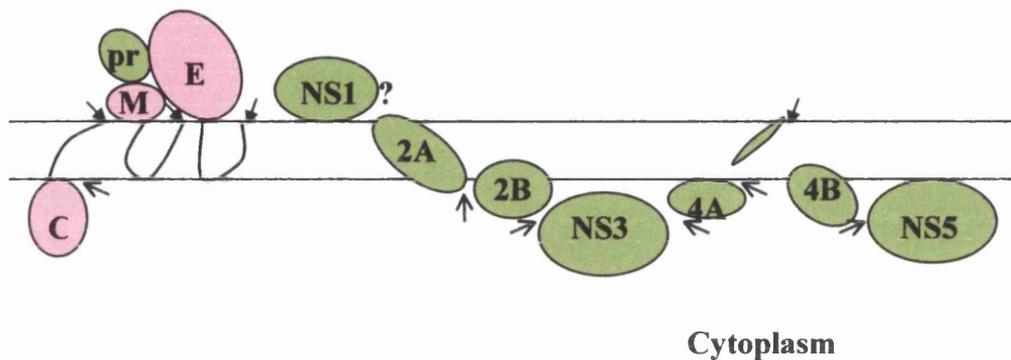
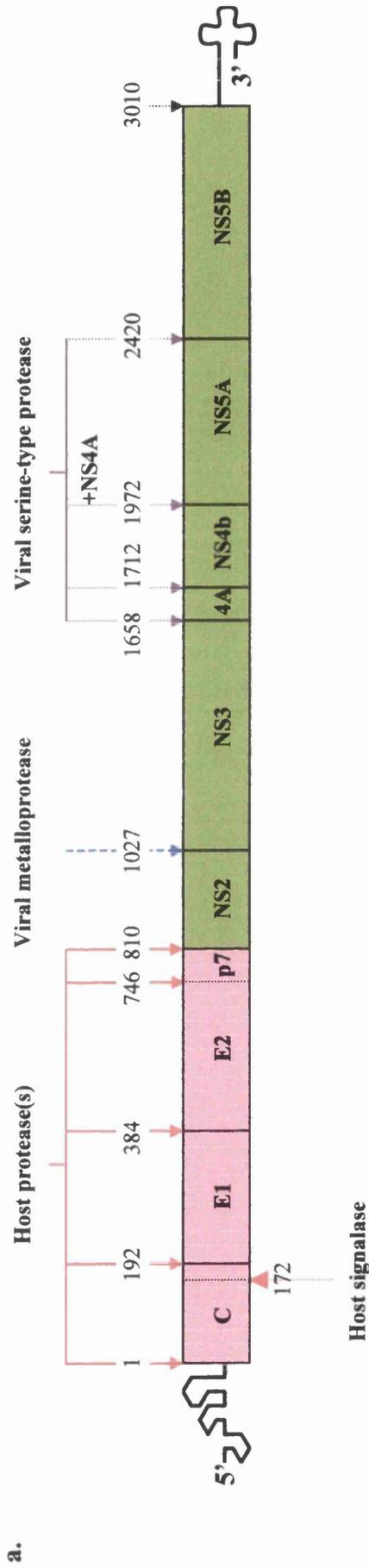


Fig 1.3. Model for the membrane topology of the flavivirus proteins. A diagram of the flavivirus polyprotein cleavage products with respect to the ER membrane is shown. The proteins are drawn to scale (areas are proportional to the number of amino acids) and arranged in order (left to right) of their appearance in the polyprotein. Mature structural proteins are shaded and C-terminal membrane spanning segments of M and E are indicated. Also shown are the cleavage sites for host proteinase (↔), the viral serine protease (→), and an unknown protease responsible for cleavage at the NS1/2A site (↘). (Taken from Rice, 1996)



b.

C	E1 E2 (E2-p7)	p7	NS2	NS3	NS4A	NS4B	NS5A	NS5B
Core (Trans-activator?)	Envelope proteins	?	C-terminus metallo-protease	Protease Helicase Oncogene-like function?	Co-factor of NS3 Required for NS5A (P)	?	Phospho-protein ? PKR modulator	RNA dependent RNA polymerase

Fig 1.4. a. Genome structure of HCV. The numbers above the schematic represent the N-termini of each protein on the HCV polyprotein, except the number 3010, which indicates the C-terminus of the polyprotein. Possible proteases involved in the cleavages are indicated. The proteolytic function for the viral metalloprotease spans the C-terminus of NS2 and the N-terminus of NS3, and a serine-like protease is encoded in NS3. **b.** Functions and putative functions of each protein are shown. [(P) – phosphorylation] (Adapted from Shimotohno & Feinstone, 1997)

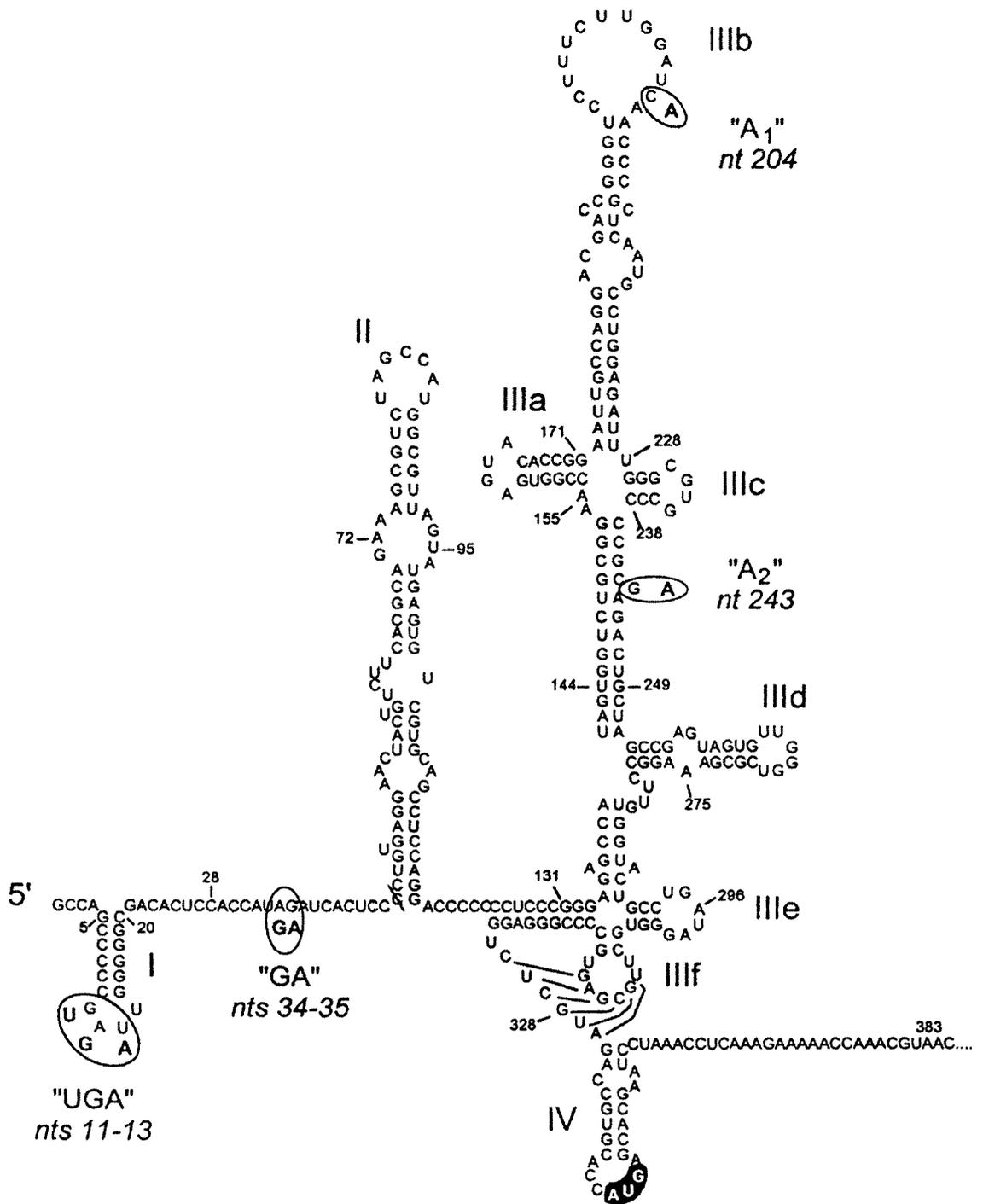


Fig 1.5. Predicted model for the secondary and tertiary RNA structure of the 5'UTR of HCV. The segment immediately downstream of the ORF of HCV-1b is included. Major structural domains are labeled I, II, III, and IV. The initiator AUG codon in stem-loop IV is highlighted. The circled nucleotides indicate differences between the sequences of HCV-1b and 1a, which are clustered at four loci: UGA, GA, A₁, and A₂. (Taken from Honda *et al.*, 1999)

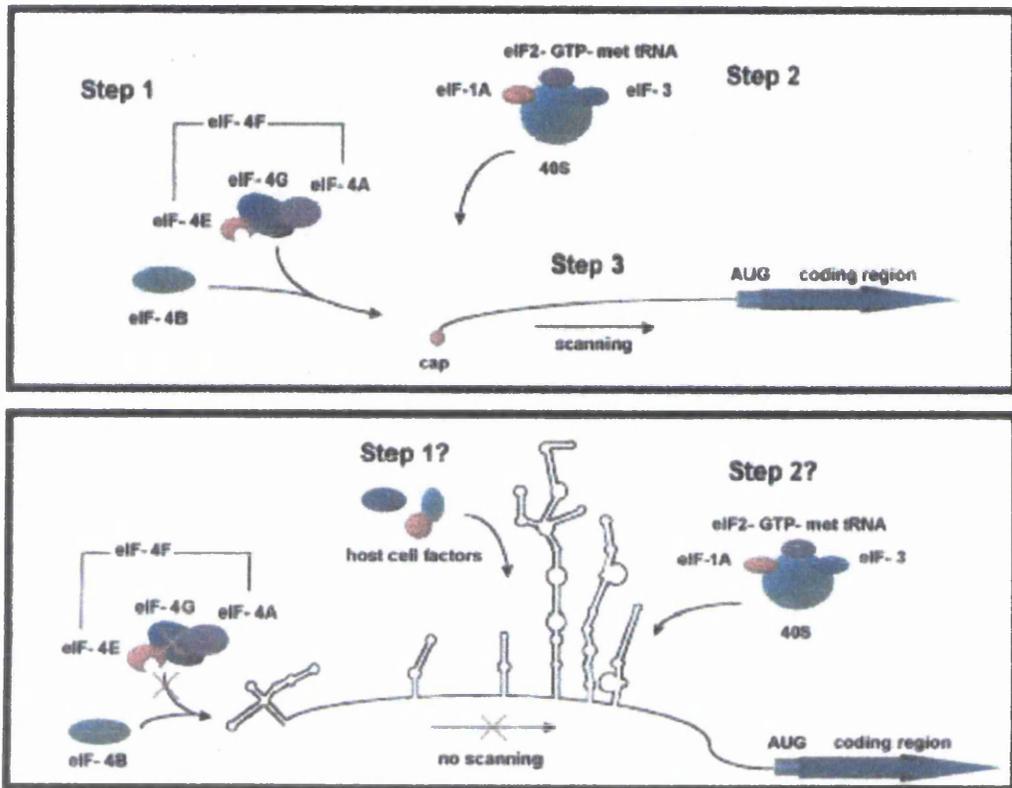


Fig 1.6. Overview of cap-dependent and cap-independent translation initiation. In cap-dependent translation (top), the eIF-4F complex recognises the m^7GTP cap at the 5' end of an mRNA. ATP-dependent RNA-RNA unwinding activities of eIF-4A are stimulated by an additional factor, eIF-4B. After binding the mRNA, the 40S ribosome and ternary complex typically initiate translation at the first-encountered AUG codon in an appropriate sequence context. Following poliovirus infection (bottom), cap-dependent translation in the host cell is abolished by the proteolytic cleavage of eIF-4G. Picornavirus RNA lacks a m^7GTP cap structure and, therefore, initiates translation in an end-independent manner. The 43S preinitiation complex may be recruited to the IRES by viral RNA structure determinants and host cell protein factors. Complex RNA and protein interactions appear to allow the 40S ribosome and ternary complex to bypass multiple upstream AUG sequences, thus directing viral polyprotein synthesis at the appropriate downstream AUG. (Taken from Stewart & Semler, 1997)

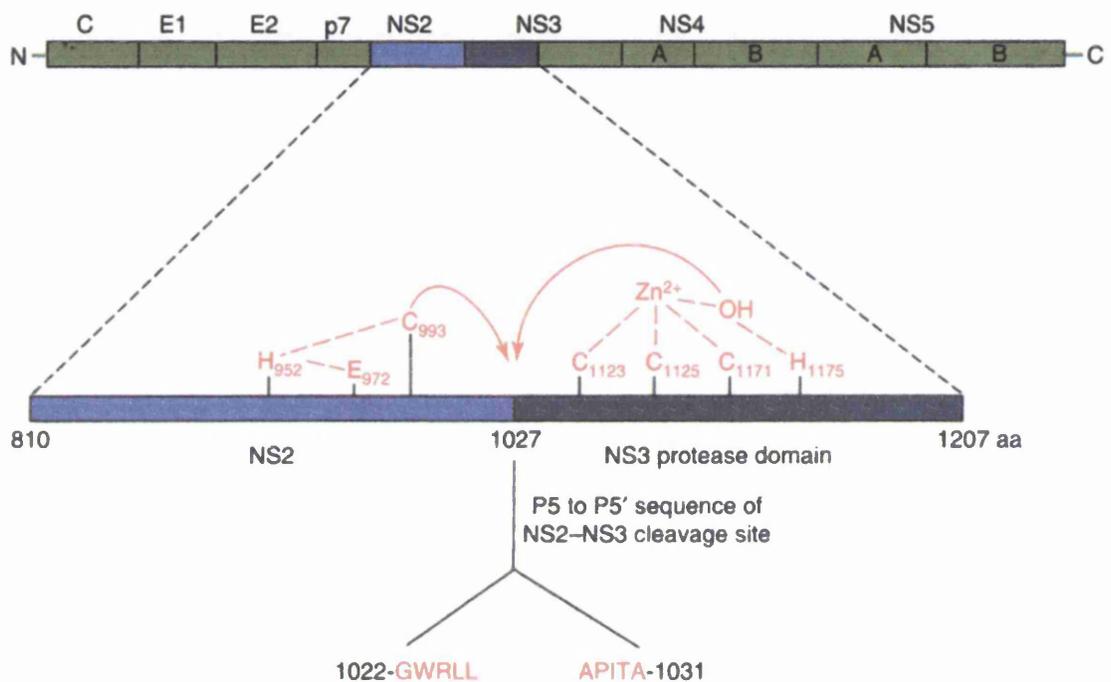


Fig 1.8. The HCV polyprotein and proposed model of NS2-NS3 autoproteolysis. The zinc divalent cation in the NS3 proteinase domain is illustrated as are amino acid residues in NS2 that are potentially involved in catalysis through a cysteine proteinase mechanism. Amino acid residues are numbered according to the HCV-1a polyprotein sequence. The P5 to P5' sequence (amino acid residues 1022-1031) of the NS2-NS3 cleavage site is included. The curved arrows indicate the NS2-NS3 cleavage site between residues 1026 and 1027. (Taken from Wu *et al.*, 1998)

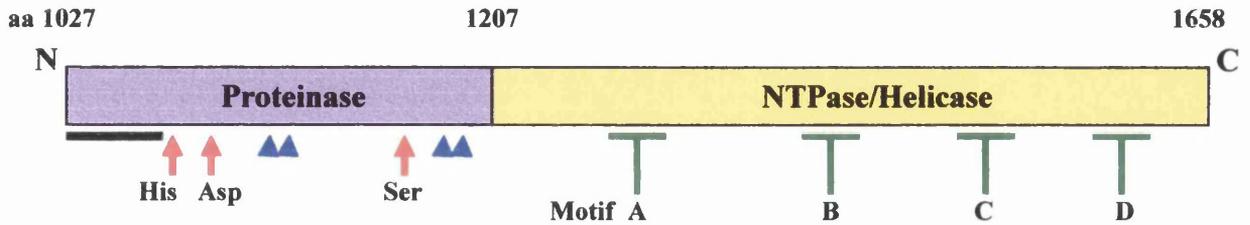


Fig 1.9. Schematic showing the features identified in NS3. The NS3 serine proteinase is located in the N-terminal 180 aa. The 22 N-terminal residues of this region are involved in binding with NS4A, indicated by (—). The catalytic triad is composed of H (1083), D (1107) and S (1165), which are indicated (↑). Finally, the NS3 proteinase structure is maintained by a zinc ion that is held by co-ordinates provided by C residues at 1125, 1127, and 1171, and His residue at 1175; these are indicated also (▲). Four distinct motifs are identified in the NS3 NTPase/helicase region designated A, B, C, and D. Motif A is A-X-X-X-X-G-K-S and motif B is the DEAD box subfamily sequence of D-E-C-H. Motifs C and D are T-A-T and Q-R-R-G-R-T-G-R, respectively.

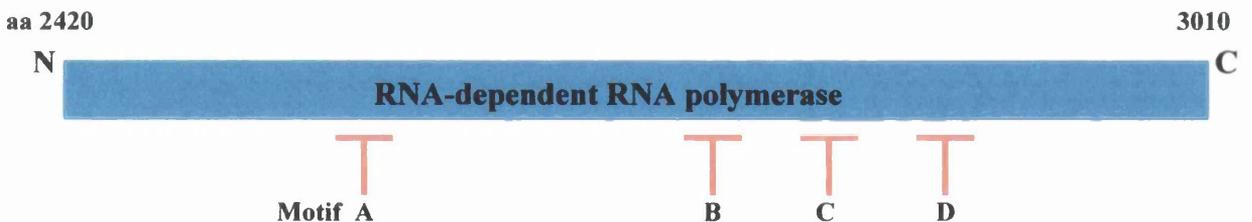
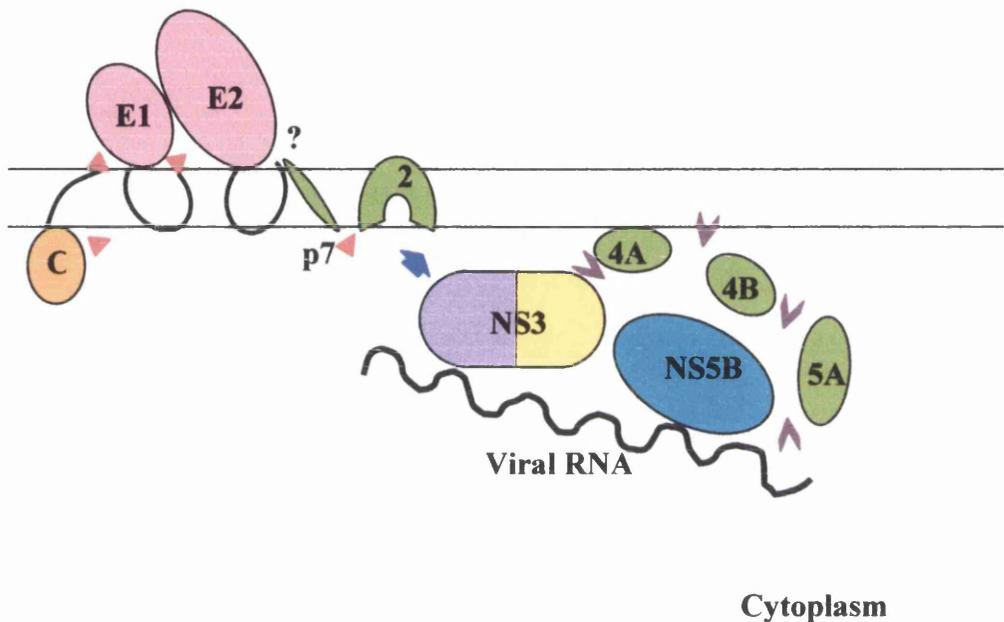


Fig 1.10. Schematic representation of NS5B. The four motifs shown within NS5B protein have the following sequences (in brackets) and predicted functions. Upper and lower case letters represent well- and poorly-conserved residues respectively. Motif A (DtrcfD) nucleotide binding and catalysis, motif B (sGvltTscgN) template binding and primer positioning, motif C (GDD) NTP binding and catalysis and motif D (amtry), NTP binding and catalysis. Motif C is characteristic of most polymerases. In motif D, a lysine residue, highly conserved in other polymerases and reverse transcriptases, is replaced by an arginine residue in NS5B. (Lohmann *et al.*, 1997)

ER Lumen



Cytoplasm

Fig 1.11. Model for the membrane topology of the HCV proteins. The diagram shows the HCV cleaved products with respect to the ER membrane. The proteins are processed according to the positions shown in Fig 1.4a. Mature structural proteins are shown in a different colour (■) from the non-structural proteins (■), although NS3 proteinase/helicase domains and NS5 RdRp are shaded according to Fig 1.9 and Fig 1.10 respectively. Cleavage from the polyprotein occurs through the action of cellular proteinases (▼), NS2/3 metalloprotease (▼), and by NS3/4A serine proteinase (▼). The precise function for NS4B and NS5A in the replication of viral RNA and/or packaging of RNA and assembly is unknown.

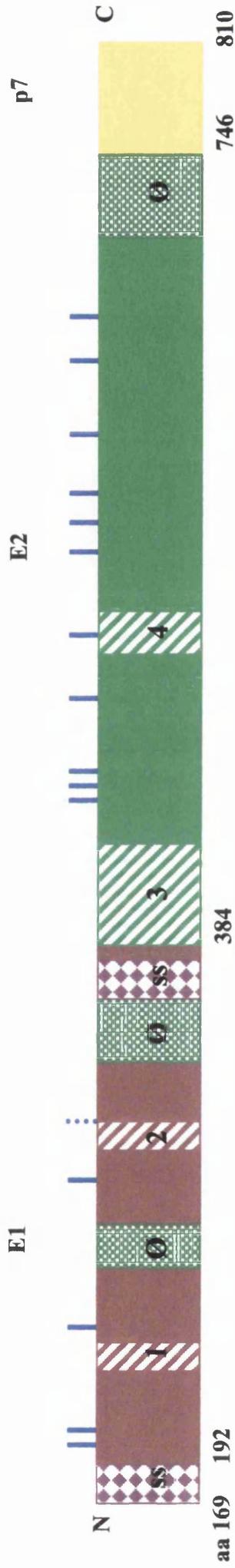


Fig 1.13. Properties of E1 and E2 glycoproteins. E1 and E2 sequences correspond to amino acid residues 192-383 and 384-745 of the HCV polyprotein, respectively. The signal sequence (ss) required for directing E1 and E2 proteins into the ER lumen are located at the C-terminal end of core protein (between residues 169-191) and E1 respectively. The hydrophobic regions are indicated by chequered green boxes (Ø) where the E2 C-terminal end hydrophobic region behaves as a transmembrane domain and may also act as a signal sequence to direct p7 into the ER. Striped regions, 1, 2, 3, and 4 indicate variable regions. Region 3 corresponds to HVR-1 in E2. The blue bars represent the glycosylation sites in E1 and E2 encoded by the infectious clone, H77. The site at position 325 in E1 (indicated by the dotted line) may not be glycosylated (Meunier *et al.*, 1999). The region representing p7 is shown in yellow.

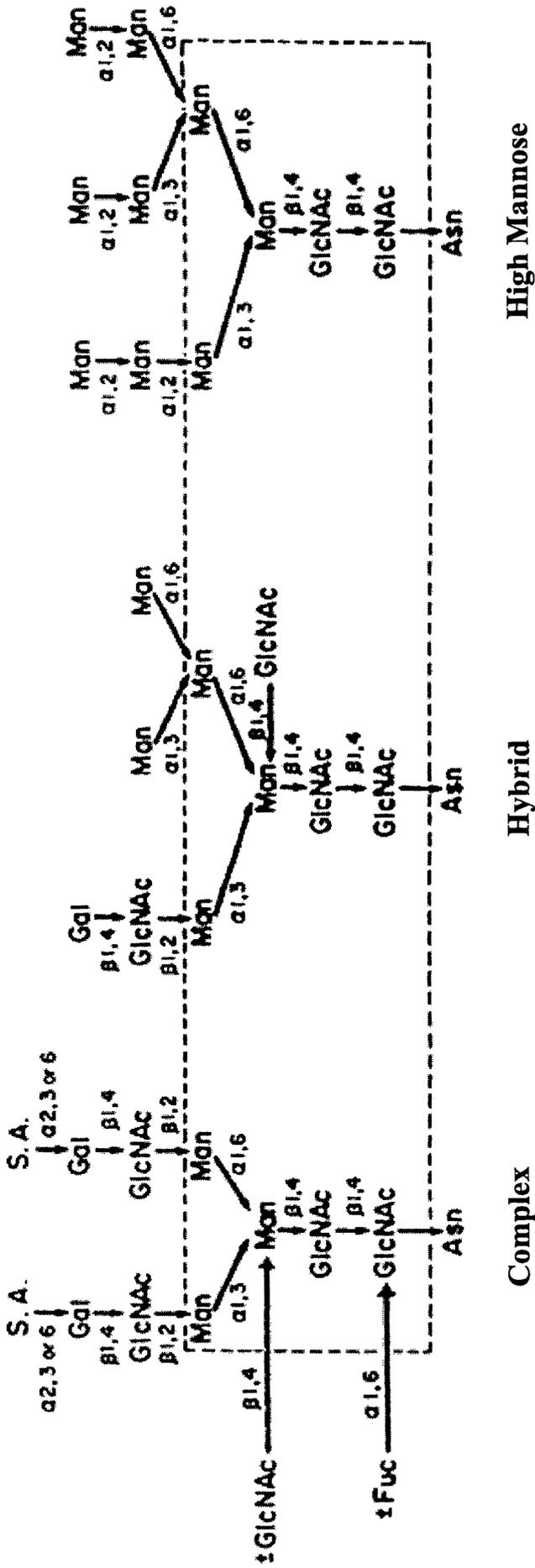


Fig 1.14. Structure of the major types of asparagine-linked oligosaccharide. The boxed area encloses the pentasaccharide core common to all N-linked structures. Abbreviations are as follows: Fuc, fucose; Gal, galactose; GlcNAc, N-acetylglucosamine; Man, mannose; S.A., sialic acid. (Taken from Kornfeld & Kornfeld, 1985)

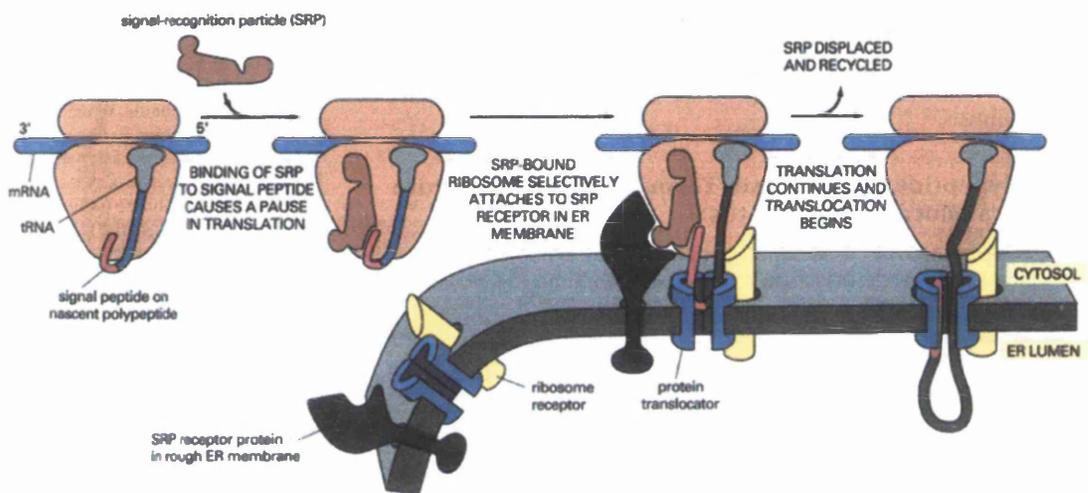


Fig 1.15. ER signal peptides and SRP direct ribosomes to the ER membrane. The SRP and the SRP receptor are thought to act in concert. The SRP binds to the exposed ER signal peptide and to the ribosome, thereby inducing a pause in translation. The SRP receptor in the ER membrane, which is composed of two different polypeptide chains, binds the SRP ribosome complex. In a poorly understood reaction that involves multiple GTP-binding proteins, the SRP is released, leaving the ribosome on the ER membrane. A multisubunit protein translocation apparatus in the ER membrane then inserts the polypeptide chain into the membrane and transfers it across the lipid bilayer. (Taken from Alberts *et al.*, 1994).

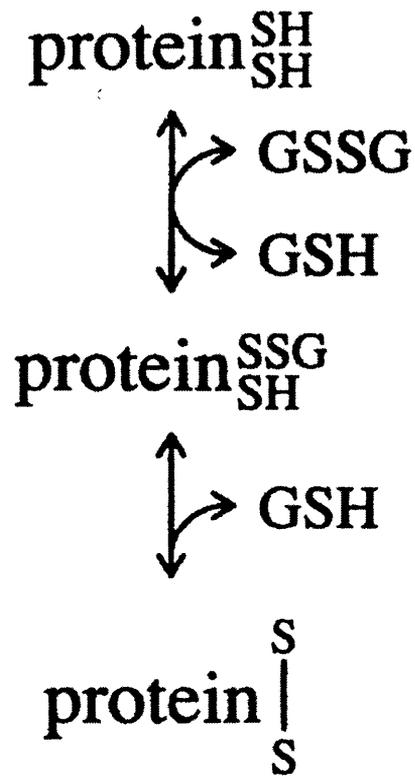


Fig 1.16. Simplified disulphide redox reaction. Formation of one disulphide bond requires two sequential thiol-disulphide exchanges involving a mixed-disulphide intermediate. Glutathione exists as the thiol (GSH) and disulphide (GSSG) forms in the cell.

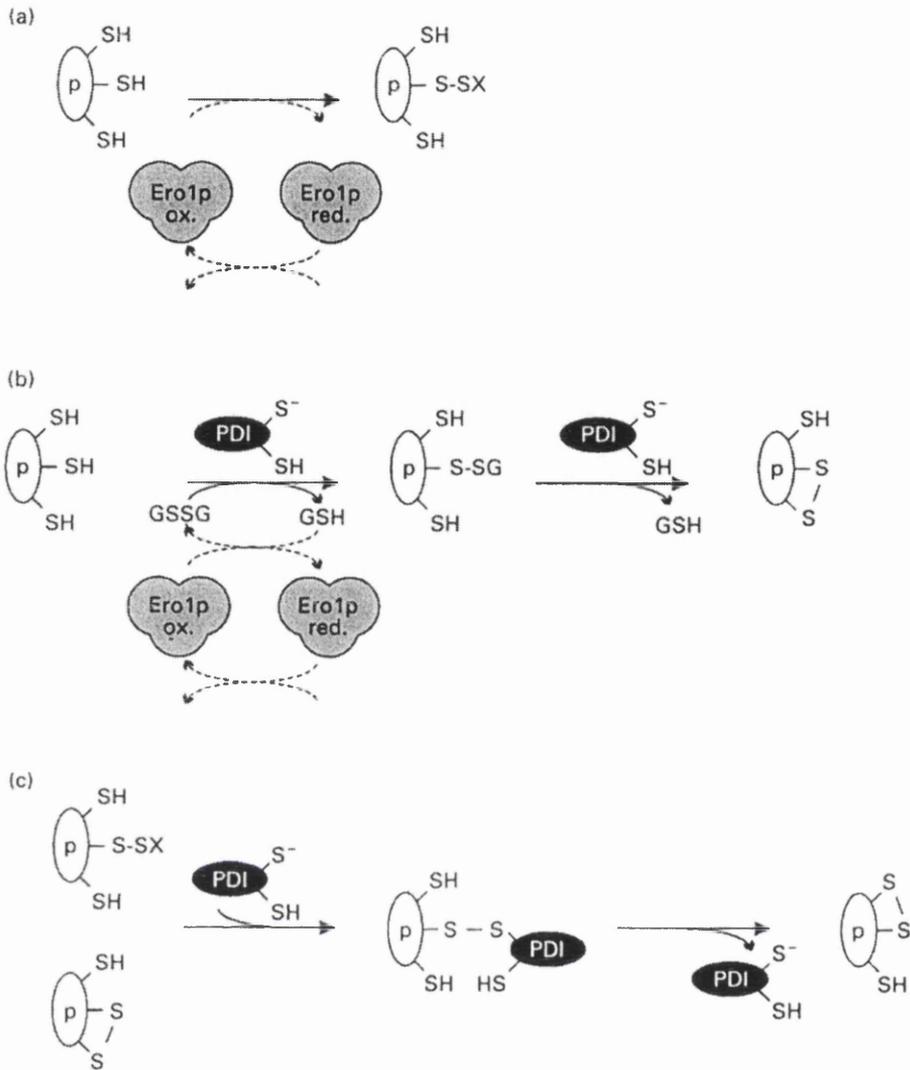


Fig 1.17. Possible mechanisms of disulphide-bond formation, reduction and isomerisation. Disulphide-bond formation in a substrate protein (p) may be catalysed directly (a) or indirectly (b) by Ero1p. PDI may also catalyse this step (b). Disulphide-bond isomerisation is catalysed by PDI using either the mixed protein/GSH disulphide (b) or some other inter- or intra-molecular disulphide (c) as substrate. S⁻ indicates the more reactive, N-terminal cysteine residue in the PDI active site. The direct oxidation pathway in (a) followed by an isomerisation pathway (c) obviates the requirement for glutathione. In this scheme, the source of the oxidising equivalents (ox) for the regeneration of oxidising Ero1p (Ero1p ox.) from the reduced form (Ero1p red.) is unknown. (Taken from Ferrari & Söling, 1999).

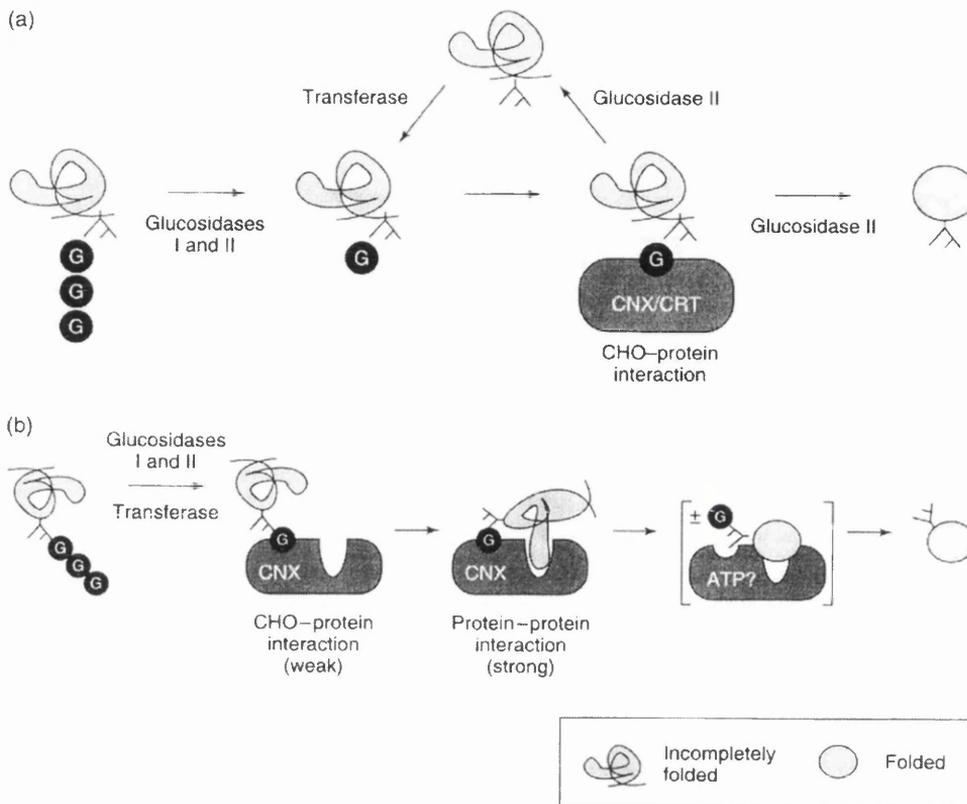


Fig 1.18. Models for calnexin (CNX) and calreticulin (CRT) binding to their substrate glycoproteins. (a) Lectin only model. Glucose residues (G) in the core oligosaccharides are trimmed. When trimmed to the monoglucosylated form, the oligosaccharides mediate the binding of the glycoprotein to calnexin and calreticulin, functioning as lectins. Release of the substrate glycoprotein depends on hydrolysis of the remaining glucose residue by glucosidase II. If the glycoproteins are not completely folded, the high-mannose glycans are selectively re-glucosylated by GT (see Fig 1.19), allowing the oligosaccharides to rebind the chaperones. Once glycoproteins reach their mature conformation, they are no longer recognised by GT. (b) Dual-mode model. The monoglucosylated oligosaccharides are merely needed to bring glycoproteins into contact with calnexin. After the initial carbohydrate-mediated association, stronger protein-protein interactions occur between the chaperone and peptide elements exposed on the surface of the incompletely folded substrate protein. These stabilise the complex until the peptides are hidden in the folded protein. Thus, release of the substrate depends on conformational changes in substrate glycoprotein that eliminate the protein-protein interaction. (Taken from Helenius *et al.*, 1997)

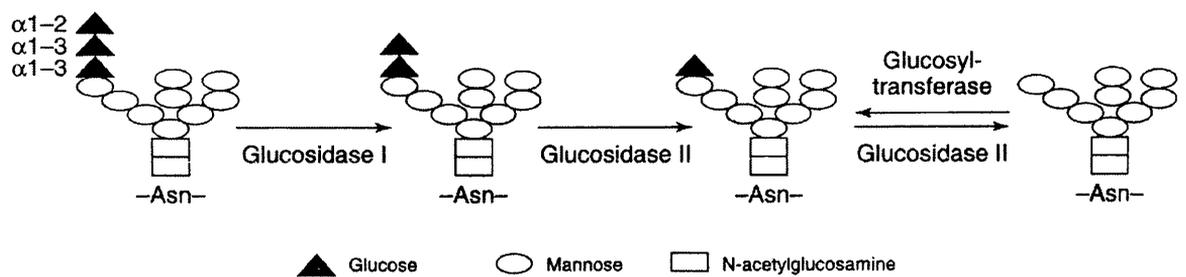


Fig 1.19. N-glycosylation and glucose trimming in the ER. The core oligosaccharide ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$) is synthesised in the ER membrane and transferred to the consensus N-glycosylation site (Asn-X-Ser/Thr) on nascent polypeptide chains. Immediately after transfer, the glucose residues are removed by the sequential action of glucosidases I and II. Glucosidase I removes the terminal $\alpha 1-2$ -linked glucose, whereas glucosidase II removes the two remaining $\alpha 1-3$ -linked residues. Fully glucose-trimmed high-mannose glycans can be re-glucosylated by UDP-glucose:glycoprotein glucosyltransferase (GT) if present on incompletely folded or assembled proteins. (Taken from Helenius *et al.*, 1997)

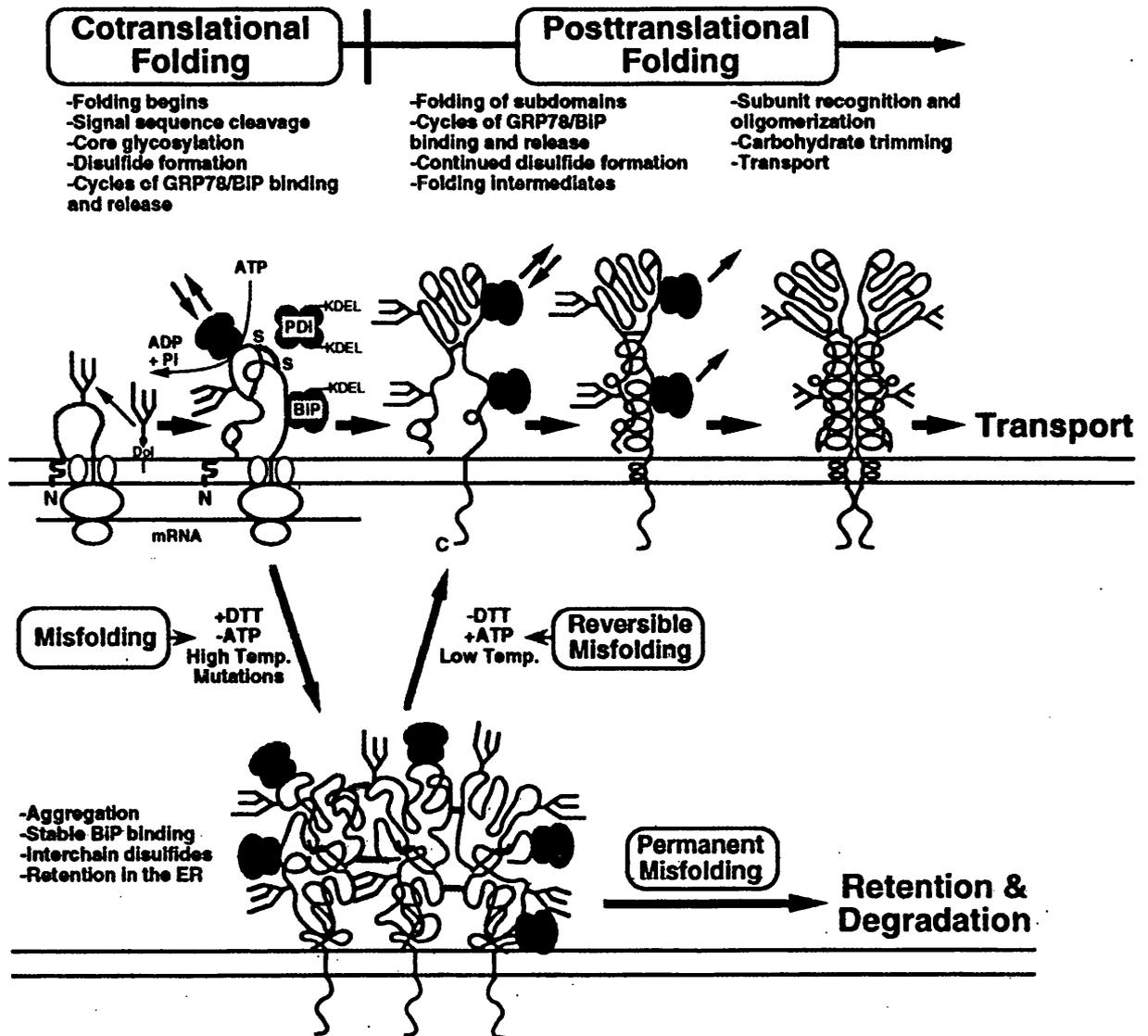


Fig 1.20. Folding and assembly of viral membrane proteins in the ER. The schematic shows the folding pathway utilised by a typical viral membrane protein. Proteins begin to fold co-translationally and N-linked oligosaccharides are added to the nascent chain from dolichol phosphate. Disulphide bonds also form (red lines), assisted by protein disulphide isomerase (PDI). The molecular chaperone BiP associates with many folding polypeptide chains at multiple sites. Other chaperones such as calnexin and calreticulin are also involved. Sequential folding from the beginning of the polypeptide chain to its end, with independently folding sub-domains, also account for the efficiency of folding *in vivo*. The third molecule shows the completed polypeptide with a folded top subdomain and a stem domain which is largely in an unfolded state. When the monomeric subunit is extensively folded, subunit-subunit recognition and assembly occurs, with dissociation of BiP. Only correctly folded and assembled proteins are generally transported to the Golgi apparatus. A typical misfolding pathway is also shown. Proteins can misfold due to inefficient folding or due to a number of factors, including mutations, heat shock, and energy depletion. Misfolded proteins often, but do not always, aggregate. The aggregates may be disulphide-bonded, as shown, and typically display permanent and stable association with BiP. Misfolded proteins are retained in the ER and are eventually degraded, unless they are rescued. Instances in which misfolding is reversible include temperature-sensitive folding mutants and situations in which the ER environment has been experimentally altered. Restoration of ATP can enable misfolded protein aggregates to dissociate to monomers which then fold normally. (Taken from Doms *et al.*, 1993)

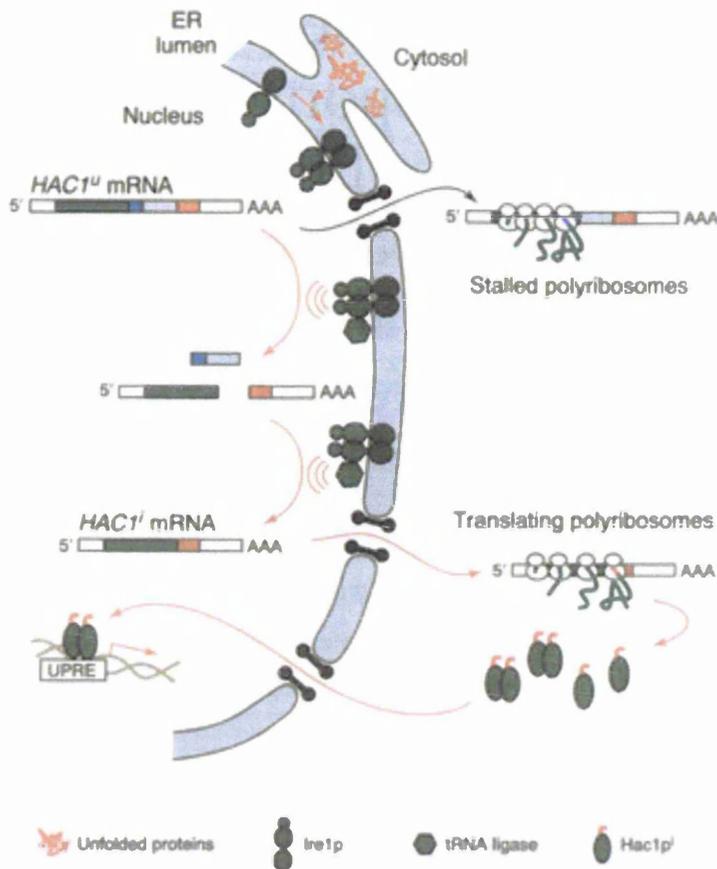


Fig 1.21. Model of the unfolded protein response (UPR) signalling pathway. Accumulation of unfolded proteins in the ER lumen triggers activation of Ire1p, which cleaves *HAC1^u* mRNA at both splice junctions. tRNA ligase then joins the two exons to produce the spliced form of the message, *HAC1ⁱ* mRNA. Both forms of *HAC1* mRNA exit the nucleus and associate with polyribosomes. However, only the spliced form gives rise to protein, Hac1pⁱ, which then enters the nucleus and upregulates genes containing the unfolded protein response element (UPRE). Whether Ire1p is indeed resident in the inner nuclear membrane as depicted in this model has not been determined experimentally and hence remains speculative. (Taken from Sidrauski *et al.*, 1998)

Chapter 2 – Materials and Methods

2.1.1. Bacterial Strains

<i>Escherichia coli</i> (<i>E. coli</i>) Strain	Phenotype
DH5 α	F ['] /endA1 hsdR17 (r _k ⁻ m ⁺) supE44 thi1 recA1 gyrA (NaI ^r) relA1 Δ (lacZYA-argF) U169 (ϕ 80dlac(lacZ)M15)
Top 10 F' (Invitrogen)	F ['] {lacI ^q , Tn10(Tet ^R)} mcrA Δ 9mrr-hsdRMS-mcrBC) ϕ 80dlac(lacZ) Δ M15 Δ lacX74 deoR recA1 araD139 Δ (ara-leu)7697 galU galK rpsL(Str ^R) endA1 nupG
GM48	dam ⁻ dcm ⁻ strain (obtained from Dr R Thompson)

2.1.2. Vectors

pGEM1 (Fig 3.3, panel a)	Promega
pSFV1 expression vector (Liljestrom & Garoff, 1991b; Fig 3.3, panel b)	Life Technologies, Inc.

2.1.3. Kits and Enzymes for DNA/Protein Modification

Advantage DNA polymerase, dNTPs, PCR buffer	Clontech
DNA purification kit	Qiagen
Endoglycosidase H _f , PNGase F	New England Biolabs
Restriction enzymes	New England Biolabs and Boehringer Mannheim
SP6 RNA polymerase, T7 DNA polymerase, <i>in vitro</i> TnT kit, RNasin, NTPs, Wizard Kit	Promega
T4 DNA ligase, alkaline phosphatase	Boehringer Mannheim

2.1.4. Cells

Proteins were expressed in BHK (clone C13) cells (Macpherson & Stoker, 1962).

2.1.5. Cell culture growth media

Glasgow minimal Eagles Medium (GMEM), new born calf serum (NBCS), penicillin/streptomycin and non-essential amino acids were supplied by Gibco-BRL.

Components of Eagles media used during radiolabelling were produced in-house by the media department, as were PBS, versene, trypsin, L-broth, and tryptose phosphate (TP) broth.

2.1.6. Radiochemicals

³⁵ S-L-Methionine	1175 Ci/mmol (10 μ Ci/ μ l)	Amersham
³⁵ S-Pro-Mix™ [³⁵ S-L-methionine/cysteine]	~1000Ci/mM (14.3 μ Ci/ μ l)	
Mannose, D-[2- ³ H(N)]	23.3Ci/mM (1 μ Ci/ μ l)	DuPont

2.1.7. Antibodies

Polyclonal antisera and MAbs against E1 and E2 (some of the characterisation of these antibodies was conducted as part of this study; summarised in Appendix 5)	A. Patel
Anti-E2 MAbs H2, H33, H47, H53	J. Dubuisson (Deleersnyder <i>et al.</i> , 1997)
Anti-gD MAb 4846	A. Cross and H. S. Marsden
MAbs E1F and E2G	T. Rice
Human MAbs 1A7, 2H1, 4F7 and 4F1	M. Cardosa (Siemoneit <i>et al.</i> , 1995)
Anti-calnexin polyclonal antibody (against the N-terminus of canine calnexin)	StressGen
Penta-His MAb	Biotechnologies Corp. Qiagen
Anti-rabbit protein-A-HRP, anti-mouse-HRP, anti-rabbit Texas Red raised in donkey, anti-mouse-FITC raised in sheep and goat, anti-rabbit Cy5 raised in goat	Sigma

2.1.8. Chemicals

The following companies were used for the supply of the indicated chemicals:

ECL, Hybond nitrocellulose membrane	Amersham
Boric acid, bromophenol blue (BPB), EDTA, glycine, sodium chloride, sodium dodecyl sulphate (SDS), sucrose	BDH
Ampicillin	Beecham Research
30% acrylamide, TEMED, Coomassie Brilliant blue and ammonium persulphate	Bio-Rad
Tris	Boehringer Mannheim
En ³ Hance	DuPont
Ammonia solution	Fisher Scientific Ltd.
Dried skimmed milk	Marvel
Ecoscint A scintillation solution	National Diagnostics
ATP (for DNA ligation)	Pharmacia
Butan-2-ol, chloroform, ethanol, glycerol, hydrochloric acid, isopropanol, methanol	Prolabo
Nickel-NTA agarose	Qiagen
Agarose, β -ME, BSA, DTT, EtBr, imidazole, NEM, PFA, phenol, PMSF, protein-A sepharose, RNase A, Triton X-100, Tween 20, cycloheximide, TCA, Trypan blue	Sigma

2.1.9. cDNA clones

Clones encoding the Glasgow strain core, E1 and E2 proteins were provided by R. M. Elliot, M. McElwee and J. McLauchlan. A plasmid, pCV-H77C was provided by J. Bukh (Ogata *et al.*, 1991; Yanagi *et al.*, 1997) that carries the full-length cDNA sequence of infectious HCV strain H77. The clone encoding the HSV-1 gD glycoprotein was provided by V. Preston.

2.1.10. Photographic reagents

Developer and fixer were bought from Kodak for the Kodak X-OMAT. Kodak X-OMAT S-Film was used.

2.1.11. Solutions

30 % acrylamide solution	30% acrylamide/0.8% (w/w) <i>N,N'</i> -methylene bisacrylamide cross-linker (ratio 37.5:1)
Agarose gel loading buffer (6x)	5xTBE, 50 % sucrose, 1µg/ml bromophenol blue
Binding buffer	25mM Tris.HCl, pH 8.0, 300mM NaCl, 1 % (v/v) Triton X-100, 1mM PMSF, 40mM imidazole, 20mM NEM
Cell lysis buffer	0.2M NaOH, 1% (w/v) SDS
Cell resuspension buffer	50mM Tris.HCl, pH 7.5, 10mM EDTA, 100µg/ml RNase A
Column wash solution	80mM potassium acetate, 8.3mM Tris.HCl, pH 7.5, 40µM EDTA, 55% (v/v) ethanol
Coomassie stain solution	Fix solution containing 2% (w/v) Coomassie Brilliant blue
Destain solution	5% (v/v) methanol and 7% (v/v) acetic acid in dH ₂ O
Eagles A	1.5mM CaCl ₂ .2H ₂ O, 1mM MgSO ₄ .7H ₂ O, 0.1ml concentrated HCl
Eagles B	50% (v/v) salts/plus, 40% (v/v) amino acids/plus, 3.2% (v/v) vitamins (gassed with CO ₂ to pH 6.5)
Eagles (low met)	7:1 (v/v) Eagles A: Eagles B (one-fifth normal concentration methionine), 2% (v/v) NBCS
Eagles (low met/cys)	7:1 (v/v) Eagles A: Eagles B (one-fifth normal concentration methionine and cysteine), 2% (v/v) NBCS
Fix solution	50% (v/v) methanol, 7%(v/v) acetic acid in dH ₂ O
Gel loading buffer (x3)	29% SGB, 6% (w/v) SDS, 2M β-ME (excluded for non-reducing conditions), 29% glycerol, 1µg/ml bromophenol blue

Immunoprecipitation (IP) buffer	50mM Tris.HCl, pH 7.6, 200mM NaCl, 1mM EDTA, 1% (v/v) Triton X-100, 20mM NEM, 1mM PMSF
L-Broth	10gr NaCl, 10gr Bactopeptone and 5gr yeast extract per litre
L-broth agar	L-broth, plus 1.5% (w/v) agar
Ligase buffer (5x)	250mM Tris.HCl, pH 7.6, 50mM MgCl ₂ , 5mM DTT, 25% PEG 8000
Neutralisation buffer	1.32M potassium acetate, pH 4.8
PBS	PBS (A) plus 6.8mM CaCl ₂ .2H ₂ O and 4mM MgCl ₂ .6H ₂ O
PBS (A)	170mM NaCl, 3.4mM KCl, 10mM Na ₂ HPO ₄ , 1.8mM KH ₂ PO ₄ , 25mM Tris.HCl pH 7.2
Resolving gel buffer	1.5M Tris.HCl, pH 8.8, 0.4 % (w/w) SDS
Running gel buffer	40mM Tris, 185mM glycine, 0.1 % (w/w) SDS
Stacking gel buffer	0.5M Tris.HCl, pH 6.9, 0.4 % (w/w) SDS
TBE (10x)	10mM Tris.HCl, pH 8.0, 0.3M boric acid, 1mM EDTA
TE	10mM Tris.HCl, pH 8.0, 1mM EDTA
Towbin buffer	25mM Tris.HCl, pH 8.3, 192mM glycine, 20% (v/v) methanol
Trypsin solution	0.25% (w/v) Difco trypsin dissolved in PBS (A), 0.005% (w/v) phenol red
Versene	0.6mM EDTA in PBS (A), 0.002% (w/v) phenol red
Wash buffer	Binding buffer containing 50mM imidazole

2.2.1. Manipulation of DNA

2.2.1.1. Small Scale Purification of DNA (mini-preps)

Single colonies of transformed bacteria were inoculated into 2ml of L-broth containing 10µg/ml ampicillin and incubated with shaking overnight at 37°C. Aliquots of 200µl of cultures were added to an equal volume of cell lysis buffer and mixed by inversion. 200µl of neutralising buffer was then added and the solution was again mixed by inversion. The cell debris was removed by centrifugation at 13 000 g for 2 min at room temperature. The supernatant was transferred to a fresh tube and mixed with 500µl of isopropanol to precipitate DNA. Plasmid DNA was pelleted by centrifugation at 13 000 g for 2 min at room temperature. The pellet was dried and resuspended in 20µl of dH₂O. 5µl of this solution of DNA was used for restriction enzyme digestions.

2.2.1.2. Large Scale Purification of DNA (maxi-preps)

200ml of L-broth containing 10µg/ml ampicillin was inoculated with bacteria and incubated with shaking overnight at 37°C. Cells were pelleted by centrifugation at 2 700 g for 10 min at 20°C. The cell pellet was resuspended in 10ml of cell resuspension solution and mixed gently with 10ml of cell lysis buffer. Following incubation at room temperature for 10 min, 10ml of neutralisation solution was added to the lysate and mixed gently for a further 10 min. Cell debris was pelleted at 13 000 g for 10 min at 20°C. The supernatant was filtered and mixed with 1/2 volume of isopropanol. After a 30 min incubation at -20°C, precipitated DNA was pelleted by centrifugation at 13 000 g for 10 min at 20°C. The DNA was resuspended in 2ml of TE buffer and mixed with 10ml of Wizard™ Maxiprep DNA purification resin. The resin/DNA mix was transferred to a maxicolumn attached to a vacuum manifold port. The resin/DNA was washed firstly with 25ml of column wash solution and secondly with 5ml of 80% ethanol. Preheated dH₂O at 65°C was applied to the resin, incubated for 1 min and the DNA was eluted by centrifugation at 2 700 g for 5 min. The eluate was filtered through a 0.22µm Micropore syringe filter and ethanol precipitated. DNA was subsequently pelleted by centrifugation at 13 000 g at 20°C, vacuum dried and resuspended in 200µl of H₂O before storage at -20°C.

2.2.3. Oligonucleotide Synthesis and Purification

Oligonucleotide synthesis was performed on an automated Cruachem PS250 Oligonucleotide Synthesiser operated by David McNab and Alex Orr. Oligonucleotides were synthesised by cyanoethyl phosphoramidite chemistry on controlled pore glass (CPG) columns. The first nucleoside of the 3' end of the oligonucleotide is attached to the CPG surface via an ester linkage and a hydrocarbon spacer. The 5'OH of the nucleoside is protected with a dimethoxytrityl group (DMT) to allow the correct polarity of chemical synthesis (3'→5'). For the sequential addition of each nucleotide the following steps were repeated; a) deprotection by removal of the 5'DMT acid labile group on the nucleoside attached to the column to generate a 5'OH, b) the phosphoramidite to be added is activated by tetrazole for formation of a covalent bond with the 5'OH group, c) chains with failed additions are capped, and d) oxidation of the phosphite internucleotide bond by iodine.

To remove synthetic oligonucleotides from the CPG, the ester linkage was cleaved with ammonia (specific gravity 0.88) resulting in a free 3'OH. The protecting groups on oligonucleotides were removed by incubation at 55°C for 5 h in ammonia and the DNA was lyophilised under vacuum. An aliquot of the lyophilised oligonucleotides was resuspended in 100µl dH₂O followed by purification using phenol/chloroform extraction. After ethanol precipitation and vacuum drying, oligonucleotides were resuspended in 200µl of dH₂O.

2.2.4. Quantitation of Plasmid DNA and Oligonucleotides

A Beckman Du®-62 Spectrophotometer was used to determine DNA concentrations by measuring the absorbance at 260nm, assuming $1 A_{260}=50\mu\text{g/ml}$ for double-stranded plasmid DNA and $1 A_{260}=20\mu\text{g/ml}$ for single-stranded oligonucleotides. In addition, purity of nucleic acids was confirmed by readings at 260nm and 280nm (A_{260}/A_{280}). A ratio of approximately 1.8 indicated that the preparation was relatively free of protein contaminants. *In vitro* transcribed RNA was also measured at 260nm assuming that $1A_{260}=40\mu\text{g/ml}$.

2.2.5. Restriction Enzyme Digestion of DNA

Small scale restriction enzyme digestion of DNA to identify recombinant plasmids was carried out at 37°C (or the temperature specified by the supplier) in 20µl volumes using 10 units of each enzyme per µg DNA in the buffer supplied with the enzyme. Reactions were incubated typically for 1 h. Large scale reactions for preparative isolation of DNA fragments were carried out in 100µl volumes using 50 units of enzyme per 15µg DNA. Digestion reactions were terminated by addition of 1/5 volume of agarose gel loading buffer. The common restriction enzymes used in this study are shown in Appendix 6 along with their recognition sequences.

2.2.6. Electrophoretic Separation and Isolation of Digested DNA

2.2.6.1. Agarose Gel Electrophoresis

DNA fragments produced by restriction enzyme digestion or PCR were resolved by agarose gel electrophoresis. Slab gels (120 x 90mm) of 1% agarose were used for separation of fragments larger than 200 bp and 1.5% agarose gels were employed for smaller fragments. Slab gels of approximately 0.5cm thick were prepared in 0.5xTBE containing a final concentration of 0.5µg/ml EtBr and run at 100V in 0.5xTBE buffer. 100 bp DNA ladder standard size markers (Promega) were used where appropriate. Following electrophoresis, DNA was visualised either under short-wave UV light or, for preparative gels, under long-wave UV light. Photography was carried out using an Appligene Imager.

2.2.6.2. Purification of DNA from Agarose Gels

Agarose slices containing appropriate DNA fragments were excised from gels under long-wave UV transillumination and the DNA recovered using a QIAGEN gel purification kit. Excised agarose slices were dissolved in QC buffer (provided by supplier) and incubated at 50°C until the gel slice had dissolved. To aid purification, a volume of isopropanol, equivalent to the weight of the gel slice, and 5µl of 3M sodium acetate (pH 5.0) was added to the solution of dissolved gel. This solution was applied to a QIAGEN column by centrifugation at 7 000 g. DNA on the column was washed with 500µl QC buffer and

750µl of a solution QE (provided by supplier) containing 80% ethanol. DNA was eluted with 30µl of dH₂O by centrifugation (10 000 g for 1 min) into a fresh microfuge tube.

2.2.6.3. Purification of Linearised DNA and PCR Fragments

Linearised DNA and PCR fragments were purified using the QIAGEN Nucleotide purification kit. The restriction enzyme digestion was mixed with column application buffer and applied to the QIAGEN column by centrifugation at 7 000 g. Bound DNA was washed with 750µl of solution QE containing 80% ethanol and eluted with 30µl dH₂O as described in Section 2.2.6.2.

2.2.6.4. Phenol/Chloroform Extraction

Proteins were removed from restriction enzyme digestions by adding an equal volume of a 1:1 solution of phenol/chloroform. This mixture was then vigorously mixed and the organic layer containing the proteins was separated from the aqueous layer containing the nucleic acids by centrifugation at 13 000 g for 1 min. The aqueous layer was transferred to an equal volume of chloroform and mixed to remove any residual phenol. Following centrifugation, the aqueous layer was subjected to ethanol precipitation for concentrating the nucleic acids.

2.2.6.5. Ethanol Precipitation

Nucleic acids were precipitated by mixing with ethanol and 5M ammonium acetate to achieve final concentrations of 70% and 0.5M respectively. Following precipitation at -20°C for at least 1 h, precipitated nucleic acids were collected by centrifugation at 13 000 g for 10 min, dried and resuspended in 10-50µl of dH₂O.

2.2.7. Ligation Reactions

Purified vector and DNA fragments, prepared by restriction enzyme digestion and purification, were ligated in a 1:3 ratio. Reactions were carried out in a 25µl volume of 1xligase buffer with 2 units of T4 DNA ligase containing 1mM ATP. Reactions were

carried out at 16°C for 15-20 h. Ligated DNA was ethanol precipitated and resuspended in 5 µl of dH₂O prior to transformation of competent *E.coli*.

2.2.8. Preparation of Competent *E coli* Cells

2.2.8.1. Electrocompetent DH5α

Plasmids were mostly grown and maintained in *E.coli* strain DH5α. To prepare electrocompetent bacteria, a 50ml overnight culture grown in L-broth from a single colony was used to inoculate 1l of L-broth. This culture was grown for 3-3.5 h at 37°C in a shaking incubator, until the OD₅₅₀ was approximately 0.5-0.6. The culture was cooled on ice for 30 min and the cells were pelleted by centrifugation at 2 700 g for 15 min at 4°C. Cells were washed twice with ice-cold, sterile dH₂O and pelleted before resuspending in ice-cold sterile dH₂O containing 10% glycerol. Cells were pelleted again at 2 700 g for 15 min at 4°C. The pellet was finally resuspended in 4ml dH₂O containing 10% glycerol and aliquoted in 40µl volumes for storage at -70°C.

2.2.8.2. Competent GM48

Plasmid DNAs lacking dam/dcm methylation modification (required for certain restriction enzymes) were grown in *E.coli* strain GM48. A 2ml culture grown overnight in L-broth from a single colony was used to inoculate 100ml of pre-warmed L-broth. This culture was grown in a shaking incubator at 37°C to an OD₂₈₀ = 0.5-0.6 to obtain bacterial cells that were growing during the log phase. Following a 5 min incubation of the bacterial culture on ice, the cells were pelleted at 2 700 g for 10 min at 4°C. The cell pellet was resuspended in 50ml ice-cold 100mM CaCl₂ and incubated on ice for 2 h. The cells were pelleted, resuspended in 1ml of ice-cold 100mM CaCl₂ and stored on ice until transformation with plasmid DNA.

2.2.9. Transformation of Competent *E coli* Cells

2.2.9.1. Electroporation of DH5α and TOP10F' Cells

An 80µl aliquot of electrocompetent DH5α was thawed on ice before the addition of 2µl of the products of a ligation reaction. This mixture was transferred to a cold electroporation cuvette (0.1cm gap, Bio-Rad) and electroporated by a single pulse at 1.6kV and 25µF (Bio-

Rad Gene Pulser II). The cells were resuspended in 1ml of L-broth and incubated at 37°C for 1 h to allow expression of the antibiotic resistance gene. 100µl of this culture was plated on to a L-broth agar plate containing 100µg/ml ampicillin and incubated at 37°C overnight.

Occasionally, commercially available TOP10 F' cells (Invitrogen) were used to obtain plasmid DNA. For transformation of TOP10 F' cells, ligation reactions were ethanol precipitated and the pelleted DNA was resuspended in 5µl of dH₂O. A 2µl aliquot of the ligation products was mixed with 20µl of electrocompetent TOP10 F' cells and incubated on ice for 1 min. The transformation mixture was transferred to an ice-cold electroporation cuvette (0.1cm gap) and electroporated with a single pulse at 1.6kV and 25µF. The cells were resuspended in 480µl of L-broth media and incubated at 37°C for 1 h with shaking. A 50µl aliquot of transformed cells was plated onto a L-broth agar plate containing 100µg/ml ampicillin and incubated at 37°C overnight.

2.2.9.2. Transformation of GM48 Bacteria

1µg of purified plasmid DNA was added to 100µl of competent cells (prepared as in 2.2.8.2) and the mixture was incubated on ice for 40 min. Transformed bacteria were placed at 42°C for 4 min before plating out on L-broth agar plates containing 100µg/ml ampicillin. Plates were placed at 37°C overnight.

2.2.10. Automated DNA Sequencing

Automated DNA sequencing was carried out (by L. Taylor) using an ABI PRISM BigDye™ terminator cycle sequencing ready reaction kit. Sequencing was performed using an ABI PRISM™ 377 DNA Sequencer.

2.2.11. PCR Amplification of DNA

A PCR kit from Clontech was used for all PCR reactions. Typically reactions were performed in a final volume of 50µl, containing 1xPCR reaction buffer (provided by supplier), 0.2mM dNTPs, and 2.5 units Advantage DNA polymerase. The DNA template

and primers were used at final concentrations of 0.5ng and 1pg per μl respectively. Prior to the PCR cycles, reactions were typically heated at 94°C for 3 min to allow complete denaturation of the nucleic acids. PCR reactions were carried out under the following reaction conditions: a) strand-separation at 94°C for 30 sec, b) primer annealing at 56-60°C (temperature varied depending on primer) for 30 sec, c) strand elongation from primers at 72°C for 30-60 sec (times varied with length of PCR product) for 30 cycles, d) a final elongation step at 72°C for 5 min. PCR reactions were carried out in a Techne (Progen) PCR machine.

2.2.12. *In Vitro* Transcription of Linearised DNA

Recombinant pSFV1 DNAs (15 μg) were linearised with *SpeI* and purified by phenol/chloroform extraction and ethanol precipitation. After pelleting, DNA was resuspended in 15 μl dH₂O. *In vitro* transcription reactions were typically performed in a final volume of 50 μl as follows: NTP mix (1mM ATP, UTP, CTP, 0.5mM GTP), Transcription buffer (40mM Tris, pH 7.9, 6mM MgCl₂, 2mM spermidine and 10mM NaCl), 1mM m⁷G(5')ppp(5')G RNA cap analogue, 1mM DTT, 2 units/ μl human placental ribonuclease inhibitor (RNasin), 2 units/ μl SP6 RNA polymerase (as described by Liljestrom & Garoff, 1991). 1.5 μg of linearised template DNA was added to reactions and incubated at 37°C for 1-2 h. 2 μl of the *in vitro* transcription reaction was analysed on a 1% agarose gel before proceeding to use the RNA for electroporation into BHK cells.

2.2.13. Maintenance of BHK Cells

BHK cells were cultured in GMEM media containing 10% new born calf serum (NBCS), 4% TP broth and 100 units/ml of penicillin/streptomycin (ETC10) at 37°C with 5% CO₂ in 160cm² tissue culture flasks. Cells were harvested at confluency, with a yield of 2-4x10⁷ cells per flask.

2.2.14. Electroporation of Recombinant SFV RNA

Cells approaching confluency in a 160cm² flask were washed with versene at room temperature and removed by trypsin treatment. Approximately 10ml of ETC10 media was

added to the detached cells and a single cell suspension was achieved by pipetting several times with a narrow hole 5ml pipette. The cells were pelleted by centrifugation at 400 g for 3 min at 15°C. The cell pellet was resuspended in 10ml PBS (A) and pelleted as before. BHK cells were finally resuspended in PBS (A) to a final concentration of $\sim 1 \times 10^7$ cells/ml. 0.8ml of resuspended cells was placed in an electroporation cuvette (0.4cm gap) with 20 μ l of recombinant pSFV RNA. The cell/RNA suspension was mixed by inversion and electroporated at 1.2kV and 25 μ F (Bio-Rad Gene Pulser II) at room temperature. The cells were pulsed twice with an inversion step between the pulses. Electroporated cells were transferred to 10ml of ETC10. Thereafter, 35mm tissue culture dishes were seeded with 2×10^6 cells (in 3ml medium) and 13mm coverslips in 24-well plates were seeded with 10^5 cells (in 0.8ml medium). Cells were incubated at 37 °C.

2.2.15. Trypan Blue Staining

A 500 μ l aliquot of electroporated cell suspension was diluted with 0.5% trypan blue stain in PBS at a ratio of 1:1 and incubated for 5 min at RTM after gently mixing. The exclusion of this dye from live cells allowed stained dead cells to be counted under a light microscope.

2.2.16. Metabolic Labelling of Electroporated Cells

2.2.16.1. Radiolabelling with ^{35}S -amino acids

For longer periods of radiolabelling, cells were incubated at 37°C for 4 h then medium was replaced with Eagles medium containing one-fifth normal concentration of methionine [Eagles (low met)]. Following further incubation at 37°C for 30 min, medium was again replaced with Eagles (low met) containing 10 μ Ci/ml ^{35}S -methionine. Cells were radiolabelled for 12 h at 37°C before harvesting (see Section 2.2.17).

For pulse/chase experiments, cells were radiolabelled 12 h after electroporation. Firstly, media was replaced with Eagles containing one-fifth normal concentration methionine and cysteine [Eagles (low met/cys)], and incubated for 30 min at 37°C. Cells were then pulse-labelled for 2-15 min with fresh Eagles (low met/cys) containing 400 μ Ci/ml ^{35}S -methionine/cysteine (Pro-Mix). To chase radiolabelled proteins, labelling media was

replaced with Eagles media containing 5x excess methionine/cysteine and incubation at 37°C was continued for up to 80 min depending on the nature of the experiment.

2.2.16.2. ³H-Mannose Labelling

4 h post-electroporation, cell culture media was replaced with PBS containing 2% (v/v) NBCS, and 1% non-essential amino acids (PBSM). Following a 30 min incubation period at 37°C, this solution was replaced with PBSM containing 100µCi/ml ³H-mannose and cells were incubated again at 37°C for 12 h.

2.2.17. Preparation of Cell Extracts and Isolation of Protein from Eukaryotic Cells

Cells were washed twice with 1ml ice-cold PBS, containing 20mM N-ethylmaleimide (NEM), an alkylating agent which blocks free sulphydryl groups on cysteine residues. Cells were lysed using 500µl of either binding buffer or immunoprecipitation buffer also containing 20mM NEM. Complete lysis was achieved by incubation on ice for 10 min and the cell debris was pelleted by centrifugation at 13,000 g for 10 min at room temperature. The cleared supernatants were transferred to fresh microfuge tubes. Cell extracts were either used directly (e.g. for Western blot analysis) or proteins of interest were isolated according to either of the following procedures.

2.2.17.1. Immobilised-Metal Affinity Chromatography (IMAC)

Ni-NTA agarose was equilibrated by washing three times in binding buffer and resuspended to achieve a 50% slurry. For the isolation of histidine-tagged protein, 100µl of equilibrated Ni-NTA agarose was added to supernatants from lysed cells and the binding was allowed to occur for 1 h at 4°C on a rotator. Ni-NTA agarose was pelleted by brief centrifugation to 1 000 g and washed three times with 500µl of wash buffer. Bound proteins were eluted with 20µl of gel loading buffer.

2.2.17.2. Immunoprecipitation

For immunoprecipitation, 1-5 μ l of the appropriate antibody was added to cell lysates and the mixture incubated at 4°C overnight. 10mg of protein-A sepharose per precipitation was allowed to swell and equilibrated overnight in IP buffer at 4°C. Prior to the addition of 100 μ l of equilibrated protein-A sepharose to the cell lysate-antibody mixture, the equilibrated protein-A resin was washed three times in IP buffer and resuspended in an equal volume of the same buffer. The protein-A-antibody binding was allowed to occur for 1 h at 4°C on a rotator. Subsequently, the washing (using IP buffer) and protein elution stages (using gel loading buffer) were performed as for those described for IMAC.

2.2.18. TCA Precipitation

An aliquot of radiolabelled crude cell extract was blotted onto filter paper discs and placed in 15% TCA. Radiolabelled proteins were precipitated for 20 min and then the discs were washed twice in methanol before being placed in 1ml of scintillation solution. TCA precipitated samples were counted using a PACKARD Liquid Scintillation Analyser 1600 TR.

2.2.19. Deglycosylation of Proteins Using Endo H_f and PNGase F

Lysates from cells were denatured in 0.5% (w/v) SDS, 1% (v/v) β -ME at 100 °C for 10 min. Sodium citrate was added to 50mM followed by 2000 units of endo H_f or 1000 units of PNGase F. For PNGase F digestions 1% (v/v) NP-40 was added to the reaction mix following the denaturation step to inactivate SDS. The reactions were incubated at 37°C for 1 h.

2.2.20. Coupled *In Vitro* Transcription and Translation (*In Vitro* TNT)

Cell free protein synthesis was carried out using the TNT coupled transcription/translation system from Promega. Reactions were set up in a final volume of 25 μ l containing 12.5 μ l

TnT rabbit reticulocyte lysate, 10xTnT, 2 units/ μ l T7 RNA polymerase, 1mM amino acid mixture minus methionine, 1 μ Ci/ μ l 35 S-methionine, 2 units/ μ l RNasin and 1 μ g of linearised template DNA. Recombinant pGEM constructs containing E1 and E2 sequences were linearised at the *Sall* site in E1 (see Appendix 1). Reactions were carried out at 30°C for 60 min and stopped by the addition of 15 μ l gel loading buffer before analysis by SDS-PAGE.

2.2.21. Protein Analysis by SDS-PAGE

Proteins were resolved by electrophoresis through SDS polyacrylamide gels using the discontinuous buffer system (Laemmli, 1970) and a Bio-Rad Miniprotein II apparatus. The glass plates (10 x 8cm) were assembled according to the manufacturer's instructions using 1.5mm wide spacers. The cavity created by this gel assembly could hold approximately 8ml of resolving gel. Resolving gel mix was prepared containing 1 x resolving gel buffer and 10 or 12% (v/v) acrylamide solution in dH₂O. Acrylamide was diluted from a 30% stock solution consisting of acrylamide and *N,N'* methylene bis-acrylamide (ratio 37.5:1). 12% gels were used to analyse endo H_r and PNGase F-treated material while untreated material was analysed on 10% or 12% polyacrylamide gels. The cross-linking of acrylamide and bisacrylamide was activated by the addition of 0.06% (v/v) APS and 0.05% TEMED. This solution was immediately poured into the gel assembly to approximately 2cm from the top of the plates and overlaid with butan-2-ol. After polymerisation the top surface of the resolving gel was washed with resolving gel buffer.

Stacking gel mix was prepared containing 1 x stacking gel buffer and 5% acrylamide solution in H₂O. Polymerisation was initiated by addition of 0.08% (v/v) APS and 0.08% TEMED. This mixture was placed on the resolving gel and a 1.5mm wide comb was inserted to allow loading wells to form within the polymerised stacking gel. Following polymerisation, the comb was removed and the wells were washed with running buffer. The gel was placed in a tank and the reservoirs filled with running buffer.

Samples that were analysed under reducing electrophoretic conditions alone were denatured and reduced in gel loading buffer containing β -ME by heating to 95°C for 5 min before loading. Alternatively, for examination of proteins under reducing and non-reducing conditions, samples were initially denatured in gel loading buffer lacking β -ME

by heating at 95°C for 5 min. Subsequently, the samples were divided into two parts and DTT to a final concentration of 20mM was added to one part. The reduction of disulphide bonds on proteins in this sample by DTT was promoted by a further incubation step at 95°C for 5 min before loading. When reduced and non-reduced samples were analysed on the same polyacrylamide gel, NEM to a final concentration of 20mM was added prior to loading to prevent reduction of non-reduced samples by diffusion of the reducing agent.

Molecular weight markers (Bio-Rad) were resolved alongside the samples to allow protein size determination. Gels were run at 150V until the bromophenol blue dye reached the bottom of the resolving gel. Thereafter, the gels were removed from the apparatus and treated in either Coomassie Brilliant blue stain solution for 15 min followed by destaining in destain solution. Alternatively, gels were immersed in fix solution for 30 min then in En³Hance for 30 min and finally washed in water for 15 min. Subsequently, destained or gels treated in En³Hance were dried under vacuum and exposed to Kodak XS-1 X-ray film for the detection of radiolabelled proteins.

2.2.22. Western Blot Analysis

2.2.22.1. Electroblothing to Nitrocellulose

Proteins resolved by SDS-PAGE were transferred to nitrocellulose membrane essentially following the method of Towbin *et al.* (1979) in a Bio-Rad mini transblot apparatus. A blotting sandwich was set up with the gel in contact with the nitrocellulose membrane and both were placed between sheets of Whatman 3mm paper cut to the same size as the gel. This was in turn placed between fibre pads. All materials were soaked in Towbin buffer prior to assembly of the blotting layers. Electrotransfer was carried out at 50mA for 4 h at 4°C.

2.2.22.2. Immunodetection

After transfer, nitrocellulose membranes were blocked at 4°C overnight in 5% skimmed milk (Marvel) prepared in PBS (A). Membranes were then washed three times in PBS (A) containing 0.05% (v/v) Tween-20 (PBST) for 10 min per wash prior to incubation with the primary antibody (diluted in 1% BSA in PBST) for 2 h. The anti-E2 MAbs (ALP98) was

used at a dilution of 1:1000, anti-E2 antiserum (R141) at 1:200, anti-E1 antiserum (R528) at 1:200, Penta-His MAb at 1:500, and gD MAb (4846) at 1:2000. Membranes were then washed as described above, before their incubation for 1 h with either anti-mouse IgG or protein-A conjugated to horseradish peroxidase diluted in 1% BSA in PBST. Following a final series of three washing steps, proteins on the membranes were detected using the Amersham enhanced chemiluminescence (ECL) system. The chemiluminescence reagent was prepared by mixing the two supplied reagents at a 1:1 ratio. This mixture was placed on to the membranes and agitated for 3 min. Finally, membranes were wrapped in cling film and exposed to Kodak XS-1 film for 5-45 sec. All the washing and incubations procedures were performed at room temperature with shaking.

2.2.23. Removal of Antibodies and Immunoglobulins from Membranes

Blots were stripped before probing with a different primary antibody by incubation at 55°C for 45 min in 60mM Tris, 10mM β -ME and 2% (w/v) SDS. Membranes were then washed three times for 10 min per wash in PBST at room temperature with shaking, before being placed in blocking solution and re-analysed with another primary antibody as described above.

2.2.24. Immunofluorescence

Electroporated cells on coverslips were fixed in methanol (at -20°C) and incubated for 20 min at -20°C. Alternatively, non-permeabilised cells were prepared by fixing with paraformaldehyde (PFA) for 30 min at 4°C. Permeabilised cells were fixed under the same conditions in PFA containing 0.1% (v/v) Triton X-100. Cells were rehydrated by incubation for 10 min with PBS and blocked with PBS containing 2% (v/v) NBCS (PBSN) for 10 min. 200 μ l of the relevant primary antibody diluted in PBSN was added to cells and incubated for 1.5 h at room temperature. The anti-E2 MAb was used at a dilution of 1:200, anti-E1 antiserum at 1:100, anti-gD MAb was used at 1:400 and anti-calnexin antiserum was used at 1:100. Cells were washed four times for 5 min per wash with PBSN at RTM and then incubated with diluted secondary antibody (200 μ l). Anti-mouse IgG (conjugated with FITC raised in sheep) and anti-rabbit IgG (conjugated with Texas red raised in donkey), both were used at a dilution of 1:100. For double-labelling, cells were incubated with two primary antibodies and as secondary antibodies anti-mouse conjugated to FITC

(diluted to 1:100) and anti-rabbit conjugated to Cy5 (diluted to 1:400) both raised in goat, were employed. Cells were again washed as described above followed by a final wash in dH₂O to remove residual saline solution. Coverslips were removed from the wells, gently dried and mounted onto slides with 5µl of Citifluor mounting fluid (supplied by Chem. Lab. Canterbury).

Fluorescence microscopy was performed using a Nikon MICROPHOT-SA microscope and confocal microscopy was performed using a ZEISS LSM 510 microscope and accompanying software.

2.2.25. Preparation of PFA Fixative

A quantity of PFA to make a final solution of 4% was heated to 60°C in a volume of dH₂O slightly less than 2/3 the desired final volume of fixative. This temperature was maintained while 1 drop of 2M NaOH was added to the solution to promote the solubilisation of PFA. Subsequently 1/3 volume of 3x PBS was added and the pH of the solution was adjusted to 7.2 with 1M HCl. The final volume was made up by the addition of dH₂O and the solution was filtered using a 0.22µm Millipore filter. The solution was cooled on ice before applying to cultured cells.

2.2.26. Computer Software

An AGFA StudioScan II si scanner and Adobe Photoshop 4 and Adobe Illustrator 7.0 programs were used to process the figures. EndoNote 2 Plus was used for creating a bibliography. Analysis of DNA and protein sequences was carried out using the programs in the Genetic Computer Group sequence analysis package version 7. The ExPASy-PeptideMass program (<http://www.expasy.ch/cgi-bin/peptide-mass.pl>) was used to predict molecular weights.

Chapter 3 – Establishment of the SFV Expression System

3.1. Introduction

In the absence of any efficient *in vitro* HCV replication system, studies conducted on the glycosylated proteins E1 and E2 have employed various heterologous expression systems. Throughout this study, the SFV system was used to achieve expression of the HCV glycoproteins. No previous publications have described the use of this system to investigate HCV proteins. However, several studies have employed successfully SFV vectors to express proteins encoded by other viruses including the neuraminidase and HA proteins of influenza virus (Zhou *et al.*, 1994; Berglund *et al.*, 1998). The SFV vector offered a number of advantages and part of the project was to establish the suitability of the system for expression purposes.

At the start of the project, antibodies against the glycoproteins of strain Glasgow had not been raised and initial characterisation of the expressed proteins was performed with limited amounts of reagents gifted by others. Since one of the primary aims of the project was to study interactions between E1 and E2, substantial amounts of reagents were necessary to conduct appropriate assays. In light of this, a tag consisting of six histidine residues was fused to the N-terminus of E1. This tag provided a means to purify the protein via IMAC (Fig 3.1; Porath, 1975). In this study Ni-NTA agarose was employed which is a complex of Ni-NTA coupled to Sepharose® CL-6B via a spacer arm (Fig 3.2.).

The aim of this chapter is to describe the construction of a recombinant pSFV plasmid encoding E1 and E2 as a polyprotein with a histidine tag fused to the N-terminus of E1. This construct was used, firstly to optimise the SFV replicon system for the expression of these proteins and secondly to establish whether complex formation occurred between E1 and E2 from strain Glasgow. Some preliminary characterisation of E1 and E2 from this strain was also carried out to compare their behaviour with those of other strains described by other investigators. For completion purposes, some of the experiments described in this chapter make use of antibodies against the glycoproteins of strain Glasgow which subsequently became available and were kindly provided by Dr. A. Patel.

3.2. Construction of Histidine-Tagged E1 (E1_{his})

Plasmids containing the coding region for the core, E1 and E2 proteins of HCV strain Glasgow were obtained by combining fragments from two constructs called core.pTZ18 and 5'-ΔNS2 (provided by M. McElwee and R. Elliott). Core.pTZ18 possesses nucleotide residues 337-915 of the HCV strain Glasgow genome and 5'-ΔNS2 contains residues 1-2895. DNA fragments from these plasmids were combined in a vector called pGEM1 (Fig 3.3, panel a). The resultant plasmid contained nucleotide residues 337-2895 of the HCV strain Glasgow genome and therefore encodes the core, E1, E2, and p7 proteins as well as the N-terminal region of NS2 of this strain (Appendix 1). For further cloning purposes, *Bgl*III sites were introduced at *Eco*RI and *Hind*III sites which were upstream and downstream from the HCV sequences respectively. The resultant plasmid was called pgHCV.CE1E2. These flanking *Bgl*III sites permitted convenient sub-cloning of the resulting fragments into the pSFV1 vector at a unique *Bam*HI site (Fig 3.3, panel b). A construct expressing a histidine-tagged form of E1 was made using the following strategy. Firstly, the HCV coding sequences from E1 to NS2 (amino acids 195-851), were excised as a *Fsp*I-*Hind*III fragment. This fragment was reinserted into pGEM1 along with an oligonucleotide encoding the histidine tag and a unique *Kpn*I site to allow for the introduction of the E1 signal sequence (Fig 3.4). Subsequently, oligonucleotides encoding amino acids 169-192 of the signal sequence for E1 were inserted (Fig 3.5). The *Bgl*III sites engineered into the pGEM1 vector, which flanked the sequences encoding E1-E2-p7-NS2 ORF, were then employed to generate a fragment that was directly introduced at the *Bam*HI site in the pSFV1 vector; this gave a construct called pSFV/E1_{his}E2. A recombinant SFV construct encoding the untagged Glasgow HCV glycoproteins was named pSFV/E1E2_{Gla} and was provided by J. McLauchlan. The only difference between these two plasmids was the absence of histidine tag sequences in E1 in pSFV/E1E2_{Gla}.

3.3. *In Vitro* Transcription of SFV RNA

The SFV replication system relies on *in vitro* transcription of RNA from linearised recombinant vector template, which, following transfection into eukaryotic cells, is replicated *in cis* by encoded SFV non-structural proteins (Section 1.6). The recombinant

template DNA is linearised at the *SpeI* site situated downstream of the SFV genomic poly(A) tail (Fig 3.3, panel b). Nucleotide sequences beyond the poly(A) stretch direct the synthesis of anti-sense RNA and two promoters for 43S and 26S RNAs direct transcription of positive-sense molecules. Thus, input RNA molecules introduced into cells are amplified and two populations of mRNA are synthesised.

In vitro transcription reactions were set up according to details provided in section 2.2.11, where 1.5µg of linearised template DNA was used per reaction. The amount of RNA consistently synthesised *in vitro* was approximately 60µg per 50µl reaction (as determined by absorbance spectrometry). A 2µl aliquot from a typical transcription reaction analysed on a 1% agarose gel and compared to approximately 1µg of linearised template DNA showed a major RNA band representing SFV transcripts synthesised from the SP6 RNA polymerase promoter to the *SpeI* cleavage site, a length of approximately 11kb (Fig 3.6). The minor bands and smeary appearance of the RNA are presumed to be a consequence of a combination of partial degradation, incomplete transcription and higher order structures formed by the RNA. In addition, RNA does not migrate as a tight band through an agarose matrix due to the secondary structures adopted by individual molecules.

3.4. Effect of Electroporation on Survival of BHK Cells and Efficiency of Electroporation of SFV RNA

As mentioned previously the SFV system had not been fully established in the laboratory at the start of this project. Although a number of studies elsewhere had used this system successfully in BHK and COS-K1 cells (Liljestrom & Garoff, 1991b; Berglund *et al.*, 1993), the optimal conditions for the use of BHK cells for this study had to be established. Initially, the affect of electroporation on BHK cells was examined. The trypan blue dye exclusion method, which specifically stains dead cells, was used to determine the amount of cell death that occurred after each of the pulse steps during electroporation. Data presented in Table 3.1 shows that from $\sim 1 \times 10^7$ cells used per electroporation reaction, approximately 8% of cells are stained by trypan blue at the first pulse step and 18% after the second pulse. Thus, approximately 70% of cells survive the electroporation conditions that were adopted (i.e. 2 pulses).

In order to identify the optimal amount of RNA required to achieve highest transfection efficiency, 7.5, 15, 30 and 60 µg of RNA transcribed from pSFV/E1_{his}E2 construct were electroporated into competent BHK cells (Table 3.2). Expression levels of HCV E2 were detected in these electroporated cells by indirect immunofluorescence using an E2-specific antibody, ALP98 (Fig 3.7). Comparison of cells showing fluorescence to those observed under phase contrast indicated that less than 50% of cells expressed the protein with 15µg or less RNA. By contrast, expression of E2 was found in over 80% of cells when they were electroporated with 30µg or 60µg RNA (Table 3.2.). These transfection efficiencies were reproducibly achieved, demonstrating that half (25µl) of the *in vitro* transcription reaction was sufficient to obtain high levels of electroporation. It was estimated that about 3×10^5 molecules of RNA per cell were available using 25µl of an *in vitro* transcription reaction.

3.5. Solubilisation of HCV Glycoprotein E1_{his}

A hydrophilicity prediction profile (Fig 3.8) indicated that both E1 and E2 have a C-terminal hydrophobic region that constitutes their transmembrane domains and thus they are presumed to be type 1 membrane-bound proteins. Conditions allowing the release of the HCV glycoproteins from associated membranes were established by examining their solubility following cell lysis by three detergents; deoxycholic acid, an ionic detergent, and NP-40 and Triton X-100, which are both non-ionic detergents. Cells electroporated with pSFV/E1_{his}E2 RNA were lysed in 500µl of Tris-salt buffer [25mM Tris.HCL (pH 8.0), 300mM NaCl] containing these detergents at a final concentration of 0.1%, and the samples were spun at 13 000g for 10 min at room temperature to separate the cell debris from the soluble material. The cell debris was resuspended in 500µl gel loading buffer and equivalent portions of this material were separated by SDS-PAGE alongside the soluble fraction. The samples were then probed for histidine-tagged E1 using a Penta-His antibody by Western blot analysis (Fig 3.9, panel a). This monoclonal antibody detected a protein of approximately 30 kDa that corresponds to the size of E1 identified in the literature (Grakoui *et al.*, 1993b; Ralston *et al.*, 1993). The data showed that non-ionic detergents NP-40 and Triton X-100 at concentrations of 0.1% released E1 from the membranes into the soluble material more efficiently (panel a, compare lanes 5 and 7 with lanes 6 and 8), than deoxycholic acid (panel a, lanes 3 and 4). The use of Triton X-100 in buffer solutions was further characterised.

The optimal detergent concentration to be included in any buffer used for purification of the glycoproteins (Section 2.2.15) was determined by testing Tris-salt buffers containing Triton X-100 at concentrations of 0.01, 0.1 and 1%. The cell debris was again separated from the soluble material and an equal volume of both fractions was analysed by Western blotting, once again probing for the histidine tag in E1 (Fig 3.9, panel b.). The buffer containing 1% Triton X-100 was most effective for releasing E1 from membranes (lane 6 compared with lane 5). At this stage it was assumed that the efficient solubilisation of E1 achieved with buffer containing 1% Triton X-100 was sufficient also for the solubilisation of E2, and this was confirmed in later experiments (Fig 3.15.).

3.6. Radiolabelling of HCV Glycoproteins and Immunodetection of E2

The previous section has shown the detection of E1 by indirect immunofluorescence and Western blot analysis. The expression of this glycoprotein, and E2, was further examined by radiolabelling with ^3H -mannose and ^{35}S -methionine. For ^3H -mannose, the radioactive sugar residue is incorporated during the biosynthesis of high-mannose oligosaccharide by oligosaccharyl transferase (Kornfeld & Kornfeld, 1985). In turn, these structures are transferred *en bloc* by the enzyme to the asparagine residue at N-linked glycosylation sites of only those proteins translocated to the ER (Section 1.4.5). Thus, this could provide a highly specific means of detecting the E1 and E2 glycoproteins. By contrast, ^{35}S -methionine is directly incorporated into the polypeptide chain (Fig 3.10).

Cells electroporated with pSFV/E1_{his}E2 and pSFV/E1E2_{Gla} RNA or mock-electroporated cells were incubated overnight in media containing either ^3H -mannose or ^{35}S -methionine as described in section 2.2.16. These were lysed and the level of radioisotope incorporation was quantified by TCA precipitation followed by scintillation counting (Section 2.2.18.). Equivalent TCA precipitated radioactive counts were analysed by SDS-PAGE for both the ^3H -mannose- and ^{35}S -methionine-labelled samples (Fig 3.11). The ratio of TCA precipitable counts between ^3H -mannose- and ^{35}S -methionine-labelled crude extracts was approximately 1:14. Crude extracts from mock-electroporated cells showed a high number of ^3H -mannose incorporated glycoproteins (lane 6). In cells electroporated with recombinant pSFV RNA, novel species of ~65 kDa and ~35 kDa were readily apparent and there was a significant reduction in the number of cellular

glycoproteins detected (Fig 3.11, lanes 4 and 5). Based on other studies these novel species correspond in size to E2 (~65 kDa) and E1 (~30 kDa; Grakoui *et al.*, 1993b, Ralston *et al.*, 1993). The diffuse nature of the ³H-mannose-labelled bands is characteristic of glycoproteins with multiple glycosylation sites. Also, in the extract from pSFV/E1_{his}E2-electroporated cells, the E1 species is slightly larger than that made from pSFV/E1E2_{Gla}; this is presumably a consequence of the presence of the histidine tag in E1. Radiolabelling with ³⁵S-methionine over the same time interval did not reveal any novel species in crude extracts which corresponded to E1 and E2 (Fig 3.11, compare lanes 1 and 2 with lane 3). Another labelled band of approximately 16 kDa (lane 5) was detected in the extract from pSFV/E1E2_{Gla}-electroporated cells. The origin of this band was unclear but was not further examined.

For any given expression system, it is important to establish when the system initiates protein production and the levels of expression achieved at different time periods after transfection or electroporation. Therefore, pSFV/E1_{his}E2- and mock-electroporated cells were metabolically labelled with ³⁵S-methionine/cysteine for 3 h intervals over a 25 h incubation period. The first labelling period was conducted 4 h post electroporation. Immediately after the labelling period, cells were washed and lysed in binding buffer (Section 2.2.17). An equal volume of the soluble material from each time period was analysed by SDS-PAGE (Fig 3.12). Data showed that the level of host cell protein synthesis in pSFV/E1_{his}E2-electroporated cells started to decline by 10-12 h post electroporation (lane 3) and this effect was enhanced at later time points (lanes 4-7). Indeed, host protein synthesis was considerably reduced by 19 h after electroporation with the recombinant SFV RNA (lanes 6 and 7). By contrast, mock-electroporated cells, which had been electroporated at the same time in the absence of RNA transcripts, showed no alteration in host protein synthesis (lanes 8-14).

Crude cell extracts obtained at 4, 7, 10, 13, 15 and 18 h after electroporation with pSFV/E1_{his}E2 RNA were also analysed for expression of E2 by Western blot analysis using the E2-specific MAb, ALP98 (Fig 3.13). This revealed that E2 expression was not apparent at 4 h after electroporation (lane 1). However, the protein was readily detected by 7 h after electroporation (lane 2) and peak levels were reached by 13 h (lane 4). There does not appear to be a significant difference between the levels of expression in cells harvested at 13 h compared with cells harvested at 15 or 18 h post electroporation (compare lanes 4-6), by which point host cell protein production is reduced (compare Fig 3.12, lanes 4 and

5). The only other notable difference was the rise in heterogeneity of E2 at later times following electroporation. It is assumed that E1 expression would follow the same pattern, since it is made co-translationally with E2 as the N-terminal part of the polyprotein.

3.7. Isolation of E1_{his} and E1E2 Complexes by IMAC

Since purification and detection of E1 was to be accomplished by IMAC, conditions for specific binding and elution of histidine-tagged E1 from Ni-NTA resin had to be established. In addition, E1 and E2 glycoproteins from a number of HCV strains have been reported to form a heteromeric complex. Therefore, the histidine-tagged purification system would also allow identification of any E1E2 complexes formed by strain Glasgow glycoproteins through co-elution of E2 with E1.

Initially, to address the question of Ni-NTA agarose specificity, cells were electroporated with RNA transcribed from constructs pSFV/E1_{his}E2 and pSFV/E1E2_{Gla}, in addition to mock-electroporated cells. Following an incubation period of 16 h, cells were lysed in 500µl binding buffer containing 5mM imidazole, an analogue of histidine with a high affinity for nickel ions (Fig 3.14). The soluble material was separated from the insoluble cell debris by centrifugation, and then applied to Ni-NTA agarose. Unbound material was collected and bound proteins were eluted with gel loading buffer. Subsequently, equal volumes of soluble material from each extract and the cell debris fraction, resuspended in 500µl gel loading buffer, were analysed alongside equal volumes of unbound material and total bound material eluted from Ni-NTA resin. Analysis was conducted by Western blotting for E1 using an E1-specific MAb, E1F. This data revealed that variable amounts of E1 were detected in the cell debris, soluble and the non-Ni-NTA bound supernatant fractions of samples derived from constructs pSFV/E1_{his}E2 (Fig 3.15, panel a, lanes 1-3, respectively) and pSFV/E1E2_{Gla} (lanes 5-7) but E1 was not found in mock-electroporated samples (lanes 9-11). Significantly, E1 was detected in the Ni-NTA bound and eluted material from the pSFV/E1_{his}E2 cell extract but not from the pSFV/E1E2_{Gla} cell extract (Fig 3.15, panel a, compare lanes 4 and 8). Thus, histidine-tagged E1 is able to associate specifically with Ni-NTA agarose while the untagged counterpart protein does not bind to the resin.

To establish whether assembly of a heteromeric complex occurred between strain Glasgow E1 and E2, co-elution of E2 with the histidine-tagged E1 was examined. Therefore, the same nitrocellulose membrane used for the analysis of E1 (Fig 3.15, panel a) was stripped of immunoglobulins and re-probed using E2-specific MAb E2G (panel b). The detection pattern for E2 was similar to that for E1, although substantially less E2 was identified in the insoluble cell debris samples, which may be a reflection of the efficient solubilisation of the protein. However, E2 was readily apparent in the soluble fractions of the cell extracts and the unbound material for pSFV/E1_{his}E2 and pSFV/E1E2_{Gla} samples (Fig 3.15, panel b, lanes 2, 3, 6, and 7). Interestingly, E2 was only detected in Ni-NTA-eluted material derived from crude extracts containing the histidine-tagged E1 (compare lanes 4 and 8). Since E2 is specifically eluted along with histidine-tagged E1, this indicates that the strain Glasgow glycoproteins are able to form a complex.

3.8. Optimisation of Ni-NTA Binding of E1_{his}

Having established that a complex between strain Glasgow glycoproteins could be detected via the nickel-NTA binding assay, optimal conditions for the isolation of histidine-tagged E1 by Ni-NTA agarose binding were determined. Mock-electroporated and pSFV/E1_{his}E2 RNA-electroporated cells were radiolabelled from 4-12 h post-electroporation and lysed in binding buffer containing various concentrations of imidazole. Ni-NTA binding of the soluble fraction from cell extracts was performed in buffers (Section 2.2.17) with imidazole at final concentrations of 5mM, 10mM, 20mM, 30mM and 40mM. Binding in the absence of imidazole also was conducted for control purposes. Samples were eluted with gel loading buffer and analysed by SDS-PAGE (Fig 3.16). Two major species were isolated by Ni-NTA purification from extracts of pSFV/E1_{his}E2-electroporated cells that correlated in size with E1 and E2; this again confirmed formation of E1E2 complexes (lanes 7-12). In addition, use of various concentrations of imidazole during the binding step showed that in the presence of 40mM imidazole, interaction of histidine-tagged E1 with Ni-NTA was not abolished and the level of non-specific binding was reduced to a minimum (lane 12). This is observed also in samples eluted from mock-electroporated crude extract at the same imidazole concentration (lane 6), where a significant reduction in the level of non-specific protein binding was evident. In the absence of imidazole in the binding buffer, very little tagged E1 was eluted (lane 7). In all subsequent experiments, binding buffer contained 40mM imidazole.

To examine whether the co-elution of proteins non-specifically bound to Ni-NTA agarose could be further reduced, removal of bound E1 was attempted with buffers containing various concentrations of imidazole. Experimental conditions were as stated for the experiment above, except that a 30 μ l aliquot of binding buffer containing imidazole at final concentrations of 200mM, 400mM, 600mM, and 800mM was used as a first elution step. Residual proteins remaining on the resin were removed with the same volume of gel loading buffer. The eluted samples were again analysed by SDS-PAGE (Fig 3.17.). The data showed that although E1 and E2 could be removed from the column with each concentration of imidazole, no significant difference in the level of non-specific protein elution was achieved with increases in imidazole concentration in the elution buffer (lanes 3, 5, 7 and 9). This is particularly obvious when compared to the material eluted directly with gel loading buffer (lane 2). The remaining lanes on the gel represent the proteins strongly bound to Ni-NTA resin, which were only eluted using gel loading buffer (lanes 4, 6, 8 and 10). In light of this data where no substantial reduction in non-specific protein is detected using imidazole as an eluant, proteins bound to Ni-NTA were eluted with gel loading buffer in subsequent experiments.

3.9. Immunolocalisation and ER Retention of E1_{his} and E2 in BHK Cells

Further comparison of the behaviour of strain Glasgow E1 and E2 with that reported for glycoproteins made by other strains of HCV was performed by indirect immunofluorescence. Cells, electroporated with pSFV/E1_{his}E2 RNA, were fixed with methanol and probed for E1 and E2 using antibodies (R528 and ALP98 respectively) along with anti-calnexin-specific antiserum as a marker for the ER. Localisation of the glycoproteins and calnexin was examined by confocal microscopy (Fig 3.18). Both E1 and E2 had a cytoplasmic, reticular distribution (panels a and b) and co-localised on merging the images (panel c). Analysis of the localisation of calnexin revealed that it was widely distributed throughout the cell (panel d). Within the same cell, glycoprotein E2 had a more restricted distribution (panel e). Merging images in panels d and e (panel f) shows that E2 co-localises with calnexin but emphasises the wider distribution of the ER chaperone. This suggests sub-compartmentalisation of the HCV glycoproteins in the ER.

Translocated proteins that are processed in the ER are normally destined for secretion or the plasma membrane by transport through the secretory pathway. However, other proteins, termed ER-resident proteins, are either retrieved from the *cis*-Golgi sub-compartment or are specifically retained (Section 1.4.8). The N-linked high-mannose oligosaccharides transferred onto proteins in the ER are modified on those proteins which are retrieved from the Golgi apparatus. By contrast proteins that remain in the ER are not modified (Kornfeld & Kornfeld, 1985). The oligosaccharide modifications render the retrieved proteins insensitive to endo H_f, which removes N-linked high-mannose oligosaccharides. This enzyme was employed to identify whether strain Glasgow E1 and E2 were retrieved or specifically retained in the ER.

Extracts from cells electroporated with pSFV/E1_{his}E2 RNA were treated with endo H_f, and from Western blot analysis (Fig 3.19), both E1 (panel a) and E2 (panel b) were sensitive to the enzyme (lanes 2). Compared to untreated proteins (lanes 1), endo H_f-treated proteins migrated faster on polyacrylamide gels. This analysis suggests that E1 and E2 are retained in the ER and are not retrieved from the *cis* Golgi network or transported out of the ER.

Accumulation of proteins retained in the ER can lead to their degradation by the ERAD pathway (Section 1.4.9). To examine the stability of E1 and E2 in the ER over time, cells were electroporated with pSFV/E1_{his}E2 RNA. 7 h after electroporation cells were placed in media containing cycloheximide (CHX) to inhibit any further protein synthesis. Subsequently, cells were harvested at 3 h intervals following the addition of CHX. The crude extracts derived from these cells were examined by Western blot analysis using E2-specific MAb ALP98 (Fig 3.20). The result indicated that the level of E2 detected at 7 h after electroporation (lane 1) remained constant after the addition of CHX to the media (compare lane 1 with lanes 2-5). This suggested that the E2 glycoproteins retained in the ER remain relatively stable over a period of time.

3.10. Discussion

In recent years, SFV has been explored as a system for the expression of heterologous proteins by replacing the SFV structural coding sequences with those of a gene of interest (Liljestrom & Garoff, 1991b). As with almost all expression systems, there are both advantages and disadvantages associated with the use of SFV. One of the principle reasons for selecting SFV to express the HCV glycoproteins was the ease with which

mutants could be constructed in inserted genes since the virus vector is plasmid-based. This facet of the vector was put to use in the construction of a substantial number of variants of E1 and E2 (Chapters 5 and 6). Expression is dependent on the introduction and replication in cells of RNA transcribed *in vitro* from a SFV plasmid template. Moreover, electroporation rather than transfection was employed as the preferred method for introducing RNA into cells. The efficiency of transfection of nucleic acid into cells is often variable, however, the *in vitro* transcription and electroporation methods used in this study gave consistent results from experiment to experiment. It is also possible to manipulate conditions by varying the amount of RNA and electroporation settings to modulate the number of cells that express genes inserted into the vector and this adds flexibility to the system. The levels of expression of the HCV glycoproteins were readily detectable by both Western blot analysis and indirect immunofluorescence. However, immunoprecipitation with specific antibodies (Chapters 4-6) or purification of histidine-tagged protein by IMAC was necessary to identify the proteins in radiolabelled extracts. Another advantage of the SFV vector over other viral expression systems such as vaccinia virus is that no progeny virus is produced. This may limit the alterations to cellular pathways which can occur by the expression of a large number of genes by the virus vector. For example, the vaccinia virus genome encodes about 200 potential gene products (Moss, 1996) whereas SFV synthesises only 9 proteins. However, SFV infection does induce apoptosis and the gene involved in induction is the non-structural gene nsP2 (Glasgow *et al.*, 1998). This gene is made by the SFV vector and hence apoptosis is a feature of the expression system. Apoptosis markers such as laddering of extracted DNA, characteristic of internucleosomal DNA fragmentation, and terminal deoxynucleotidyl transferase-mediated bio-UTP nick end labelling (TUNEL; Gavrieli *et al.*, 1992) have been examined in pSFV1-electroporated cells. Positive results for both of these markers were observed at 48 h after electroporation but were not detected at 24 h after electroporation (Glasgow *et al.*, 1998). Such a cytopathic effect is a disadvantage since expression in SFV-electroporated cells can continue for only a limited period. Moreover, such effects could influence the behaviour of the inserted genes that are expressed. In this study, crude cell extracts were typically prepared 12-18 h post-electroporation with recombinant SFV plasmids and therefore before the reported time at which apoptotic markers are evident. E1 and E2 proteins also were glycosylated and of the correct predicted sizes. Thus, cleavage and processing at the ER were not inhibited. Moreover, the two proteins formed a complex in agreement with reports using alternative expression systems. Recent results

using SFV to express structural proteins of a flavivirus, louping ill virus, have shown that they also are correctly processed and glycosylated (Fleaton *et al.*, 1999).

It was possible that the SFV system could have produced virus particles containing HCV structural proteins. Expression of a subgenomic portion of Moloney murine leukemia virus (MMLV) along with a region containing MMLV replication and packaging signals by the SFV vector resulted in the production of extracellular virus-like particles (Li & Garoff, 1996). These particles possessed reverse transcriptase activity and contained only MMLV structural proteins which were authentically processed. In another study, virus-like particles were detected which consisted of naked replicon RNA and spikes of VSV G protein in the absence of any other VSV proteins (Rolls *et al.*, 1994). These particles could be propagated in tissue culture. Analysis of culture media from cells expressing HCV core, E1 and E2 by SFV did not indicate that propagation-competent virus particles were produced. Moreover, no structures that could correspond to HCV virus-like particles were found by electron microscopy analysis of electroporated cells (J. McLauchlan, J. Aitken and S. Butcher, personal communication). In early experiments, co-electroporation of RNAs from the SFV helper virus vector (Section 1.6) and the recombinant SFV plasmid encoding E1 and E2 did result in the production of extracellular SFV particles. However, analysis of these particles revealed no detectable amounts of incorporated HCV glycoproteins.

In the absence of any substantial amounts of E1- or E2-specific antibodies, a histidine tag was introduced at the N-terminus of E1 for purification purposes. Although it is not possible to predict the consequence of incorporation of short amino acid tags into a polypeptide coding region, native proteins which retain their function can be isolated from cells by Ni-NTA purification. For example, vaccinia virus has been used to produce histidine-tagged human serum response factor from HeLa cells (Janknecht *et al.*, 1991). The purified protein retained the capacity to form a complex with a second cellular component, and stimulate transcription *in vitro* from a template bearing a promoter with the serum response element. Purification conditions were determined for strain Glasgow histidine-tagged E1 that did not require denaturation of the proteins for binding and elution from Ni-NTA agarose. As a consequence, E2 was shown to co-elute with tagged E1 and thus E1 and E2 from strain Glasgow are able to form a complex. These data also indicate that the presence of the histidine tag did not abolish complex formation. In a separate study using the baculovirus expression system, truncated forms of E1 and E2 which were

both histidine-tagged at their N-termini were able to form a complex (Hussy *et al.*, 1996). Moreover, truncated forms of histidine-tagged E1 and E2 expressed in bacteria were capable of binding truncated E2 and E1 respectively (Yi *et al.*, 1997a). However, binding in this latter study was performed entirely *in vitro* and therefore did not take account of any active role played by ER proteins in complex formation and protein folding.

Endo H_f analysis revealed that glycosylated E1 and E2 remain sensitive to the enzyme, indicating that the proteins are retained in the ER. The lack of any modifications to glycans is consistent with retention of the proteins in the ER as opposed to a cyclical mechanism of transport to the Golgi apparatus followed by retrieval (see Section 1.4.8). This conclusion agrees with other studies using this enzyme (Dubuisson *et al.*, 1994; Cocquerel *et al.*, 1998; Duvet *et al.*, 1998). In addition, Duvet *et al.* (1998) made use of endo D to confirm that cycling did not occur from the ER to the *cis* Golgi. If cycling between these compartments did occur, then the glycans would be modified by α -mannosidase I in the *cis* Golgi and molecules with these modifications would accumulate. Such modifications are sensitive to endo D however E1 and E2 remained endo D-resistant. Further HPLC analysis of glycans released by PNGase F treatment also indicated retention as opposed to retrieval as the mechanism for localisation of E1 and E2 in the ER (Duvet *et al.*, 1998).

From confocal microscopy analysis, the distribution of E1 overlapped exactly with that of E2, confirming that the localisation of strain Glasgow glycoproteins is similar to that detected for those of strain H (Dubuisson *et al.*, 1994). The distribution of E2 was coincident with that of calnexin, a chaperone resident in the ER. Moreover, E2 was found predominantly in the regions of the ER nearer the cell nucleus. However, calnexin had additional ER distribution where E2 was not present which was towards the periphery of the cell. Examination of E2 co-localisation with another ER marker, PDI, revealed identical results; PDI had broad ER distribution which partially overlapped with that of E2, which was localised closer to the nucleus (Duvet *et al.*, 1998). These data suggest E1 and E2 glycoproteins may be retained in subcompartments within the ER. This organelle is composed of both rough and the smooth compartments, regions which perform different cellular functions and this is reflected in the proteins they contain (Lee & Chen, 1988). While the overall protein and lipid composition of the SER and RER are similar, the translocon-associated protein complex is mainly restricted to the rough ER (Vogel *et al.*, 1990). Using cell fractionation, p63, a novel type II integral membrane protein co-

fractionates with the rough ER marker ribophorin II (Schweizer *et al.*, 1995). From immunofluorescence data, the distribution of p63 was broader than a reporter protein (cathepsin D) fused to an ER retrieval signal (KDEL) which localised more closely to the outer nuclear membrane. Although E1 and E2 are not considered to be retrieved from the Golgi to the ER, this subcompartmentalisation within the ER may suggest that E1 and E2 glycoproteins may move away from their site of synthesis.

Incompletely folded and misfolded proteins are retained in the ER by resident chaperones and foldases. Prolonged retention and accumulation of proteins induces the ER stress response which leads to the production of stress response proteins and the activation of ER-associated degradation through the ubiquitin-proteasome pathway (Section 1.4.9; Shamu *et al.*, 1994). E1 and E2 retained in the ER were relatively stable over a period of 11 h in the presence of the protein synthesis inhibitor CHX. This procedure may also have prevented the synthesis of cellular proteins required for the proper function of the ER-stress response. However, studies conducted for the degradation of p35, a neurone-specific activator of cyclin-dependent kinase-5 (CDK5), in cultured neurons were performed also in the presence CHX (Saito *et al.*, 1998). The amount of p35 in the cytoplasm of neuronal cells is balanced through synthesis and degradation pathways. In the presence of CHX, the disappearance of p35 was accelerated and was accompanied by parallel inactivation of CDK5. Although this is an example of a cytoplasmic protein, it does demonstrate that the presence of CHX does not cause disruption of degradation pathways. Thus, E1 and E2 may be relatively stable in the ER.

Studies conducted on other proteins [e.g. human asialoglycoprotein receptor (HAR) and soluble protein cathepsin D] have revealed that proteins retained in the ER require certain determinants to cause their degradation which are separable from retention signals (Alberini *et al.*, 1990; Shenkman *et al.*, 1997). HAR subunit H2a contains, next to its transmembrane segment, a charged EGHRG (Glu-Gly-His-Arg-Gly) pentapeptide in the ectodomain, which is absent from the H2b subunit, an alternatively spliced variant. A portion of H2a is cleaved to produce a soluble secreted ectodomain fragment, while H2b remains membrane bound and is transported from the ER efficiently. The portion of H2a that remains membrane bound is retained in the ER and degraded. The transfer of the EGHRG sequence to another subunit, H1, which is normally transported, resulted in its retention in the ER, but it was not degraded (Shenkman *et al.*, 1997). This suggests that the determinant for ER retention (the pentapeptide) in uncleaved H2a does not overlap

with the determinant for its degradation. Therefore, E1 and E2 also may lack specific signals to initiate their degradation also.

Data presented in this chapter has demonstrated that a heteromeric complex forms between E1 and E2 from strain Glasgow, in accord with the behaviour of glycoproteins made by other HCV strains. The nature of these complexes was further analysed and the experiments conducted are presented in the following chapter.

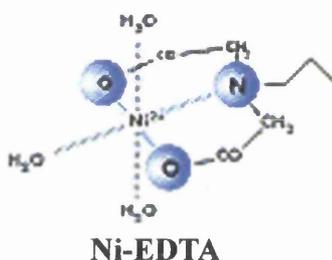
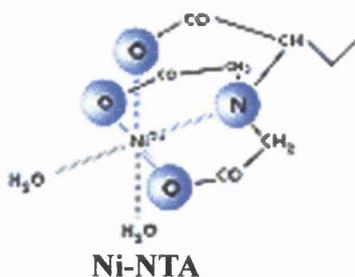
No. of pulses applied	No. of cells counted	No of trypan blue stained cells	Percentage of dead cells (%)
1	306	27	8
2	325	60	18

Table 3.1. The approximate level of cell death resulting after electroporation.

Amount of RNA used per electroporation (μg)	Total number of cells counted	No. of fluorescing cells	Percentage electroporation efficiency (%)	No. of RNA molecules available per cell ($\times 10^5$)
7.5	685	312	45	0.75
15	502	247	49	1.5
30	542	447	83	3
60	462	397	85	6

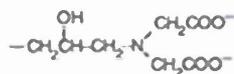
Table 3.2. The approximate level of electroporation efficiency achieved with increasing amounts of *in vitro* transcribed RNA.

a.

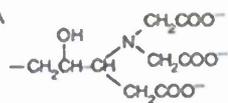


b.

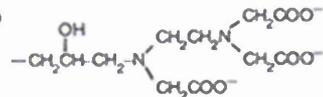
EDTA



NTA



TED



increasing
chelation
with Ni^{2+}

Fig 3.1. a. Comparison of the interaction of different metal chelation matrices with nickel ion (Ni^{2+}). b. Metal chelating agents typically used in immobilized metal affinity chromatography (IMAC). Compound molecules are shown in increasing order of metal chelating strength. Abbreviations are as follows: NTA, nitrilo triacetic acid; TED, tris (carboxymethyl) ethylene diamine. (Taken from Qiagen handbook, 1997)

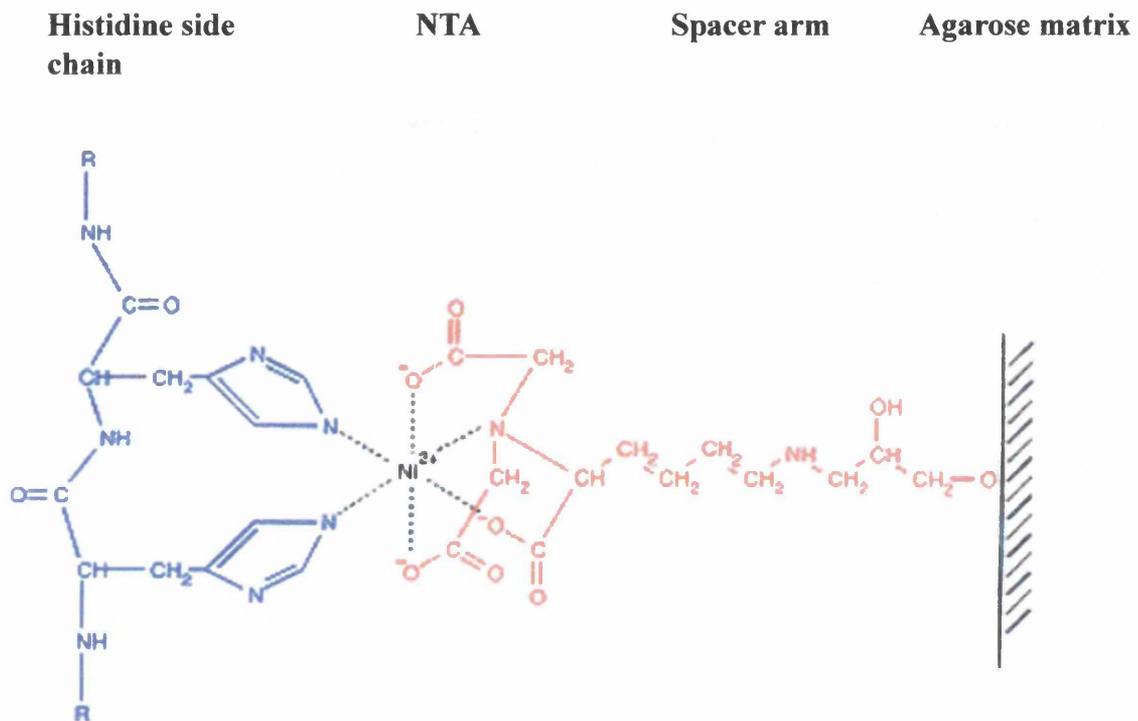
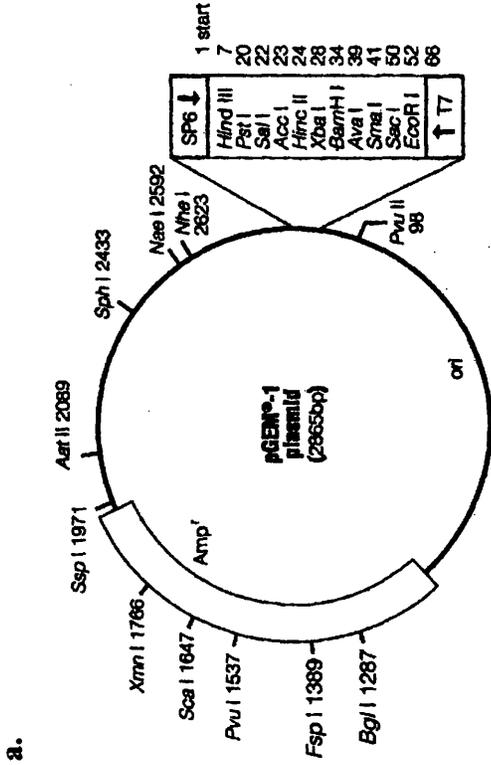


Fig 3.2. Structure of histidine-Ni-NTA complex. The nickel chelating NTA is immobilised to agarose via a spacer arm as shown. The remaining two ligand sites on the hexavalent nickel ion are occupied by the imidazole group on side chains of histidine residues. (Taken from Qiagen handbook, 1997)



c.

HCV structural protein polypeptide (amino acid 1-838)

↓

EcoRI - *BglIII* - *ApaLI* - *Eco47III* - * * * - *BglIII* - *HindIII*

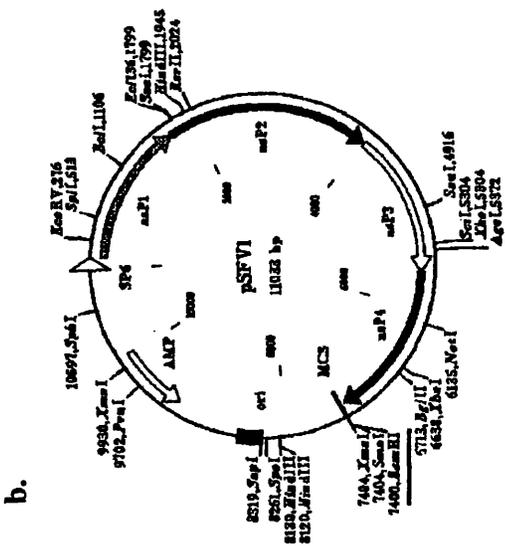


Fig 3.3. Schematic showing pGEM1 and pSFV1 plasmids. The multiple cloning region of pGEM1 plasmid (panel a) was modified to incorporate *ApaLI* and *Eco47III* restriction sites at which the HCV structural protein coding region of strain H77 (panel c) was inserted (Section 4.6). Flanking *BglIII* sites permitted the cloning of the coding regions into pSFV1 (panel b) at a unique *BamHI* site (underlined in red). Stop codons engineered downstream of the HCV ORF are indicated by ***. The ampicillin resistance gene (AMP) and SFV non-structural proteins 1-4 (nsP1-4) are shown. Figures for pGEM1 and pSFV1 are taken from the Promega catalogue and Liljestrom & Garoff, (1991b) respectively.

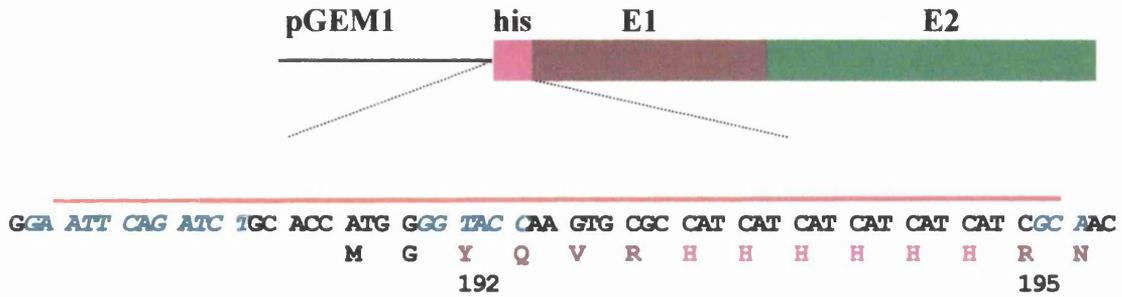


Fig 3.4. Insertion of a histidine tag at the N-terminus of E1. Schematic shows the insertion of an oligonucleotide (over-lined in red) encoding the histidine (his) tag at the *FspI* site in E1 (corresponding to codon 195) and at the *EcoRI* site in pGEM1 plasmid. This construct was named pGEM/ Δ E1_{his}E2, because it lacked sequences coding for the signal peptide for ER translocation. Nucleotides in italics represent restriction enzyme sites, including *EcoRI*, *BglIII*, *KpnI* and *FspI* (not regenerated by the oligonucleotide). E1 amino acid sequences are shown in brown, histidine tag residues are in pink and two novel amino acid residues in black include an initiating methionine residue. The schematic shows only the relevant regions of HCV sequences at which alterations were introduced.

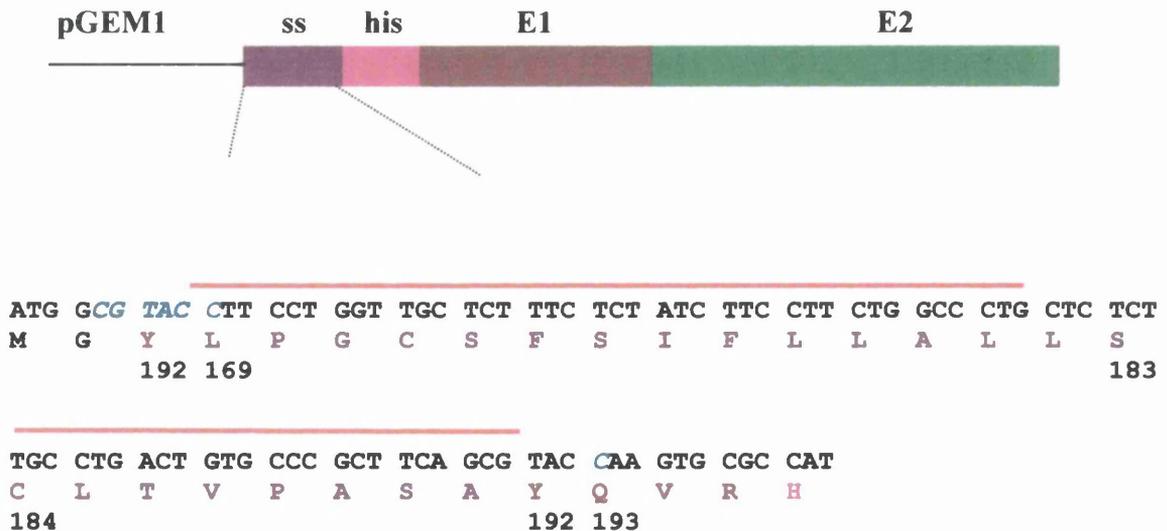


Fig 3.5 Insertion of E1 signal sequence at the N-terminus of the histidine tag in E1. The E1 signal sequence (amino acid residues 169-191 in the polyprotein) was introduced 5' of the histidine tag coding region in pGEM/ Δ E1_{his}E2 at the *KpnI* site (italicised) by introduction of an oligonucleotide (over-lined in red). This construct was called pGEM/E1_{his}E2.

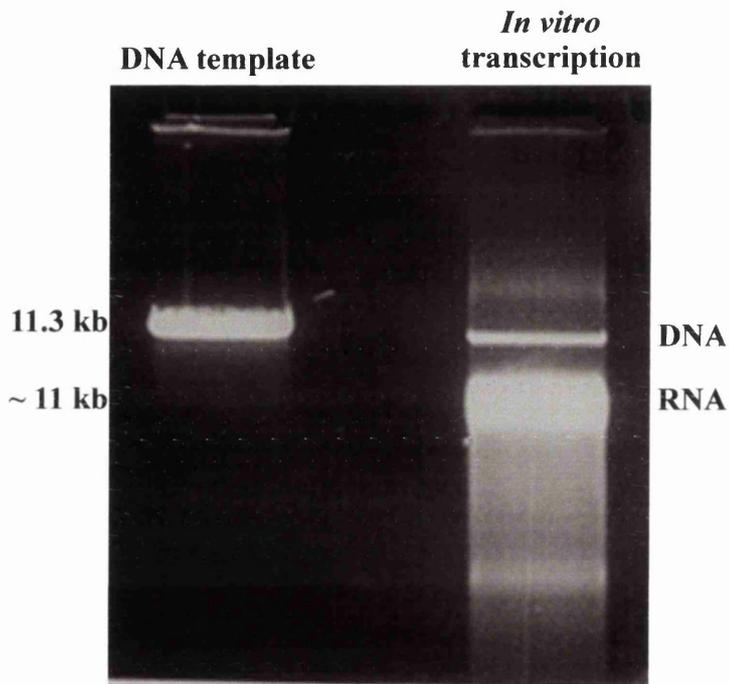


Fig 3.6. *In vitro* transcription of RNA from SFV templates. 2 μ l of the completed *in vitro* transcription reaction analysed on a 1% agarose gel alongside 1 μ g of linearised template DNA (pSFV/E1_{his}E2) is shown. The smear under the major RNA band is indicative of RNA degradation products.

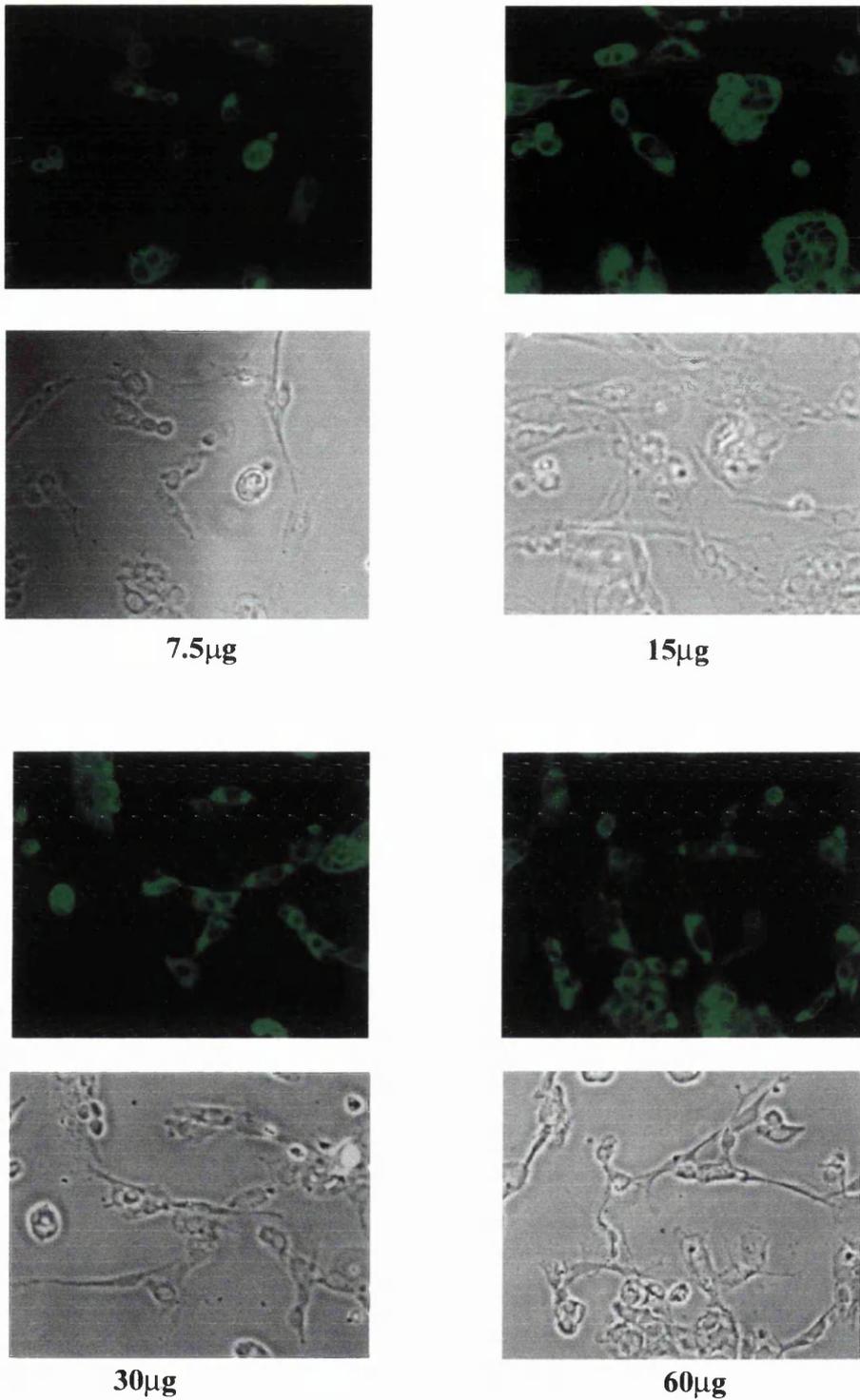


Fig 3.7. Phase contrast and indirect immunofluorescence images of electroporated cells. BHK cells were electroporated with 7.5, 15, 30 and 60 μ g of *in vitro* transcribed RNA from pSFV/E1_{his}E2. Cells were incubated for 18 h and fixed in methanol. They were processed for indirect immunofluorescence using E2-specific MAb, ALP98. Images were captured at x20 magnification.

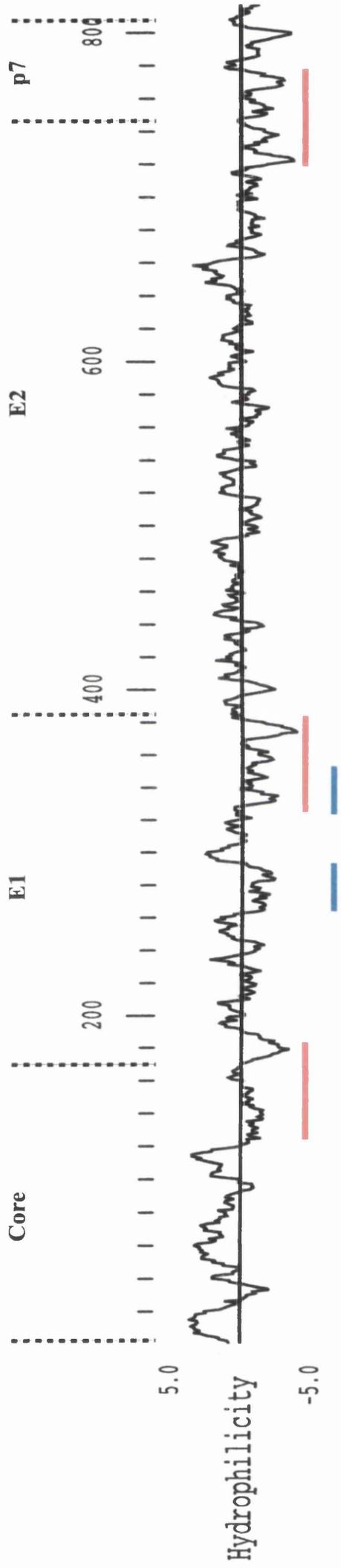


Fig 3.8. Hydrophobicity profile of the N-terminal region of the HCV polyprotein corresponding to the structural proteins. The profile corresponds to amino acids 1-838 which is processed into mature core, E1, E2 and p7 proteins. The potential signal sequence/transmembrane domains are underlined in red. Regions underlined in blue are discussed in Chapter 5, Section 5.5.

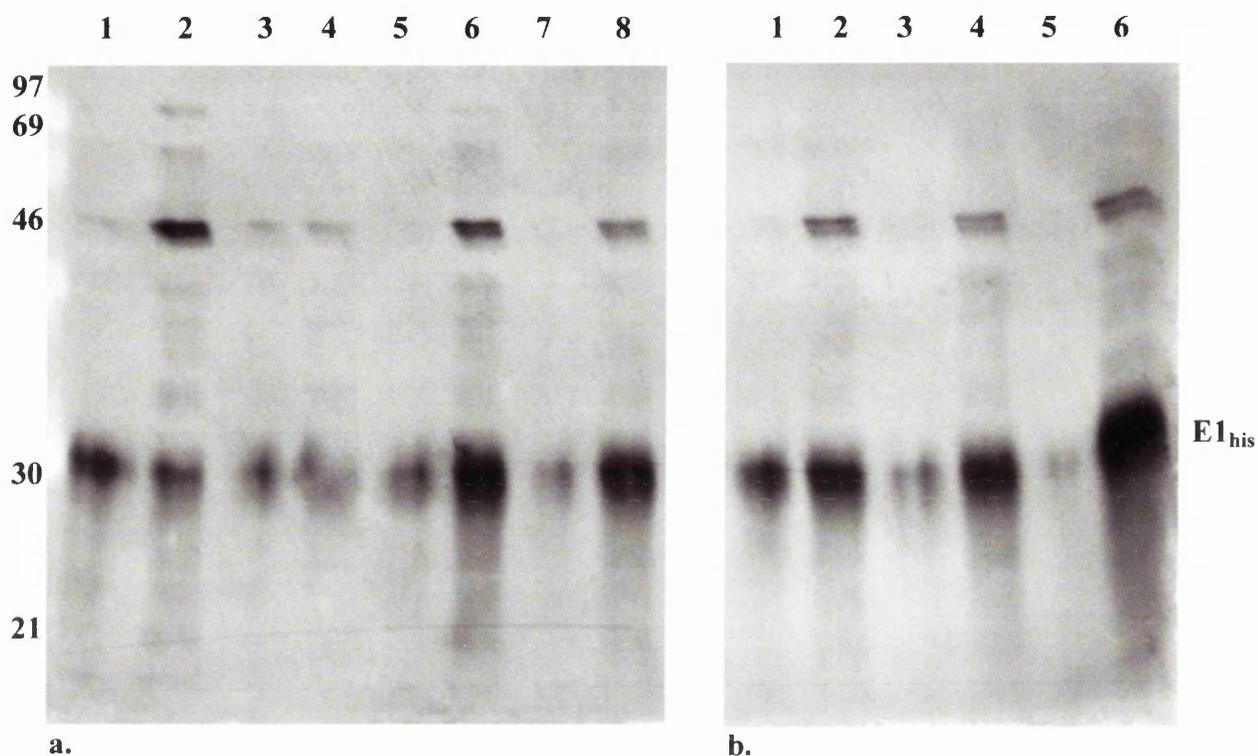
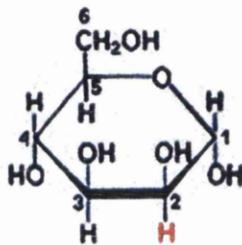
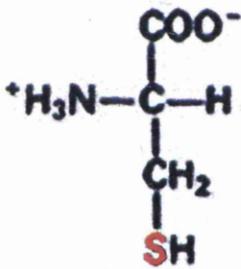


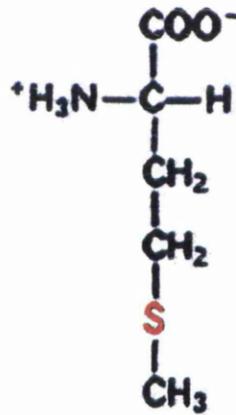
Fig 3.9. Solubilisation of E1 protein from BHK cells. **Panel a.** pSFV/E1_{his}E2 RNA-electroporated cells were incubated for 18 h and harvested in Tris-salt buffer containing either deoxycholic acid (lanes 3 and 4), NP-40 (lanes 5 and 6) or Triton X-100 (lanes 7 and 8) detergents at a concentration of 0.1%, or in buffer containing no detergent (lanes 1 and 2). The cell debris was separated from the soluble material by centrifugation and equal volumes of both fractions were analysed by Western blot using the Penta-His MAb probing for the histidine tag on E1. Lanes 1, 3, 5, and 7 show the resuspended cell debris and lanes 2, 4, 6, and 8 contain the soluble material. **Panel b.** Cells were harvested in Tris-salt buffer containing Triton X-100 at a concentration of either 0.01% (lanes 1 and 2), 0.1% (lanes 3 and 4) or 1% (lanes 5 and 6). Again the amount of E1_{his} in the cell debris (lanes 1, 3 and 5) were compared with the soluble material (lanes 2, 4 and 6). The position of E1_{his} is indicated.



a.



b.



c.

Fig 3.10. Molecular structures of radiolabelled amino acid and carbohydrate residues. Panels show: a, 2-³H-mannose; b, ³⁵S-cysteine; c, ³⁵S-methionine. The radiolabelled elements are shown in red.

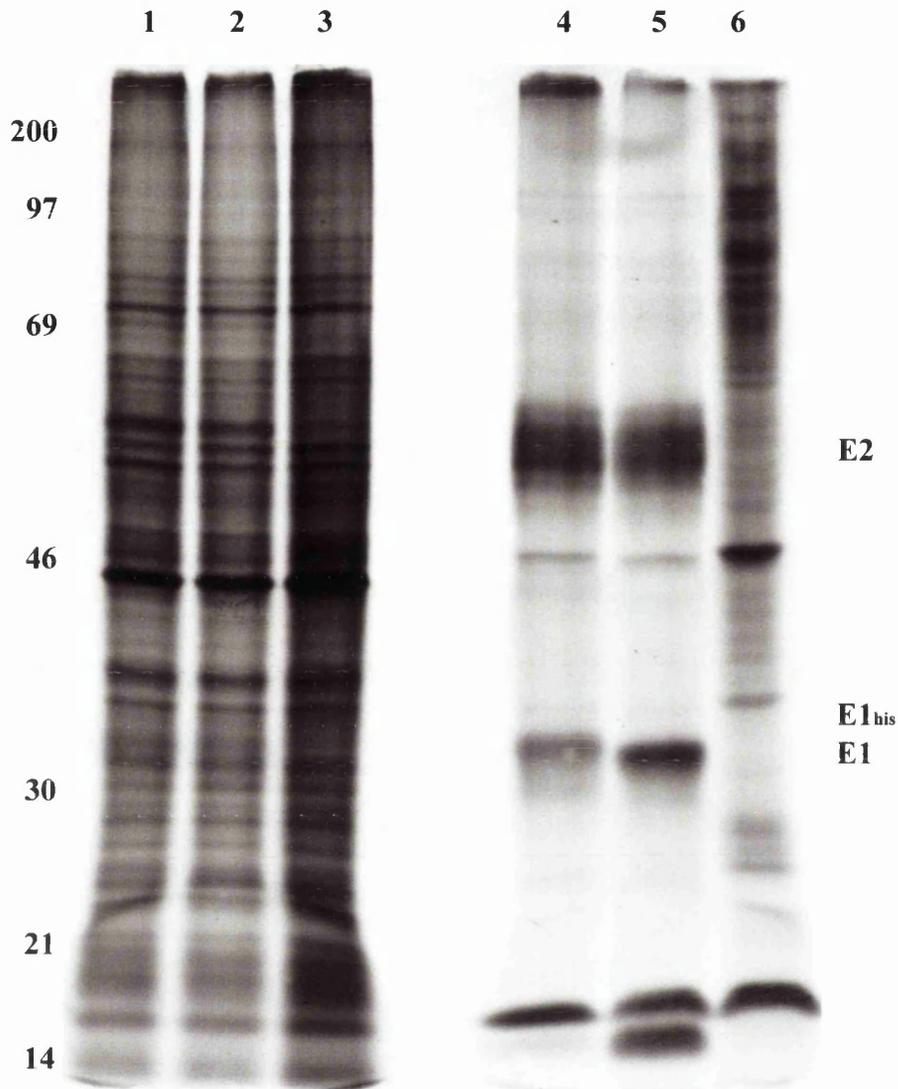


Fig 3.11. Metabolic labelling of cells with ^{35}S -methionine and ^3H -mannose. Cells electroporated with RNA from pSFV/E1_{his}E2 (lanes 1 and 4), pSFV/E1E2_{Gla} (lanes 2 and 5), and mock-electroporated cells (lanes 3 and 6) were radiolabelled 4 h after electroporation with either ^{35}S -methionine (lanes 1-3) or ^3H -mannose (lanes 4-6), followed by a 12 h incubation at 37°C overnight. Cell extracts were TCA precipitated and equivalent amounts of precipitated counts were analysed on 10% polyacrylamide gels. Bands presumed to correspond to E1_{his}, E1 and E2 are indicated.

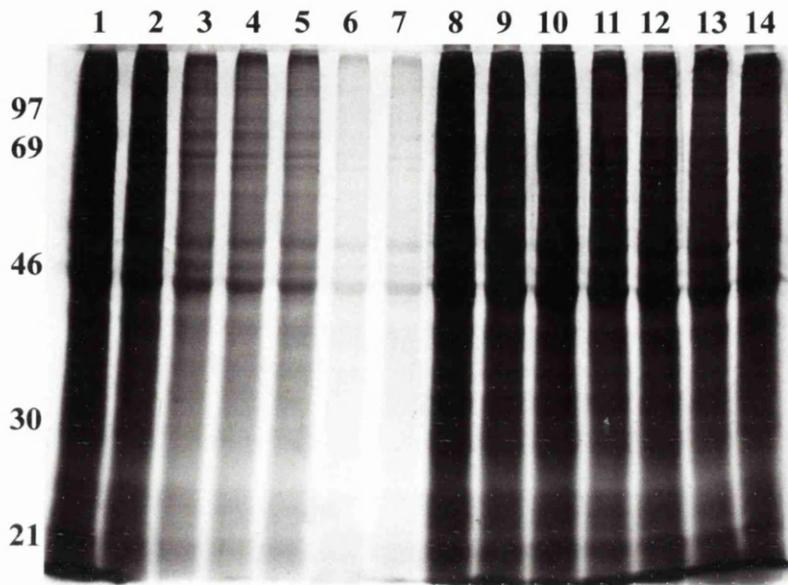


Fig 3.12. Efficiency of expression from recombinant pSFV RNA on protein synthesis. pSFV/E1_{his}E2 RNA-electroporated (lanes 1-7) and mock-electroporated cells (lanes 8-14) were seeded onto individual plates, which were ³⁵S-methionine/cysteine radiolabelled for 3 h over a time course of 25 h. The first set of plates were labeled for 4-6 h after electroporation (lanes 1 and 8), followed by 7-9 h (lanes 2 and 9); 10-12 h (lanes 3 and 10); 13-15 h (lanes 4 and 11); 16-18 h (lanes 5 and 12); 19-21 h (lanes 6 and 13); and 22-24 h (lanes 7 and 14). Immediately after the labeling period, cells were harvested and equal aliquots of crude cell extracts, obtained from each time period, were analysed on a 10% polyacrylamide gel.

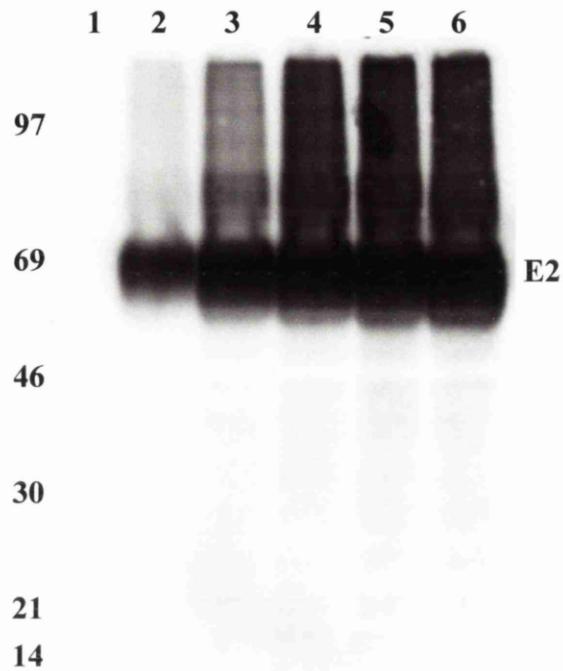


Fig 3.13. Time course of HCV E2 expression. Samples from cells electroporated with pSFV/E1_{his}E2 RNA were harvested over an 18 h time period at 3 h intervals. Samples were electrophoresed on a 10% polyacrylamide gel, transferred to nitrocellulose membrane and probed with the E2-specific MAb, ALP98. Cells were harvested after electroporation at 4 h (lane 1), 7 h (lane 2), 10 h (lane 3), 13 h (lane 4), 15 h (lane 5), and 18 h (lane 6). The position of E2 is indicated.

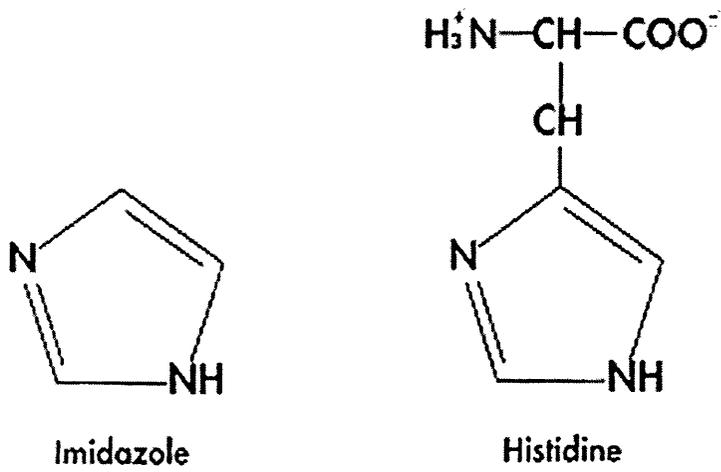


Fig 3.14. Schematic showing the structures of the side chain in a histidine residue compared with an imidazole molecule. (Taken from Qiagen handbook, 1997)

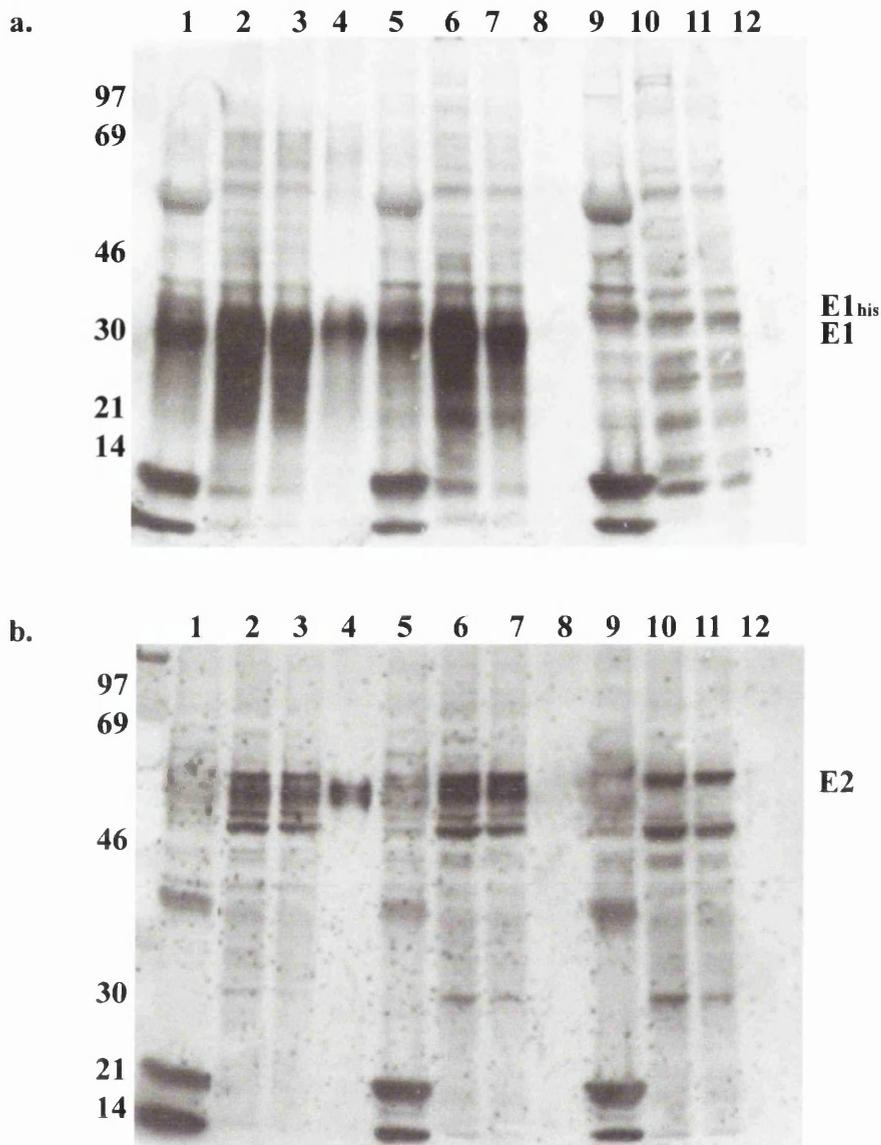


Fig 3.15. Analysis of binding of tagged E1 and E2 to Ni-NTA agarose. The cell debris (lanes 1, 5 and 9) and soluble fractions (lanes 2, 6 and 10) were prepared from cells electroporated with RNA from pSFV/E1^{his}E2 (lanes 1-4) or pSFV/E1E2^{Gla} (lanes 5-8) and mock-electroporated cells (lanes 9-12). Samples were analysed alongside proteins bound to Ni-NTA resin (lanes 4, 8 and 12) and aliquots of the unbound supernatant (lanes 3, 7 and 11). Following electrophoresis on a 12% polyacrylamide gel, proteins were transferred to nitrocellulose membrane and analysed by Western blot analysis using an E1-specific MAb E1F (diluted 1:500; panel a). Subsequently, the membrane was stripped of immunoglobulins and analysed again using E2-specific MAb E2G (diluted 1:500; panel b). The positions of E1^{his}, E1 and E2 are indicated. The characteristics of E1F and E2G are shown in Table 4.2.

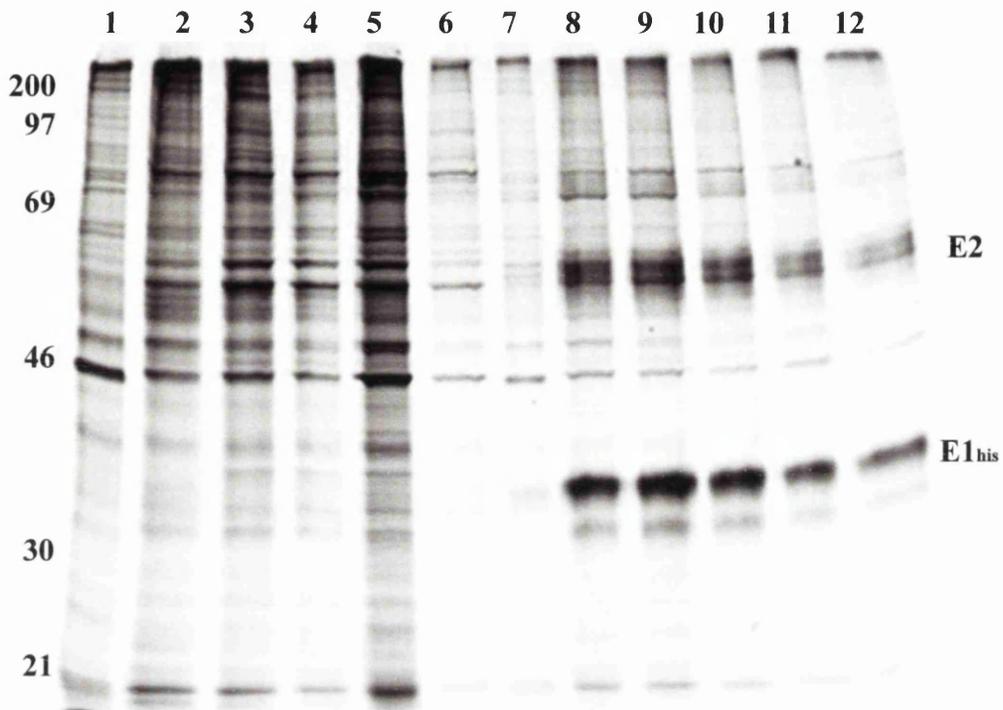


Fig 3.16. Effect of imidazole concentration on the binding of histidine-tagged E1 to Ni-NTA agarose. Cells, electroporated with RNA transcribed from pSFV/E1_{his}E2 (lanes 7-12) or mock-electroporated (lanes 1-6), were radiolabelled with ³⁵S-methionine for 4-12 h after electroporation. Crude extracts obtained from these cells were applied to Ni-NTA resin either in the absence (lanes 1 and 7) or presence of 5mM (lane 2 and 8), 10mM (lanes 3 and 9), 20mM (lanes 4 and 10), 30mM (lanes 5 and 11) and 40mM (lanes 6 and 12) imidazole. Bound proteins were eluted in loading buffer and analysed on a 10% polyacrylamide gel. Positions of E1_{his} and E2 are shown.

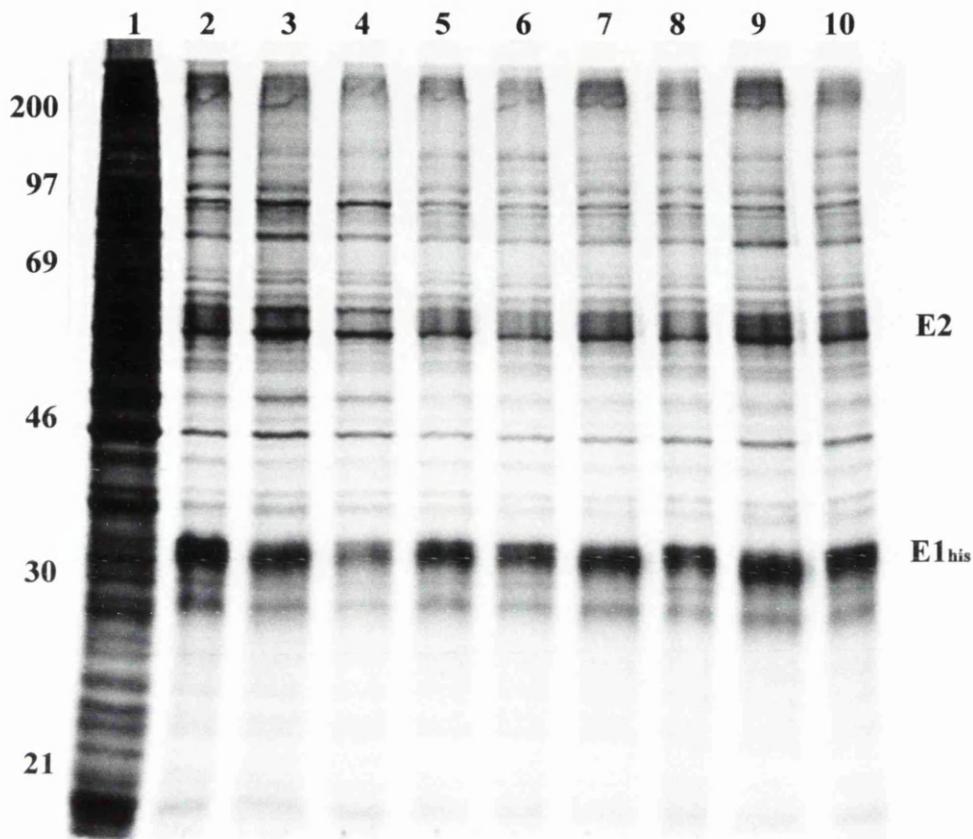


Fig 3.17. Effect of imidazole concentration on elution of tagged E1 from nickel-NTA agarose. Cells were electroporated with RNA from pSFV/E1_{his}E2 and radiolabelled for 4-12 h after electroporation. Crude extracts obtained from these cells were applied to Ni-NTA agarose and eluted in the presence of binding buffer containing imidazole at concentrations of 200mM (lane 3), 400mM (lane 5), 600mM (lane 7) or 800mM (lane 9). Gel loading buffer was then applied to the remaining resin to remove residual proteins (lanes 4, 6, 8 and 10). Lane 1 shows the crude extract applied to Ni-NTA agarose and lane 2 represents the bound material eluted directly in loading buffer. The positions of E1_{his} and E2 are indicated.

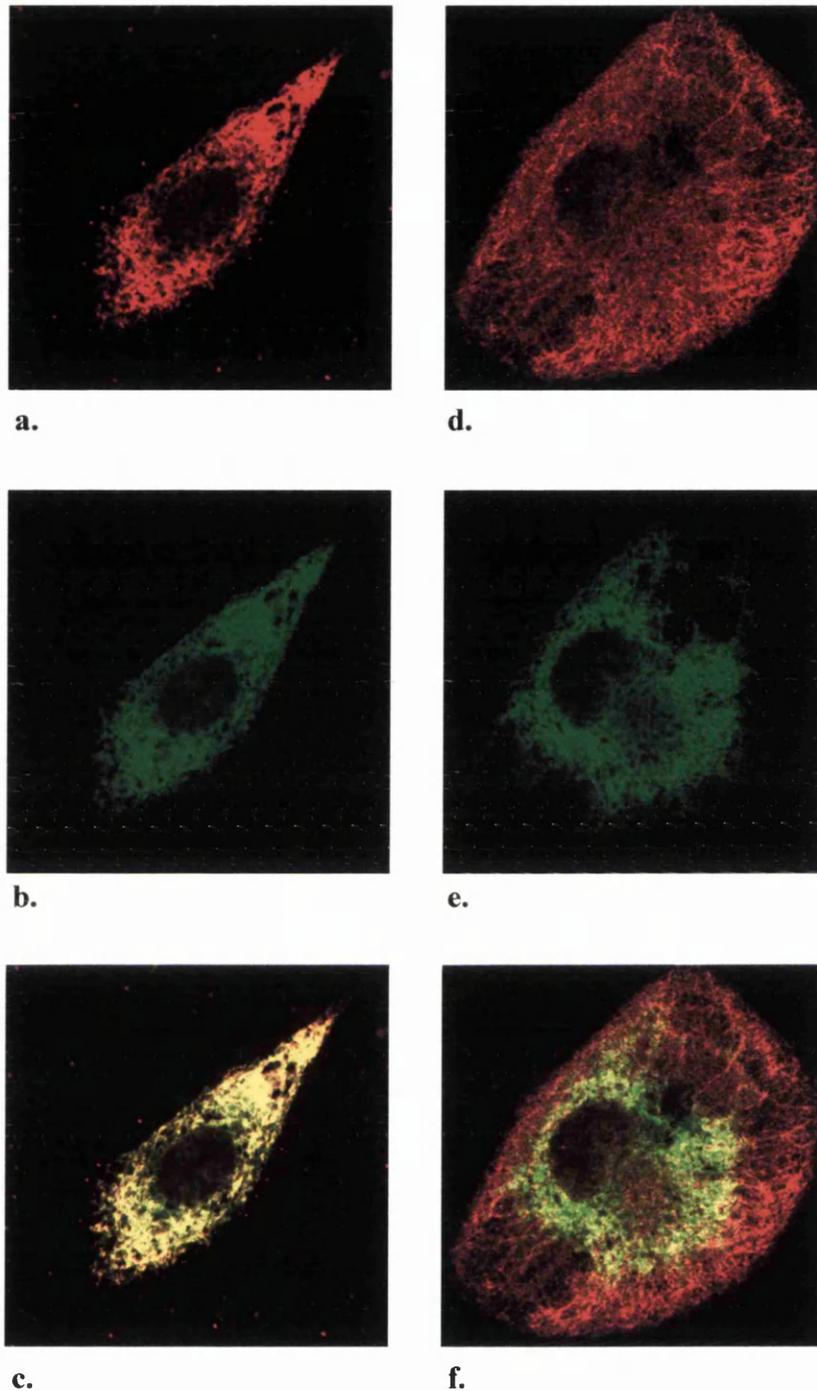


Fig 3.18. Confocal microscopy analysis of the localisation of E1, E2 and calnexin. Cells electroporated with pSFV/E1_{his}E2 RNA were fixed with methanol 16 h after electroporation and processed for indirect immunofluorescence. Images of these cells, captured by confocal microscopy, show intracellular localisation of E1 using anti-E1 antiserum, R528 (panel a), E2 using anti-E2 MAb, ALP98 (panels b and e), and calnexin using anti-calnexin antiserum (panel d). Merged images are shown for E1 and E2 in panel c and E2 and calnexin in panel f.

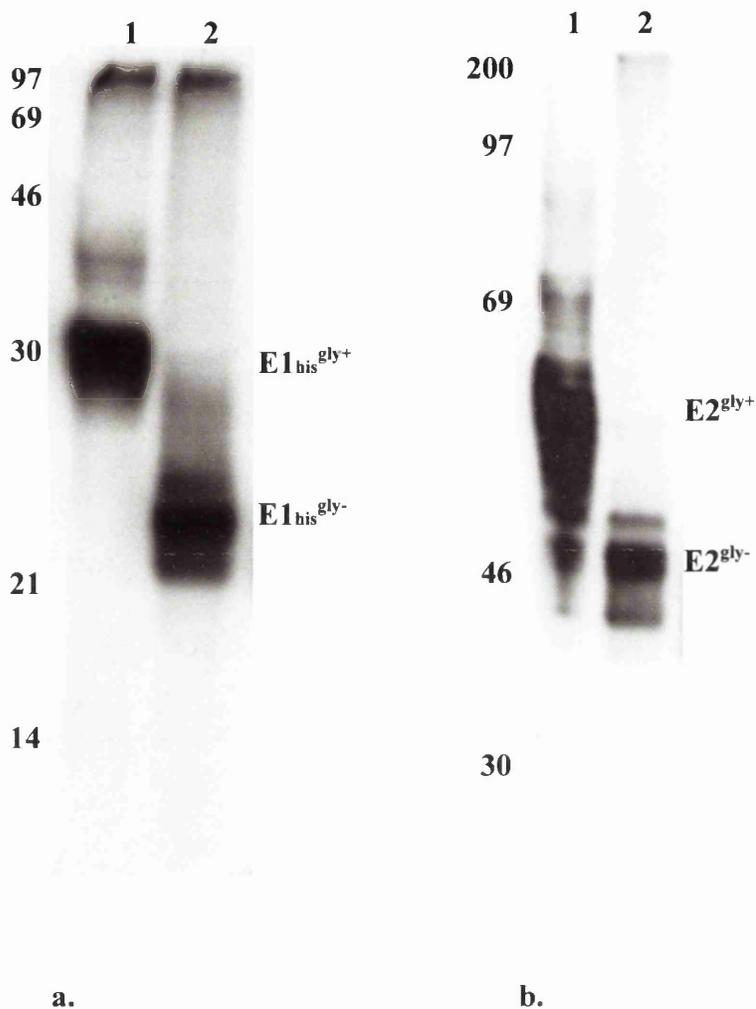


Fig 3.19. Endo H_f treatment of E1 and E2. Cells electroporated with RNA from pSFV/E1_{his}E2 were lysed 16 h post electroporation in binding buffer. Duplicate aliquots of crude extract were denatured and incubated with endo H_f. Equal volumes of endo H_f-untreated (lanes 1) and -treated (lanes 2) samples were separated by electrophoresis on a 12% polyacrylamide gel for E1_{his} (panel a) and a 9% polyacrylamide gel for E2 detection (panel b). Proteins were transferred to membrane and Western blot analysis was performed using antiserum R528 (panel a), and MAb ALP98 (panel b) to probe for E1 and E2, respectively. Glycosylated (gly⁺) and endo H_f treated (gly⁻) E1_{his} and E2 are indicated.

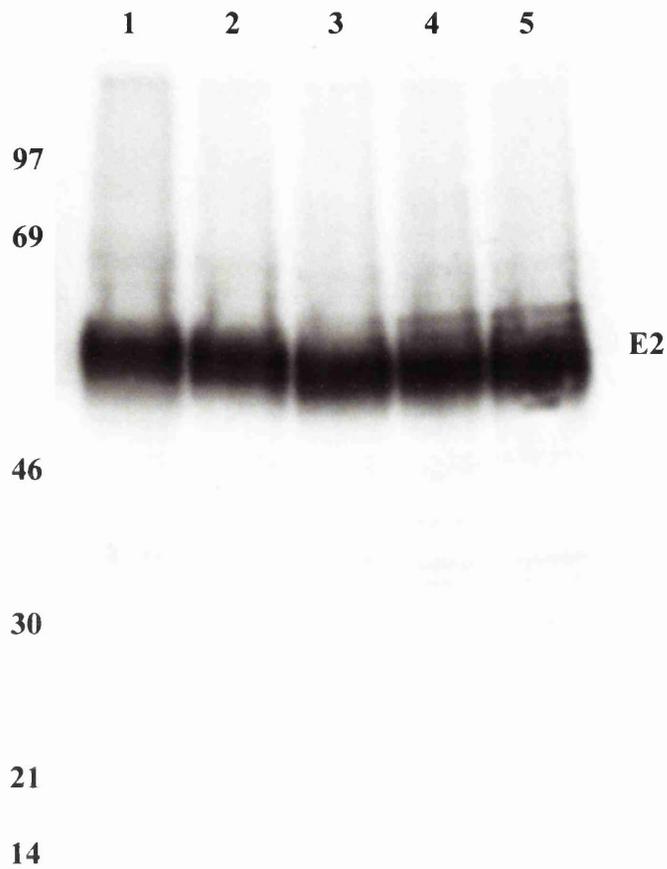


Fig 3.20. Analysis of the stability of E2 in the ER. pSFV/E1_{his}E2 RNA-electroporated cells were placed in media containing 200µg/ml CHX, to inhibit protein translation, 7 h after electroporation. Cells were harvested at 7 h (lane 1), 10 h (lane 2), 13 h (lane 3), 15 h (lane 4) and 18 h (lane 5) post electroporation. Aliquots of crude extracts were separated by electrophoresis and transferred to membrane for analysis by Western blot using E2-specific MAb ALP98. The position of E2 is indicated.

Chapter 4 - Characterisation of the Native and Aggregated Complexes Formed by E1 and E2

4.1. Introduction

At the start of this project studies by Grakoui *et al.* (1993) and Dubuisson *et al.* (1994) on strain H, a genotype 1a strain of HCV, had shown that intermolecular disulphide bonds existed in complexes formed by E1 and E2. However, Ralston *et al.* (1993) presented contrary evidence suggesting the heteromeric complex contained non-covalent E1E2 interactions. In the light of such contrasting observations, the nature of the interactions between the glycoproteins of strain Glasgow were studied by comparing their mobilities under reducing and non-reducing electrophoresis conditions. Under non-reducing but denaturing conditions, disulphide-linked proteins remain covalently attached, while non-covalently associated complexes would dissociate (see Section 1.3.5). Thus, by altering the reducing conditions in polyacrylamide gels both types of complex can be discriminated and analysed. To extend this analysis, the behaviour of E1E2 complexes was examined also *in situ* by adding a reducing agent, DTT, directly to cells. The effect of adding this reagent during protein synthesis can reversibly disrupt disulphide bonds (Braakman *et al.*, 1992b).

For comparative purposes, the behaviour of the glycoproteins produced from strain H77, which is also genotype 1a, were examined alongside the proteins produced from strain Glasgow. The cDNA clone of strain H77 is infectious in chimpanzees (Yanagi *et al.*, 1997) and was obtained later in the course of this study.

4.2. Behaviour of Strain Glasgow E1_{his} and E2 under Non-Reducing Electrophoresis Conditions

Extracts from cells electroporated with pSFV/E1_{his}E2 RNA were prepared in binding buffer containing NEM, an alkylating agent which blocks free sulphhydryl groups of cysteine residues. This compound was added to solutions to prevent aberrant disulphide bond formation between cysteine residues on proteins during preparation of crude extracts. Extracts were then mixed with Ni-NTA agarose from which bound proteins were eluted. Examination of bound material under reducing electrophoresis conditions (see Section 2.2.21) revealed the presence of E1_{his} and E2 (Fig 4.1, lane 1, compare with Section 3.8,

Fig 3.16). However, little monomeric E1 and E2 were detected under non-reducing conditions and most of the eluted material purified from Ni-NTA agarose failed to migrate into the resolving component of the gel (labelled Agg; Fig 4.1, lane 2). This indicated that the E1E2 complex contained intermolecular disulphide bonds. Therefore, the majority of the E1E2 complexes formed by strain Glasgow are characteristic of an aggregated complex and not a native complex (see Section 1.3.7).

To address the question of whether this high abundance of covalent E1E2 complex was an artefact of the time at which cells were harvested during SFV-mediated expression of the proteins, extracts were obtained at regular time points after electroporation with pSFV/E1_{his}E2 RNA. The presence of E1_{his} was examined by Western blot analysis, under both non-reducing and reducing conditions (Fig 4.2). Under reducing electrophoresis conditions, E1_{his} was not detected until 6 h after electroporation and its abundance increased substantially by 8 h (Fig 4.2, lanes 8-10). Thereafter, there was a further slight rise in amount at each of the subsequent time points (Fig 4.2, lanes 11-14). Under non-reducing electrophoresis conditions, E1_{his} did not migrate as a discrete band but rather was identified as diffuse material (Fig 4.2, lanes 1-6). These are presumably oxidised forms of E1_{his} protein in a variety of conformational states. The diffuse material was not readily detected until 8 h after electroporation and did not increase in intensity at later time points. Moreover, there was much less monomeric E1 found under non-reducing as compared to reducing electrophoresis conditions (Fig 4.2, compare lanes 1-6 with lanes 9-14). The majority of E1 detected by the antibody was found in a high molecular weight band during the time course (labelled Agg in Fig 4.2). This represented material that failed to migrate into the gel and corresponds to E1 which is covalently attached by intermolecular disulphide bonds to itself or other proteins. Thus, the low abundance of non-covalently associated E1E2 complex is not a feature of extended expression by SFV. The diffuse nature of E1 under non-reducing conditions may also explain the inability to readily observe any E1 in the non-reduced sample in the previous experiment (Fig 4.1, lane 2). Hence, E1 and E2 encoded by strain Glasgow may produce some non-covalently bound complex although they predominantly exist in a covalently-associated complex.

4.3. Behaviour of E1E2 Complexes under Reducing Conditions *In Situ*

Thus far, results have indicated that disulphide bonds do seem to have a role in complex formation since aggregated material is primarily detected under non-reducing

electrophoresis conditions. Other studies on HCV E1 and E2 (Dubuisson *et al.*, 1994; Deleersnyder *et al.*, 1997) have found similar complexes as well as significant amounts of native complex. To further examine the role of disulphide bonds in E1E2 complex formation, the properties of the complex were characterised under reducing conditions *in situ*. From previous studies, treating cells *in situ* with DTT reversibly reduces disulphide bonds within proteins (Braakman *et al.*, 1992b). Therefore, a series of experiments were devised employing this method to analyse the effect of DTT addition on the stability of pre-formed E1E2 complexes.

Cells were electroporated with pSFV/E1_{his}E2 RNA and proteins radiolabelled with ³⁵S-methionine 4 h after electroporation. Labelling was continued for a further 8 h at which point CHX was added to block translation. Following incubation for 30 min in the presence of CHX, DTT was added to cells at a final concentration of 5mM and cells were harvested after a further 30 min incubation. In parallel, another culture was treated with CHX but no DTT was added at any stage. Extracts were then prepared in the presence of NEM and, following purification by Ni-NTA affinity chromatography, bound proteins were analysed under both reducing and non-reducing conditions (Fig 4.3). The E1E2 complexes purified from cells that had not been treated with DTT behaved in an identical manner to those shown in Fig 4.1. E1 and E2 migrated as monomers under reducing conditions (lane 3) but mainly as a high molecular weight complex in the absence of reducing agent (lane 1). From the extracts of cells treated with DTT, E2 continued to co-purify with histidine-tagged E1 as both proteins are found in gels run under reducing conditions (lane 4). Significantly, under non-reducing conditions, E1 and a band presumed to be E2 no longer migrated as a high molecular weight complex but rather as monomeric proteins (lane 2). The presence of E2 in non-reducing gels was verified in a separate experiment. Cells were electroporated with pSFV/E1_{his}E2 and treated with DTT as described in the experiment for Fig 4.3. Following purification by Ni-NTA chromatography, eluted samples were analysed by Western blot analysis for E2 using an anti-E2 antiserum (R141). Under non-reducing electrophoresis conditions, E2 could be detected both as a high molecular weight species and as a monomeric protein; in a reducing gel, E2 was found to exist almost exclusively as a monomeric species (Fig. 4.4, lanes 1 and 2). This confirms that DTT treatment of cells reduces intermolecular disulphide bonds in the HCV glycoproteins but E2 and E1 remain associated. Therefore, DTT treatment of cells *in situ* can break the intermolecular disulphide bonds between E1

and E2 molecules but does not disrupt the complex they form. Hence, within the aggregated complex, non-covalent interactions do exist between E1 and E2.

Following this observation, the effect of DTT on E1E2 complexes was further characterised. In the following experiment, the result of varying the concentration of DTT added to cells on the aggregated material and the non-covalent E1E2 interactions was examined. Again, radiolabelled complexes from pSFV/E1_{his}E2-electroporated cells were isolated by Ni-NTA chromatography and analysed under non-reducing conditions to monitor the loss of high molecular weight material (Fig 4.5). At the lowest concentration of DTT (0.1mM), there was no evidence that reduction of E1 and E2 had occurred *in situ* and E1 and E2 remained as high molecular weight material at the top of the gel (lane 3). However, addition of DTT to 0.5mM to cells led to the appearance of monomeric E1 (lane 4) and at higher concentrations of DTT, there was a progressive increase in the amount of both monomeric E1 and E2 which entered the gel (lanes 5-8). These data also show that even at relatively high concentrations of DTT (5mM and 10mM), the non-covalently-linked E1 and E2 complex is stable. In subsequent experiments, DTT was added to cells at a final concentration of 5mM, a concentration that is consistent with that used in other studies (Braakman *et al.*, 1992b).

To determine the time required to fully reduce the aggregated material in 5mM DTT, cells electroporated with pSFV/E1_{his}E2 RNA were radiolabelled and then treated with DTT and CHX for various lengths of time. Analysis of isolated complexes revealed that reduction was achieved by 5 min following the addition of DTT to cells (Fig 4.6, lane 4) and is effectively complete by 20 min (lane 6). Taken together, the results in this and the previous experiment provided parameters for further studies on the effect of DTT on E1E2 complexes and gave conclusive evidence that non-covalent interactions between E1 and E2 are not dependent on intermolecular disulphide bonds for stability.

4.4. Effect of Removal of DTT from Cells on E1E2 Complexes

It is possible that *in situ* reduction of E1 and E2 by DTT addition to cells mediates a change in either the conformation of the glycoproteins or the complex which they form, such that reformation of disulphide bonds is not possible. In the previous experiments, the effect of DTT on the disruption of the pre-formed complexes was examined, where cells were incubated in the presence or absence of DTT and harvested immediately after its

removal. Here, cells expressing E1 and E2 synthesised from pSFV/E1_{his}E2 RNA were treated with CHX and DTT, after which the cultures were washed with medium containing CHX but lacking DTT and then harvested at various times following removal of the reducing agent (Fig 4.7). As described in the previous experiments, treatment of cells with DTT and CHX induced the appearance of monomeric E1 and E2 (compare lanes 2 and 3). Upon removal of DTT from cells, reformation of aggregated material could be detected within 5 min of DTT removal (lane 5) and it was the predominant species by 15 min (lane 6). Hence, the effect of DTT is reversible and disulphide bond formation recurs upon its removal. However, non-aggregated forms of the complex are not recovered, suggesting that even transient release from possible incorrect intermolecular covalent linkages does not allow re-ordering of the non-covalent interactions to generate complexes which do not aggregate.

4.5. Disulphide Bond Formation is not Required for the Initial Interactions Between E1 and E2

The data described thus far examined the function of disulphide bonds in pre-existing complexes. To address any role for such covalent linkages in initial interactions between E1 and E2, complexes formed by *de novo* synthesised E1 and E2 in the presence and absence of DTT were analysed. After incubation at 37°C for 12 h, pSFV/E1_{his}E2-electroporated cells were radiolabelled for 1 hour in the presence of DTT; a parallel culture was incubated in radiolabelling media for the same period but in the absence of DTT. Extracts made in the presence of NEM were applied to Ni-NTA agarose and the bound material was divided for analysis under non-reducing and reducing conditions (Fig 4.8). Synthesis of the E1_{his}E2 polyprotein in the presence of DTT did not inhibit its processing to produce monomeric E1 and E2 (compare lanes 3 and 4). The behaviour of the complex made in the absence of DTT was similar to that described previously (Fig 4.1, lane 2), where the majority of the complexes remained as high molecular weight material under non-reducing electrophoretic conditions (Fig 4.8, lane 2). Significantly, complexes purified from DTT-treated cultures showed that the inability to form intermolecular disulphide bonds did not prevent association of E1 and E2 (lanes 1 and 3). It is presumed that DTT also prevents intramolecular disulphide bond formation within the individual E1 and E2 glycoproteins. Hence, these data underline the ability of the two proteins to form non-covalent interactions in the absence of covalent disulphide linkages.

To further characterise the behaviour of *de novo* synthesised E1 and E2 in complex formation, electroporated cells were pulse-labelled for 2 min with ^{35}S -methionine/cysteine and then chased for up to 10 min. To prevent disulphide bond formation during the pulse/chase period, DTT was added to one set of cells 15 min prior to addition of the radiolabelling amino acids and was maintained until the cells were harvested. In this experiment, immunoprecipitations were performed with E2-specific antiserum R141 as opposed to use of Ni-NTA chromatography. Results showed E1 was associated with E2 by 2 min after the pulse in both DTT-treated and non-treated samples (Fig 4.9, panels c and d, lane 2). The presence of monomeric E1 and E2 in non-reducing gels verified that intermolecular disulphide bonds had not formed in proteins synthesised in the presence of DTT (panel a, lanes 2-4). In comparison with samples not treated with DTT, only small quantities of E2 and no E1 could be detected in monomeric form in non-reducing gels (panel b, lane 2) and the abundance of the monomeric E2 protein decreased at longer chase times (panel b, lane 4). The lack of monomeric E1 and the reduced amount of monomeric E2 results from the formation of aggregated material, which barely enters the resolving gel (panel b, lanes 2-4). It was also evident that monomeric E2 isolated from untreated cells consisted of diffuse material whereas E2 from DTT-treated cells was a discrete band (compare panels a and b, lanes 2-4). The diffuse nature of E2 in panel b is likely to result from intramolecular disulphide bond formation, which gives rise to various oxidised forms of the protein. The appearance of E2 as a discrete band in the samples from DTT-treated cells is indicative that intramolecular disulphide links are not present in the protein. From these data, it was concluded that non-covalent association of E1 and E2 occurs rapidly following synthesis and disulphide bond formation, both inter- and intramolecular, is not necessary for such interactions.

4.6. Comparison of Glycoproteins Produced from Strains Glasgow and H77

During the course of this study, it was reported that RNA synthesised from cDNAs comprised of the consensus sequence of HCV strain H77 were infectious in chimpanzees (Kolykhalov *et al.*, 1997; Yanagi *et al.*, 1997). Published data also showed that both covalently- and non-covalently-linked E1E2 complexes could be identified with the HCV H strain which was isolated from a chimpanzee infected with strain H77 and has very high

homology to H77 (Inchauspe *et al.*, 1991; Dubuisson *et al.*, 1994). Sequence comparisons of E1 and E2 between the Glasgow and H77 strains revealed that there are 9 amino acid residue differences in E1 and 30 in E2 (Appendix 2). These changes do not affect the number or the position of the cysteine residues in either of the proteins for both strains. Indeed, the cysteine residues in the glycoproteins appear to be highly conserved throughout type 1a strains of HCV (Robertson *et al.*, 1998). However, the differences in amino acid sequences at residues 476, 478 and 532 predict that E2 from strain H77 has two additional glycosylation sites as compared to strain Glasgow E2 (Ogata *et al.*, 1991; M. McElwee and R.M. Elliott, unpublished data). In E1, the variations in the amino acid sequences do not affect the predicted glycosylation sites.

The initial strategy was to clone an *ApaLI-Eco47III* fragment from pCV-H77C (supplied by J. Bukh; Yanagi *et al.*, 1997), which encoded residues 1-837 and thus contained the coding region for the structural proteins as well as sequences for p7 and the N-terminus of NS2 (Appendix 3). The strain H77 DNA fragment was inserted into pGEM1 (Fig 3.3, panel a) which had been modified to include *ApaLI* and *Eco47III* restriction sites and in-frame stop codons following the *Eco47III* site within the multiple cloning region (as shown in Fig 3.3, panel c). This construct was termed pGEM/CE1E2_{H77}. *BglIII* sites (Fig 3.3, panel c) flanking the HCV protein coding region allowed the coding sequences to be transferred into the *BamHI* site of pSFV1 (Fig 3.3, panel b); the resultant plasmid was named pSFV/CE1E2_{H77}. For later experiments on E1 (Chapter 5), the sequences encoding E1 to NS2 were excised as a *FspI-BglIII* fragment and reintroduced into the pGEM1 vector along with a PCR fragment to reconstitute the E1 signal sequence (Fig 4.10). A unique *NsiI* site was included in the forward primer and placed upstream of the E1 signal peptide coding sequence. This facilitated subsequent cloning directly into the pSFV1 vector (see Section 5.4.1). Using the flanking *BglIII* sites, the coding sequences were introduced into pSFV1, giving a construct named pSFV/ Δ E1E2_{H77}. Electroporation of RNA from this construct into cells did not give rise to detectable levels of recombinant proteins as determined by Western blot analysis and immunofluorescence studies. Sequencing data did not indicate any sequence errors which would give aberrant frame shifts or deletions. The low level of expression may have been due to the context of the AUG start codon or inappropriate secondary structure formation at the 5'-terminus of the RNA molecules, thus inhibiting translation. In light of this, the authentic start codon of the HCV polyprotein and the first 65 amino acid residues of core protein were fused to the N-terminus of the E1 signal sequence. An oligonucleotide was employed to bridge the core and E1 coding

sequences as described in Fig 4.11. The resultant construct was called pGEM/E1E2_{H77} and contained residues 1-65 and 169-837 of the strain H77 coding regions; 3 non-HCV residues were incorporated between amino acids 65 and 169. A *Bgl*III fragment carrying the HCV coding sequences was introduced into the pSFV1 vector and the construct was named pSFV/E1E2_{H77}.

To compare the E1 and E2 species made by strain H77 with those of strain Glasgow, cells were electroporated with pSFV/CE1E2_{H77}, pSFV/CE1E2_{Gla} and pSFV/E1_{his}E2 RNAs and the proteins radiolabelled with ³⁵S-methionine from 4-12 h post electroporation. Extracts prepared from cells were incubated with an E2-specific monoclonal antibody, ALP98 (Appendix 5). Electrophoresis of immunoprecipitates under reducing conditions showed that E2 made by strain H77 had a greater apparent molecular weight than E2 from strain Glasgow (Fig. 4.12, compare lane 5 with lanes 4 and 6). This increased size for strain H77 E2 is assumed to result from the two additional glycosylation sites in the protein as compared to strain Glasgow (Ogata *et al.*, 1991; M. McElwee and R.M. Elliott, unpublished data; see above). By contrast, strain H77 E1 had a lower apparent molecular weight than that for strain Glasgow (Fig. 4.12, compare lane 5 with lanes 4 and 6). The difference in mobilities for the E1 glycoproteins may not be a consequence of the presence of the histidine tag since untagged E1 made by strain Glasgow also has a greater apparent molecular weight than strain H77 E1. The predicted molecular weights of nascent and fully glycosylated E1 and E2 proteins from strains Glasgow and H77 are shown in Table 4.1. This reveals that the predicted size of unmodified E1 from both strains is identical, although the extra histidine residues in strain Glasgow E1_{his} do contribute to an increase in mass of the nascent protein. However, the difference in migration for unmodified E1 between the two strains, suggests there may be additional processing to the E1 protein from strain Glasgow. This issue is addressed further in Chapter 5 (Section 5.3.2).

Analysis of the complexes made by the two strains under non-reducing conditions revealed little, if any, monomeric E1 and E2 derived from strain Glasgow and the bulk of the complexes migrated at the top of the gel (Fig. 4.12, lanes 1 and 3). This is consistent with data presented earlier (Fig 4.1) and again shows that this property of the E1E2 complex produced by strain Glasgow is not due to the histidine tag in E1. A high proportion of the glycoproteins made by strain H77 also ran at the top of the gel, however, monomeric E2 and E1 species also could be readily observed (Fig. 4.12, lane 2). Hence, for strain H77, both covalently-linked and non-covalently-linked E1E2 complexes were detected.

In the previous experiment, MAAb ALP98, raised against E2 from strain Glasgow, was able to precipitate the glycoproteins from both strains. Further comparisons between E1 and E2 encoded by strains H77 and Glasgow were undertaken using a series of human and mouse monoclonal antibodies, which recognised conformational epitopes (supplied by M. Cardoso and J. Dubuisson, Table 4.2). Using extracts prepared from pSFV/CE1E2_{Gla} and pSFV/C1E2_{H77} RNA-electroporated cells, none of the antibodies precipitated the strain Glasgow glycoproteins (Fig. 4.13 and 4.14, lanes 1-4). However, E1E2 complexes from strain H77 were recognised by the antibodies but to variable extents (summarised in Table 4.2). In particular, two human (1A7 and 2H1; Fig 4.14, lanes 5 and 6) and one mouse antibody (H33; Fig 4.13, lane 6) had particularly high affinities. The strain-specificity of these antibodies precluded their further use in comparing the properties of the complexes made by strains Glasgow and H77. However, one antibody, H53, which reportedly has a preferred affinity for the non-covalent complex (Cocquerel *et al.*, 1998) was used in subsequent analysis of strain H77 E1E2 interactions.

4.7. Behaviour of E1 and E2 Produced by Strain H77 in the Presence and Absence of DTT *In Situ*

The previous section described data indicating that strain H77 produced both covalent and non-covalent heteromeric E1E2 complexes. To examine the effect of DTT on pre-formed complexes made by strain H77 glycoproteins, cells were electroporated with RNA from pSFV/CE1E2_{H77}, and, for comparative purposes with pSFV/CE1E2_{Gla} and pSFV/E1_{his}E2 RNAs. Following radiolabelling, cell cultures were treated with DTT. Cell extracts were prepared and complexes isolated using the E2-specific antibody, ALP98 (Fig. 4.15). This revealed that E1 continued to precipitate with E2 for strain H77 in the presence of DTT, indicative that, as for strain Glasgow, disulphide bonds are not required to stabilise the intermolecular interactions between the glycoproteins (Fig. 4.15, lanes 4-6). Under non-reducing electrophoresis conditions, there was an increase in the abundance of E2 detected as compared to analysing the equivalent untreated sample also in a non-reducing gel (compare Fig 4.12 and Fig 4.15, lanes 2). This is presumably because of E2 dissociation from the aggregate in DTT-treated samples. In agreement with the data for strain Glasgow (Fig. 4.15, lanes 1 and 3), DTT can reduce disulphide bonds in aggregated E1E2 complexes. In addition to bands specific for E1 and E2, three additional novel bands were

present in the lanes representing strain H77 samples treated with DTT (lanes 2 and 5; marked by ◆, *, and ■). The * species was found consistently for this strain when the samples were prepared after DTT treatment. The band indicated by ◆, is likely to represent glycosylated but uncleaved E1E2 precursor, while ■ may correspond to incompletely glycosylated E1. Glycosylation of E1 is addressed further in the following chapter 5 (Section 5.2.2)

Detection of E2 by ALP98 is not abolished by DTT treatment *in situ*, presumably because the antibody recognises a linear epitope (Appendix 5). To analyse the effect of DTT on detection of the HCV glycoproteins by other antibodies, including those that are conformation-specific, cells were electroporated with pSFV/CE1E2_{H77} and pSFV1 RNAs. 12 h post-electroporation parallel cell cultures were pulse-labelled for 2 min and chased for a period of 80 min in the presence and absence of DTT. Crude extracts derived from both sets of cells were analysed by immunoprecipitation with H53 as well as with anti-E1 antiserum R528, E1-conformation sensitive MAb AP497, and E2-specific ALP98 (see Appendix 5). Analysis of precipitates under reducing conditions revealed that for non-DTT-treated samples equivalent amounts of E2 were precipitated by the E2-specific MAbs H53 and ALP98 (Fig 4.16, panel b, lanes 3 and 4). In addition similar amounts of E1 (and associated E2) were precipitated by the E1-specific antibodies R528 and AP497 (panel b, lanes 1 and 2). However, for samples derived from DTT-treated cells, reduced levels of E1E2 complexes were detected by the conformation-specific E1 MAb AP497 and E2 MAb H53 (Fig 4.17, panel a, lanes 1 and 4) as compared to R528 and ALP98 precipitated material (lanes 2 and 3). This data suggests that the majority of E1 and E2 do not form the conformations recognised by AP497 and H53 in the reduced state. This indicates that disulphide bond formation may be necessary to gain and stabilise the correct conformations in E1 and E2. However, a proportion of the glycoproteins did achieve the conformation necessary for immuno-detection, in particular with H53. Therefore, this antibody, along with ALP98, was used in a subsequent experiment to monitor complex formation at various times after synthesis.

To study the role of disulphide bonds at early stages of complex formation for the H77 glycoproteins, cells electroporated with RNA from pSFV/CE1E2_{H77} were chased for various periods following pulse labelling. Parallel cultures were radiolabelled in the presence of DTT to examine the affect on complex formation in the absence of disulphide

bond formation. Immunoprecipitations were carried out on crude extracts with the two E2-specific antibodies, ALP98 and H53.

In the absence of DTT, H53 recognised a native form of the complex more efficiently than ALP98, as deduced from the higher abundance of E1 observed under non-reducing conditions (Fig 4.17, panels a and b, lane 4). The presence of E2 was obscured by other species in the immunoprecipitates. This complex was most apparent at later times in the pulse chase experiment in agreement with previously reported data (Deleersnyder et al., 1997). It was found that, in these experiments, H53 precipitated also some aggregated material (panel b, lanes 1-4). In the presence of DTT, both antibodies precipitated E1 as well as E2 even at early chase times, although ALP98 was more efficient than H53 as was observed in the previous experiment (panels a-d, lanes 5 and 6). Nonetheless, in agreement with earlier data, DTT treatment *in situ* does not prevent association between E1 and E2. From comparison of the relative mobilities of E1 in Fig 4.17, it is evident that the oxidised form of E1 is present in untreated samples while only the reduced form is detected in those samples treated with DTT (panel b, compare lanes 3 and 4 with lanes 7 and 8). This indicates that intramolecular disulphide bonds do not form in the presence of DTT, further highlighting that E1 and E2 association does not rely upon the formation of covalent linkages. This conclusion was confirmed in subsequent experiments (Section 5.4.2).

4.8. Discussion

Comparison of the mobilities of strains Glasgow and H77 E1 and E2 glycoproteins revealed a difference in size between corresponding proteins for the two strains. Although an increase in mass is predicted in E1 from strain Glasgow on account of the histidine tag, the slower migration is also observed for untagged E1 from the same strain. However, the predicted molecular weights of the nascent chains of E1 from strains Glasgow and H77 are almost identical. The predicted number of potential N-linked glycosylation sites in E1 for both strains also is conserved. E2 from strain Glasgow is predicted to have two fewer N-glycosylation sites than its counterpart from strain H77 and a change in glycosylation pattern could account for the difference in their mobilities. This cannot apply to E1, nor can the possibility of aberrant processing from the polyprotein, considering E2 is of the

appropriate size. Therefore, additional processing to the E1 protein from strain Glasgow may occur and this issue is addressed in the following chapter.

Previous studies on the E2 species have reported inefficient processing at the E2/p7 cleavage site for strain H but not for strain BK (Dubuisson *et al.*, 1994; Lin *et al.*, 1994; Dubuisson & Rice, 1996). Thus, the E2 species detected for strain H is a combination of E2 and E2-p7 proteins. Based on the apparent molecular weights of E2 from strains H77 and Glasgow, both in the glycosylated and deglycosylated states (Table 4.1), it would appear that a proportion of E2 is not cleaved from p7. The precise ratio of E2 to E2-p7 in the studies presented here is not possible to ascertain given the lack of any antibodies against p7. It is also difficult to identify the nature of the different cleavage efficiencies between strains H and BK; the amino acid sequences flanking the cleavage sites are identical in the two strains although there are 14 amino acid changes elsewhere in p7. This is in addition to the differences in amino acid sequence in E2. In light of the uncertainty as to the proportion of E2:E2-p7 made by strains Glasgow and H77 using the SFV system, it is assumed that both species could be present in extracts. For simplicity, bands of approximately 60-69 kDa identified by E2 antibodies have been referred to as E2 throughout this study with the caveat that both E2 and E2-p7 may be present in species of this size range.

Examination of complex formation by strain Glasgow E2 and histidine-tagged or untagged E1, under non-reducing conditions, revealed little evidence of native complexes and essentially only an aggregated complex was detected. This does not preclude the possibility that an antibody, which is specific for a particular conformation, may identify some native complexes formed by strain Glasgow (see Sections 5.3.3 and 6.1.4). By contrast, a native complex was detected for strain H77 glycoproteins. This suggests that the presence of the histidine tag was not in itself a cause for the reduction in native complex formation by strain Glasgow E1 and E2. The reduction could have been a consequence of differences in amino acid sequence between the strains; there are 9 and 30 residue changes in E1 and E2 respectively. Moreover, there are two fewer predicted N-glycosylation sites in E2 from strain Glasgow. Proper N-linked glycosylation and oligosaccharide trimming is necessary for glycoprotein folding as part of the ER quality control mechanism (Section 1.4.5; Hammond & Helenius, 1995). For example, influenza virus HA contains of seven N-glycosylation consensus sequences. Mutations that removed individual N-linked glycans affected the rate and efficiency of HA folding (Hebert *et al.*,

1997). Abolition of glycosylation at one site (residue 81 in HA) disrupted formation of a crucial disulphide bond between cysteine residues 67 and 76, which is close to this glycan. This inhibited folding and HA trimer formation. The additional two N-glycosylation sites present in E2 from strain H77 may also participate in protein folding, particularly as one of these sites (at residue 532) is conserved in other HCV genotypes. The second site is located within the second variable region in E2 (aa 476; see Fig 1.13). In a more recent study, analysis of the association of E1 glycosylation mutants with E2 indicated that the absence of a glycan at a single site was sufficient to impair native complex formation (Meunier *et al.*, 1999). Thus for strain Glasgow, the lack of two glycosylation sites as compared to strain H77 may account for the lower amount of native E1E2 complexes detected. Furthermore, the presence of p7 does not appear to contribute to the formation of either the native or aggregated complexes, as this peptide remains linked to a proportion of E2 from both strains Glasgow and H77. An earlier study showed that expression of aa 1-1207 of strain H by both Sindbis and vaccinia viral vectors generated native complexes that contained incompletely processed E2-p7 species (Dubuisson *et al.*, 1994).

In light of the detection of disulphide-linked aggregated complexes, the requirement for these covalent bonds was examined further by the use of DTT. Previous studies had indicated that the presence of the reducing agent, DTT, in cell culture media did not have significant effects on cellular functions such as translation, translocation, signal sequence processing, N-linked glycosylation and transport beyond the ERGIC (Braakman *et al.*, 1992a; Desilva *et al.*, 1993; Verde *et al.*, 1995). However, this agent does profoundly affect the redox potential of the ER that results in the reduction of cysteine residues in the active site of PDI, which catalyses disulphide bond formation (Section 1.4.3), and renders it inactive (Holst *et al.*, 1997). This prevents appropriate disulphide bond formation in proteins translocated to the ER lumen, including VSV G, influenza HA and CD8 ERGIC (Braakman *et al.*, 1992a; Tatu *et al.*, 1993; Verde *et al.*, 1995). These earlier studies using DTT for *in situ* reduction showed that *de novo* synthesised influenza virus HA and VSV G were retained in the ER in an unfolded state (Braakman *et al.*, 1992a). Following the removal of DTT, they folded correctly, as shown by conformation-specific antibodies. In addition, they were transported from the ER to the Golgi network, where they were assembled into proper oligomers detected on virions. Since folding of proteins can occur after removal of DTT, it was possible that E1 and E2 from strain Glasgow may fold and associate to give increased amounts of native complex upon removal of the reducing agent. However, E1E2 aggregated complexes, not native complexes, reformed relatively soon

after the removal of DTT, which suggests that manipulation of the intracellular environment does not overcome this intrinsic behaviour of E1 and E2 to form aggregates. In another study, using an alternative approach to reduce the levels of aggregated complexes, the abundance of calnexin, calreticulin or BiP was increased. However this did not induce any elevated levels of native complex (Choukhi *et al.*, 1998).

Investigation of the aggregated complexes formed by strains Glasgow and H77 E1 and E2 glycoproteins following reduction *in situ* showed that these proteins continued to co-precipitate. This was a somewhat unexpected result given that aggregated complexes were considered to represent association of E1 and E2 by disulphide links. Thus, DTT treatment has uncovered non-covalent interactions between E1 and E2 in the aggregated complex. There are a number of possibilities which may account for their continued association: a) the disruption of intermolecular disulphide bonds may have exposed regions of E1 and E2 which can then interact non-covalently, b) E1 and E2 may form homo-oligomers which retain the ability to interact non-covalently with each other, and c) in addition to covalent interactions, E1 and E2 associate via non-covalent interactions simultaneously in the aggregated material. Possibility b) is unlikely since analysis of homo-oligomerisation of E1 shows that such forms can be detected more readily upon expression of E1 alone than in combination with E2 (see Section 5.2.2). It is difficult to distinguish between possibilities a) and c) with the result presented in this chapter. However, in the following results section, mutation of cysteine residues in E1, such that disulphide bond formation is not possible, does not abolish complex formation but E2 in these complexes is in an aggregated state. Therefore, aggregated complexes can contain both non-covalent and covalent interactions between E1 and E2.

For two antibodies (AP497 and H53) that recognise conformational epitopes, disruption of disulphide bonds by DTT did reduce their capacity to recognise E1 and E2. Therefore, disulphide bond formation is important either to generate or stabilise particular conformations of the proteins. In a study using a truncated form of E2 as antigen, Lee *et al.* (1997) showed that sera from infected individuals retained the ability to recognise protein denatured by heat or SDS treatment. However, treatment of E2 with β -ME, a reducing agent, dramatically decreased recognition by most sera. This indicates an important role for disulphide bonds in generation of antigenic epitopes in the glycoproteins. The studies with DTT indicated that generation of such conformations is

not a pre-requisite for interaction between E1 and E2. However, these interactions may be those that occur in an aggregated but not a native complex.

Pulse-labelling experiments revealed that the aggregates for both strains Glasgow and H77 formed much earlier than the native complex in strain H77. The slower occurrence of native complexes has been described (Dubuisson & Rice, 1996), and it has been suggested that disulphide bond isomerisation in strain H E1 protein is rate limiting for E1E2 association in a native complex. For N-glycosylation mutants of influenza virus HA, the appearance of a conformational epitope appears more rapidly than in wild type protein (Hebert *et al.*, 1997). The mutant protein does not associate with calreticulin at early stages following synthesis and does not achieve a proper folded state (Hebert *et al.*, 1997). This was explained by the masking of calnexin-HA binding sites within the improperly folded protein. Thus for HCV E1 and E2, the speed of aggregated complex formation may prevent the ER chaperone-assisted folding mechanisms to operate (Section 1.4.4). This possibility is discussed further in Chapter 7.

Non-covalent interactions have been identified in the native E1E2 complex while both covalent and non-covalent associations are found in aggregated forms of E1E2. It is therefore possible that regions required for non-covalent interactions are identical in both types of complex, but they may be distinct from areas containing cysteine residues that are involved in intermolecular disulphide bond formation. In an attempt to distinguish between these regions, a series of E1 and E2 internal deletion mutants were made and are described in Chapters 5 and 6 respectively. Furthermore, the influence of E2 on E1 folding and vice versa was also further investigated in the following chapters.

Strain	Protein	Predicted MW of nascent protein (kDa)	Predicted MW of modified protein (~kDa)*
Glasgow	E1 (E1 _{his})	20.9 (21.9)	28.9 (29.9) [#]
	E2 (aa 384-745)	40.0	58.0
	E2-p7 (384-810)	47.1	65.1
H77	E1	20.9	28.9 [#]
	E2 (aa 384-745)	40.1	62.1
	E2-p7 (aa 384-810)	47.2	69.2

Table 4.1. Predicted molecular weights (MW) of E1 and E2 proteins made by strains Glasgow and H77. * It is assumed that addition of a high mannose oligosaccharide increased the molecular weight of nascent protein by ~2.0 kDa per glycosylation site (Kornfeld & Kornfeld, 1985). [#] It is assumed that glycosylation at amino acid position 325 in E1 does not occur as reported by Meunier *et al.* (1999).

Antibody	Type	Epitope	Specificity	Reactivity to strain	
				H77	Glasgow
H2	mouse monoclonal	Conformational	E2	+	-
H33	mouse monoclonal	Conformational	E2	+++	-
H47	mouse monoclonal	Conformational	E2	++	-
H53	mouse monoclonal	Conformational	E2	++	-
1A7	human monoclonal	Conformational	E1+E2	++++	-
2H1	human monoclonal	Conformational	E1+E2	++++	-
4F7	human monoclonal	Conformational	E2	-	-
4F1	human monoclonal	Conformational	E2	+	-
E1F	mouse monoclonal	Linear	E1	-	++++
E2G	mouse monoclonal	Linear	E2	-	++
ALP98	mouse monoclonal	Linear	E2	+++	+++

Table 4.2. Summary of the characteristics of the antibodies used in this study.

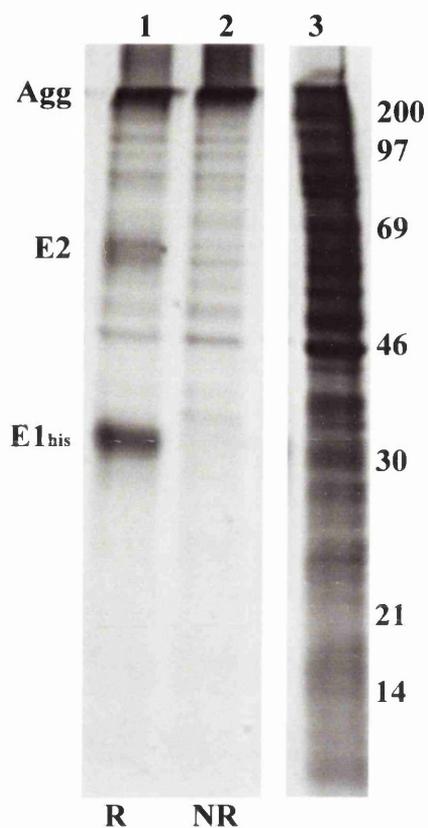


Fig 4.1. Mobilities of strain Glasgow E1_{his} and E2 examined under reducing and non-reducing electrophoresis conditions. Cells electroporated with pSFV/E1_{his}E2 RNA were radiolabelled between 4-12 h after electroporation with ³⁵S-methionine. Following lysis in the presence of 20mM NEM, soluble crude cell extract was applied to Ni-NTA agarose. Bound proteins were examined under reducing (R, lane 1) and non-reducing (NR, lane 2) conditions on a 12% polyacrylamide gel. Radiolabelled crude cell extract is shown in lane 3. The positions of E1_{his}, E2 and a high molecular weight aggregate (Agg), which fails to enter the resolving gel, are shown.

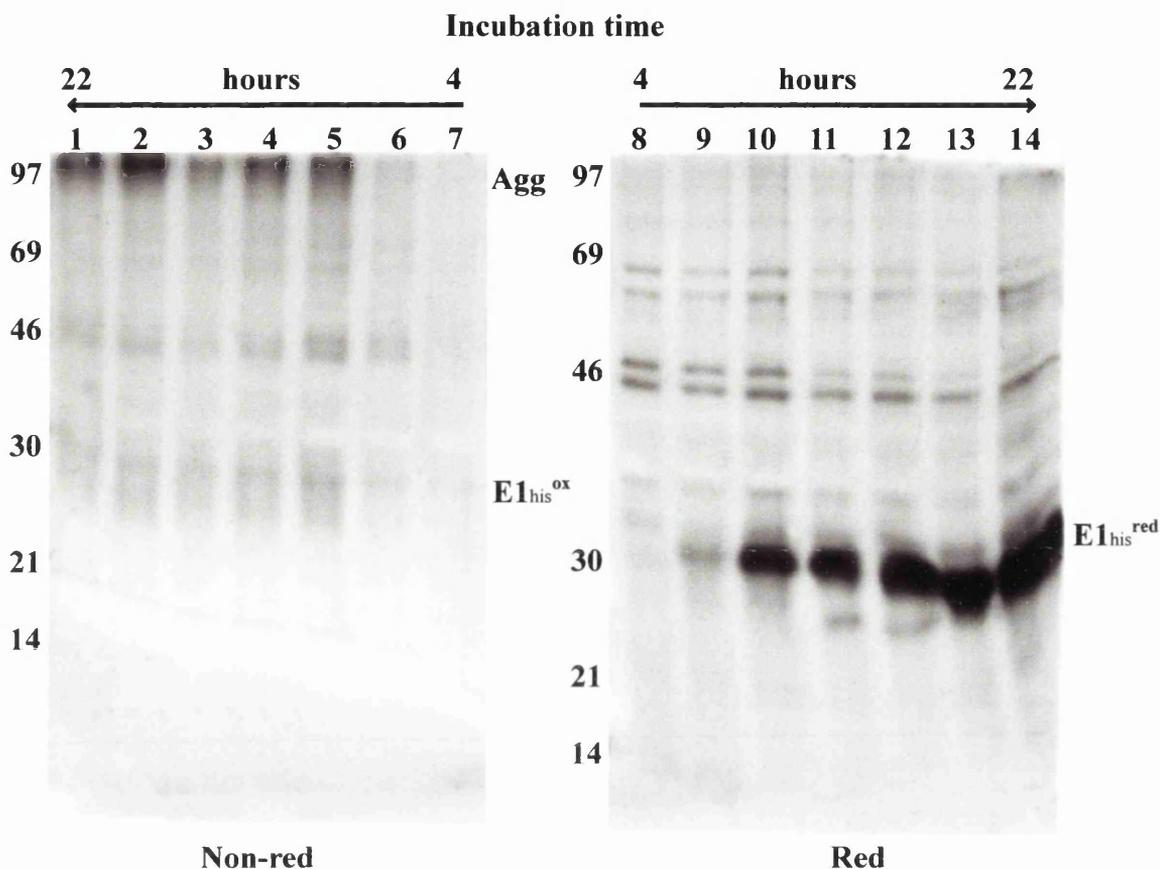


Fig 4.2. Examination of the presence of non-aggregated E1_{his} at various time points after electroporation. Cells electroporated with RNA from pSFV/E1_{his}E2 were harvested in the presence of 20mM NEM at 4 h (lanes 7 and 8), 6 h (lanes 6 and 9), 8 h (lanes 5 and 10), 10 h (lanes 4 and 11), 12 h (lanes 3 and 12), 18 h (lanes 2 and 13) and 22 h (lanes 1 and 14) post electroporation. Following removal of cell debris, duplicate aliquots of soluble material were subjected to electrophoresis on a 10% polyacrylamide gel under reducing (Red; lanes 8-14) and non-reducing (Non-red; lanes 1-7) conditions. Proteins separated under these conditions were transferred to membrane and examined by Western blot analysis using the Penta-His antibody, to detect histidine-tagged E1. The positions of the reduced (red), oxidised (ox) forms of E1_{his} and aggregated material (Agg) are indicated.

Fig 4.3.

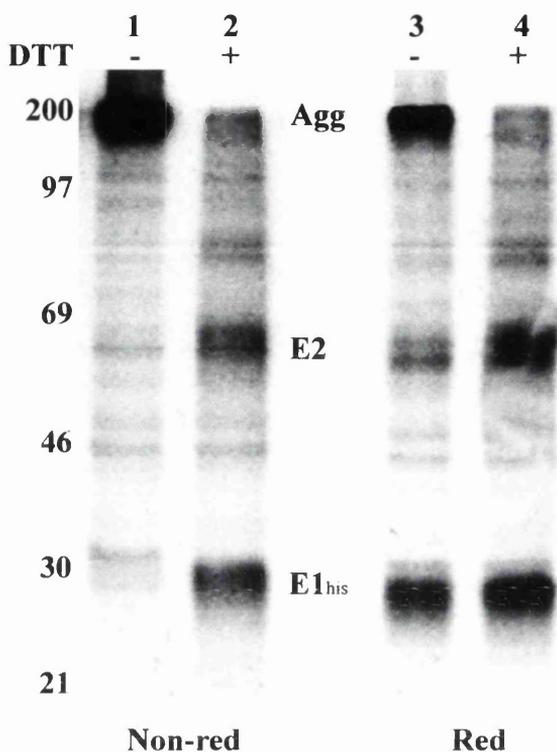


Fig 4.4.

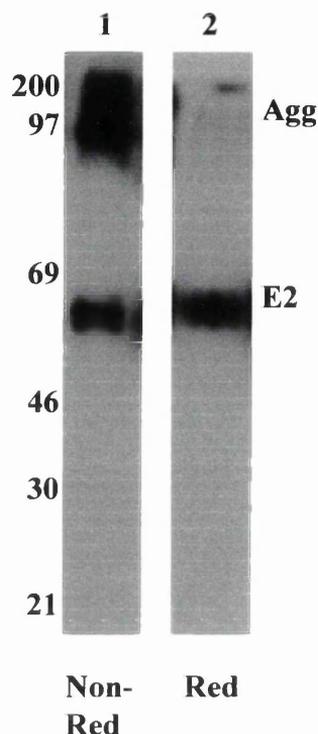


Fig 4.3. Effect of addition of DTT *in situ* on strain Glasgow E1-E2 complexes. Cells were electroporated with pSFV/E1_{his}E2 RNA, ³⁵S-methionine labelled between 4-12 h after electroporation and thereafter treated with CHX (200µg/ml) for 30 min prior to the addition of 5mM DTT. Cells were incubated in the presence of this reducing agent for 30 min. A parallel cell culture was incubated with CHX but not treated with DTT. Cell extracts were prepared in binding buffer containing 20mM NEM and then subjected to Ni-NTA affinity chromatography. Bound proteins were electrophoresed on a 10% polyacrylamide gel under non-reducing (Non-red, lanes 1 and 2) and reducing (Red, lanes 3 and 4) conditions. Samples derived from cells treated with (+) or without (-) DTT are indicated. The positions of E1_{his}, E2 and aggregated material (Agg), which fails to enter the resolving gel, are shown.

Fig 4.4. Western blot analysis of Ni-NTA eluted material derived from DTT-treated cells. Western blot analysis was performed on Ni-NTA resin eluted samples derived from cells treated with DTT and CHX. The eluted material was electrophoresed under reducing (Red) and non-reducing (Non-red) conditions. Material transferred to nitrocellulose membrane was analysed using anti-E2 antiserum, R141. The positions of aggregated material (Agg) and E2 are indicated.

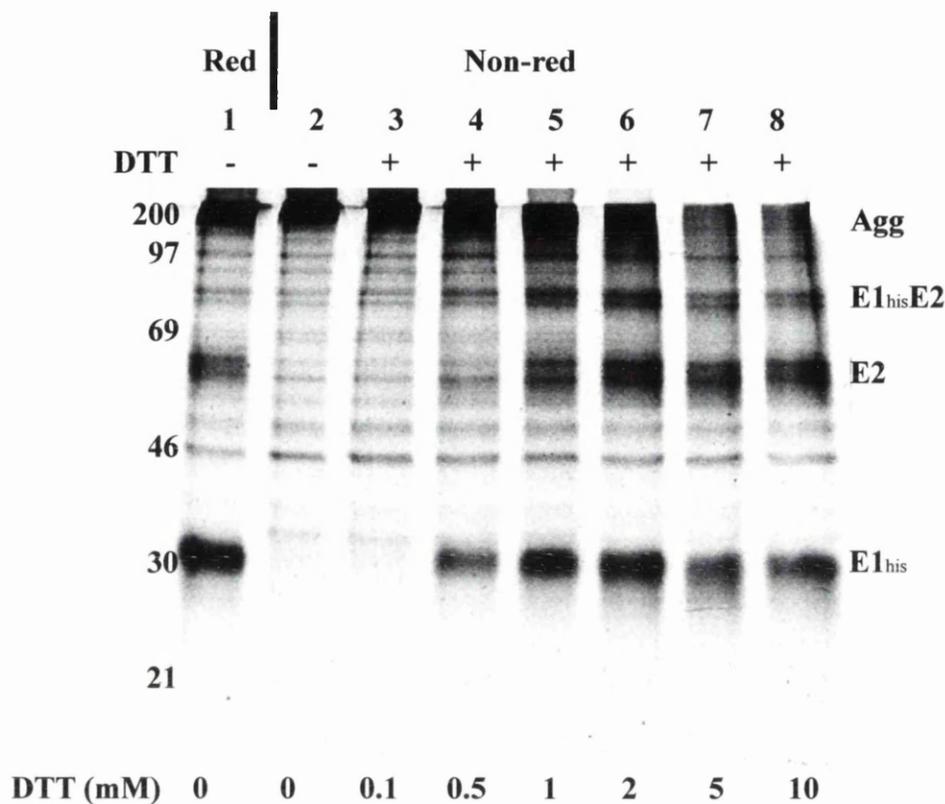


Fig 4.5. Effect on E1-E2 complexes with increasing DTT concentration *in situ*. Cells electroporated with RNA from pSFV/E1_{his}E2 were ³⁵S-labelled from 4-12 h after electroporation followed with incubation for 30 min with cell culture media containing 200µg/ml CHX. Subsequently, parallel cell cultures were incubated with increasing concentrations of DTT prior to harvesting in binding buffer containing 20mM NEM. The concentration of DTT added to each cell culture is shown below the corresponding lane. Following Ni-NTA affinity chromatography, samples were electrophoresed on a 10% polyacrylamide gel under reducing (Red, lane 1) and non-reducing conditions (Non-red, lanes 2 to 7). Lanes 1 and 2 show control samples derived from cells not treated with DTT. Positions of E1_{his}, E2, uncleaved precursor E1_{his}E2 and aggregate (Agg) are shown.

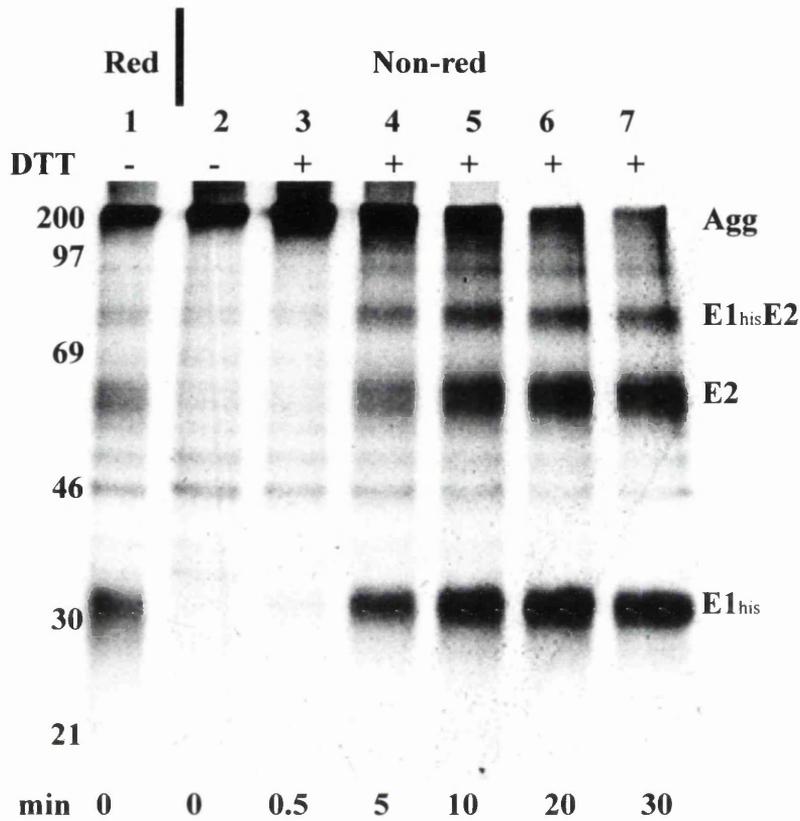


Fig 4.6. Effect on E1E2 complexes with increasing incubation time with DTT *in situ*. Experimental conditions were as described in Fig 4.5 except cells were incubated in 5mM DTT for varying times. Incubation periods are indicated below each lane. Crude extracts prepared in binding buffer were subjected to Ni-NTA chromatography and bound material was analysed under reducing (Red) and non-reducing conditions (Non-red). Lanes 1 and 2 show control samples derived from cells not treated with DTT. Positions of E1_{his}, E2, unprocessed precursor E1_{his}E2 and aggregate (Agg) are indicated.

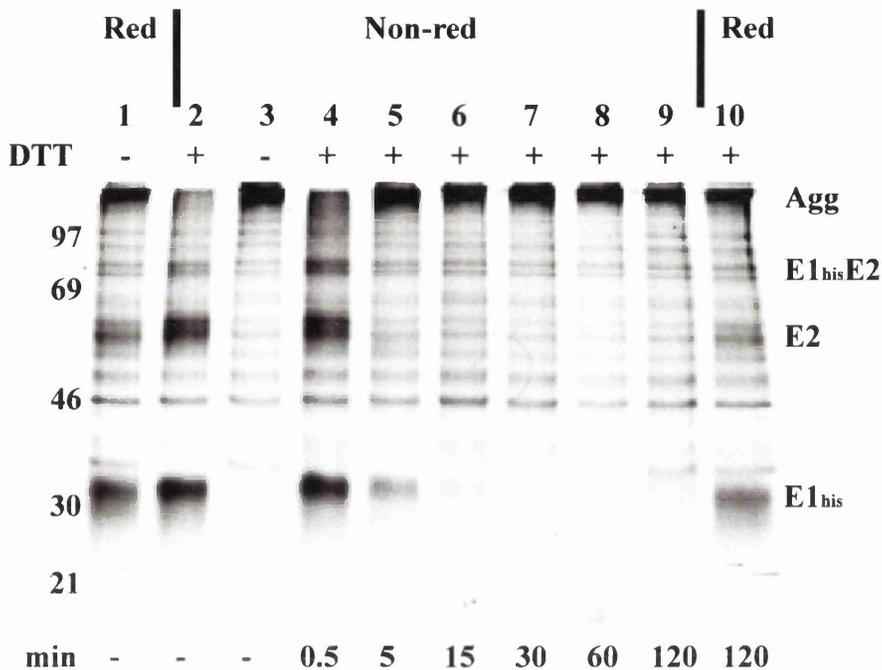


Fig 4.7. The effect of removal of DTT on E1E2 complexes. Cells were electroporated with pSFV/E1_{his}E2 RNA and ³⁵S-labelled between 4-12 h after electroporation. CHX and DTT were added to cells as described for Fig 4.3, followed by washing the cells with media lacking the reducing agent and incubation in the absence of DTT for the times indicated below each sample. Cells were lysed in binding buffer and subjected to Ni-NTA affinity chromatography. Bound proteins were analysed by electrophoresis on a 10% polyacrylamide gel under reducing (Red, lane 1) and non-reducing conditions (Non-red, lanes 2 to 9). Control samples were derived from non-treated cells (-) (lanes 1 and 3) and treated cells (+) (lane 2) harvested immediately after the removal of DTT. The sample in lane 9 was also examined under reducing conditions (lane 10). The positions of E1_{his}, E2, unprocessed E1_{his}E2 polyprotein and aggregated material (Agg) are indicated.

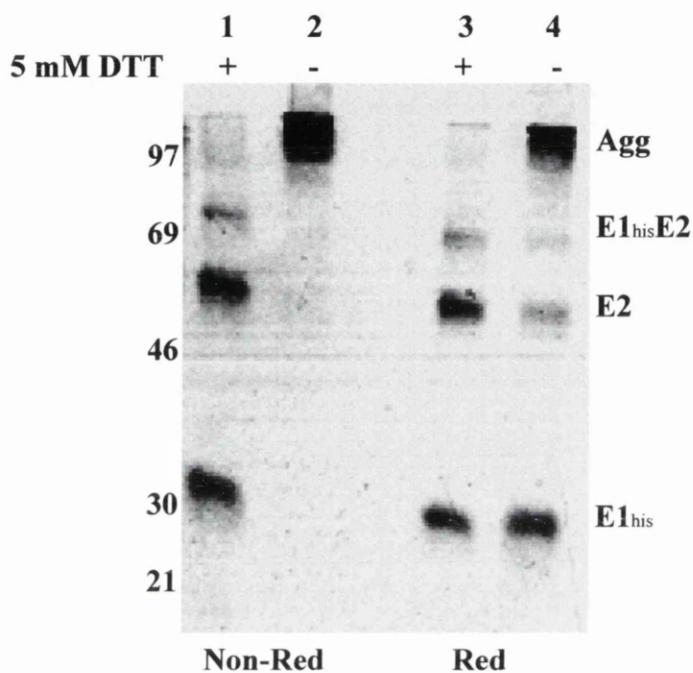


Fig 4.8. Effect of DTT on complex formation by *de novo* synthesised E1 and E2. Cells were electroporated with pSFV/E1_{his}E2 RNA and incubated at 37°C for 12 h. Subsequently, they were ³⁵S-labelled for 1 h in the absence (-) (lanes 2 and 4) or presence (+) (lanes 1 and 3) of 5mM DTT. Following cell lysis, extracts were subjected to Ni-NTA affinity chromatography and analysed under reducing (Red) and non-reducing (Non-red) conditions on a 10% polyacrylamide gel. The positions of E1_{his}, E2, E1_{his}E2 precursor and aggregated material (Agg) are indicated.

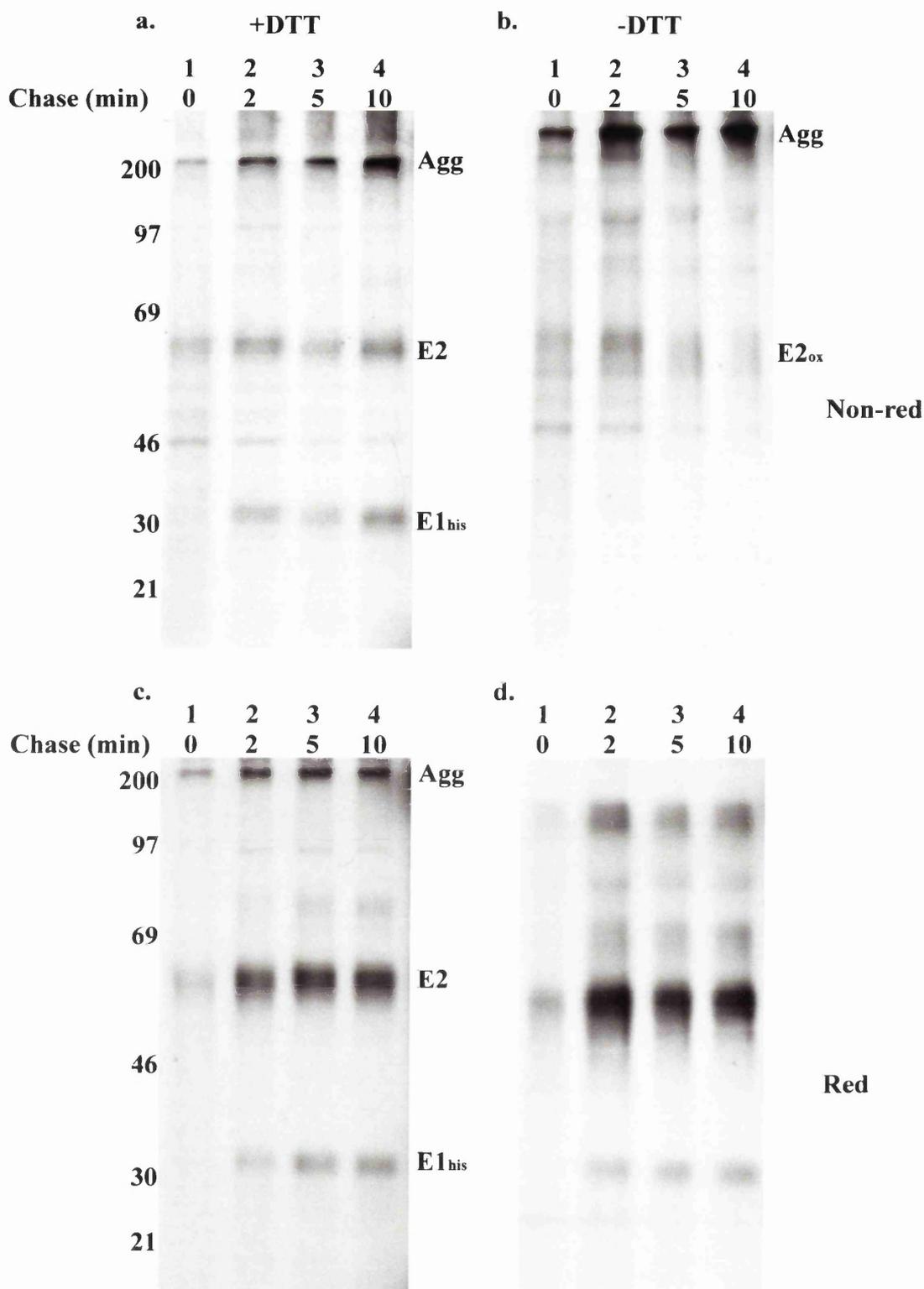


Fig 4.9. Characterisation of complexes formed by *de novo* synthesised E1 and E2 in the presence and absence of DTT. Cells were electroporated with pSFV/E1_{his}E2 RNA and incubated at 37°C for 12 h. Subsequently, cells were labelled for 2 min and then chased for the times indicated. In panels a and c, samples are from cells incubated in the presence of DTT (+DTT) from 15 min prior to labelling until they were harvested. Immunoprecipitations were performed on cell extracts with anti-E2 antiserum R141 and bound proteins were electroporated on 10% polyacrylamide gels under non-reducing (Non-red, panels a and b) and reducing (Red, panels c and d) conditions. The positions of aggregated material (Agg), E1_{his}, E2 and oxidised E2 (ox) are shown.

AAG CTT AGA TCT ATG CAT ATG CTT CCT GGT GCC TAC CAA GTG CGC AAC
M L P G A Y Q V R N
169 191 196

Fig 4.10. Cloning of the strain H77 E1 and E2 coding region into pGEM1. The *FspI*-*BglIII* segment encoding E1 to NS2 was excised from pGEM/CE1E2_{H77} and introduced into the vector backbone of pGEM/CE1E2_{H77} that had been cut with *BglIII*. A PCR fragment (over-lined in red) bridged the *FspI* (in the E1 sequence; residues in brown) and *BglIII* sites in the vector. The PCR product coded for the E1 signal peptide (residues in purple) and was generated using pGEM/CE1E2_{H77} as template. The forward PCR primer introduced *BglIII* and *NsiI* sites at the 5' terminus (nucleotides in italics) followed by an ATG start codon. The resultant construct was named pGEM/ΔE1E2_{H77}.

CAA CCT CGA GGT AGA CGT CAG CCT ATC ATG CAT ATG CTT CCT GGT TGC
Q P R G R R Q P I M H M L P G C
58 65 169

Fig 4.11. Fusion of a N-terminal region of core protein to the N-terminus of the E1 signal peptide. Two DNA fragments, a *HindIII*-*XhoI* fragment from pGEM/CE1E2_{H77} which encodes the N-terminal region of core (aa 1-58), and a *NsiI*-*EcoRI* fragment from pGEM/ΔE1E2 which encodes E1 to NS2, were inserted into pGEM1 (digested with *HindIII* and *EcoRI*) along with an oligonucleotide (over-lined in red) to bridge the *XhoI* and *NsiI* sites. This linked the N-terminal 65 residues of core (shown in blue) to 3 non-HCV residues (shown in black) followed by the signal peptide sequence for E1 (shown in purple). This construct was called pGEM/E1E2_{H77}.

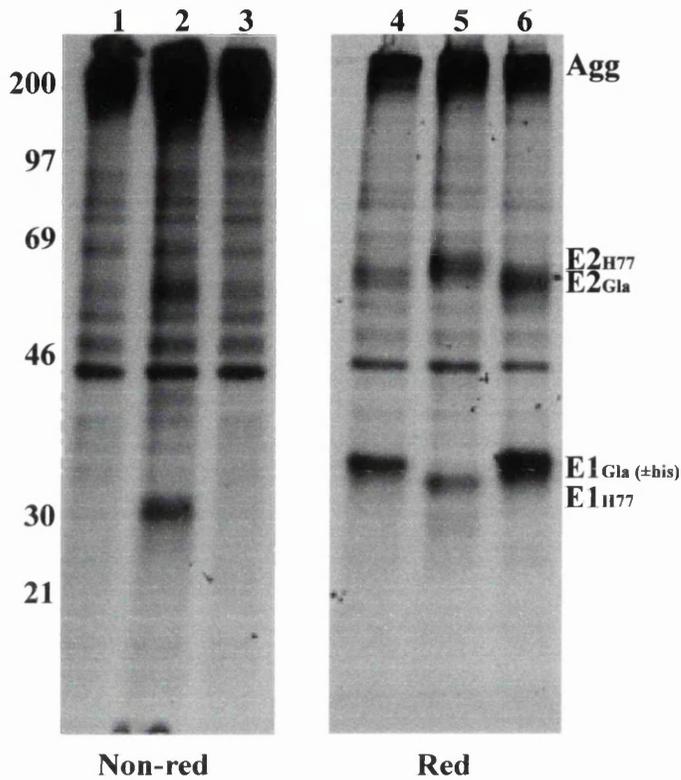


Fig 4.12. Comparison of the relative mobilities of E1 and E2 made by strains Glasgow and H77. Cells were electroporated with RNA and ³⁵S-labelled between 4-12 h post electroporation. Following lysis, the glycoproteins were immunoprecipitated from cell extracts using E2-specific MAb ALP98. Immunoprecipitates were electrophoresed on a 12% polyacrylamide gel under reducing (Red) and non-reducing conditions (Non-red). Samples were derived from cells electroporated with RNA from the following constructs: pSFV/CE1E2_{Gla} (lanes 1 and 4), pSFV/CE1E2_{H77} (lanes 2 and 5), pSFV/E1_{his}E2 (lanes 3 and 6). The positions of E1 and E2 from strain Glasgow (Gla) and H77 are indicated. The position of the aggregated material (Agg) is also shown.

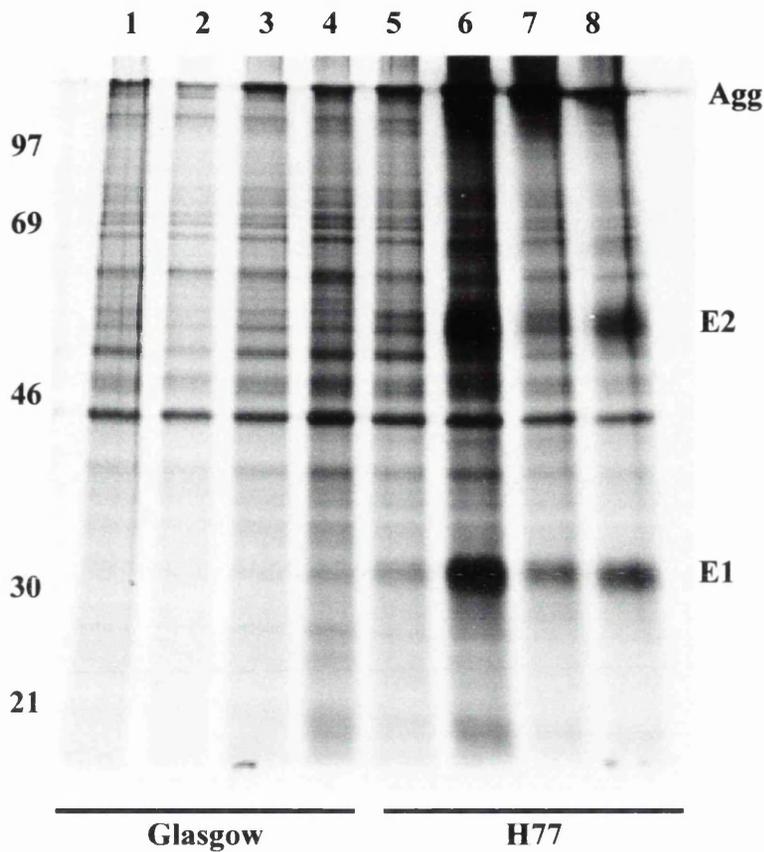


Fig 4.13. Immunoprecipitation analysis of strains Glasgow and H77 E1 and E2 proteins with E2-conformation specific antibodies. Cells electroporated with pSFV/CE1E2_{Gla} and pSFV/CE1E2_{H77} RNAs were radiolabelled from 4-12 h after electroporation. Immunoprecipitations on crude extracts were performed and precipitates were analysed under reducing conditions on a 10% polyacrylamide gel. Immunoprecipitates derived from cells expressing strain Glasgow E1 and E2 are shown in lanes 1-4 and those from strain H77 in lanes 5-8. Immunoprecipitations were performed with the following antibodies: H2 (lanes 1 and 5), H33 (lanes 2 and 6), H47 (lanes 3 and 7), and H53 (lanes 4 and 8). Positions of E1, E2 and aggregated material (Agg) are indicated.

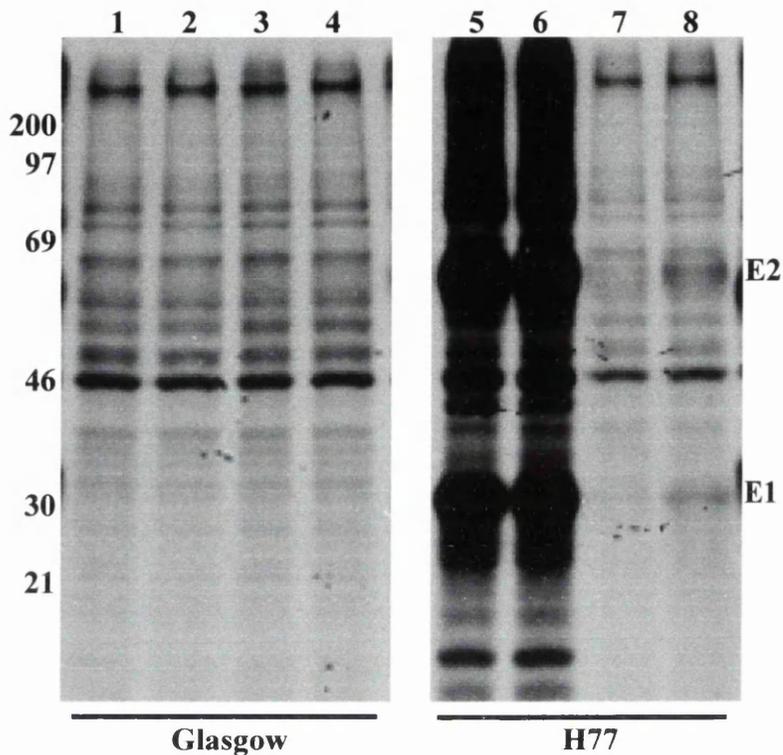


Fig 4.14. Immunoprecipitation analysis of strains Glasgow and H77 E1 and E2 proteins with human monoclonal antibodies. Cells electroporated with pSFV/CE1E2_{Gla} and pSFV/CE1E2_{H77} RNAs were ³⁵S-methionine labelled between 4-12 h after electroporation and lysed in immunoprecipitation buffer. Immunoprecipitations were performed on crude extracts using a panel of human monoclonal antibodies and are shown in the following order: 1A7 (lanes 1 and 5), 2H1 (lanes 2 and 6), 4F7 (lanes 3 and 7), and 4F1 (lanes 4 and 8). Samples were electrophoresed under reducing conditions on a 10% polyacrylamide gel. Lanes 1-4 show the precipitates obtained from extracts containing strain Glasgow proteins and lanes 5-8 show those obtained from extracts containing strain H77 proteins. Positions of E1 and E2 are indicated.

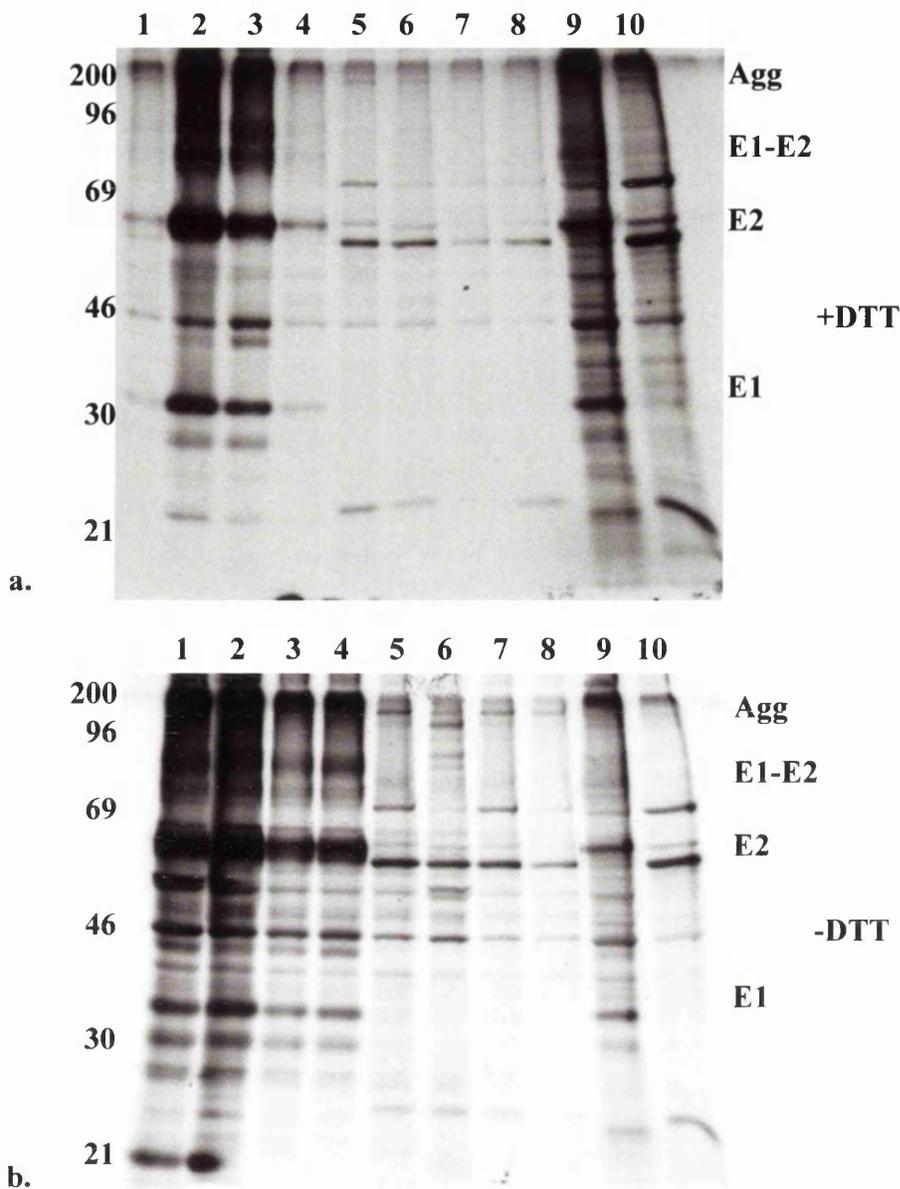


Fig 4.16. Analysis of *de novo* synthesised E1 and E2 from strain H77 in the presence of DTT. Cells were electroporated with pSFV/CE1E2_{H77} and pSFV1 RNA and incubated for 12 h. Following pulse labelling for 2 min with ³⁵S-methionine/cysteine and a chase period of 80 min in the presence (+) and absence (-) of DTT, cells were lysed in buffer containing 20mM NEM. Immunoprecipitations were performed on extracts using the following antibodies: E1-specific MAb, AP497 (lanes 1 and 5), E1-specific antiserum, R528 (lanes 2 and 6), E2-specific MAb, ALP98 (lanes 3 and 7), and E2-conformation-specific MAb, H53 (lanes 4 and 8). Immunoprecipitates were analysed on 10% polyacrylamide gels under reducing conditions. Lanes 1-4 show immunoprecipitates from extracts obtained from pSFV/CE1E2_{H77} RNA electroporated cells and lanes 5-8 from pSFV1 RNA. For control purposes, crude extracts from DTT-treated cells that were used for immunoprecipitation from pSFV/CE1E2_{H77} (lane 9) and pSFV1-electroporated cells (lane 10) are shown. Positions of E1, E2, E1-E2 precursor and aggregated material (Agg) are indicated.

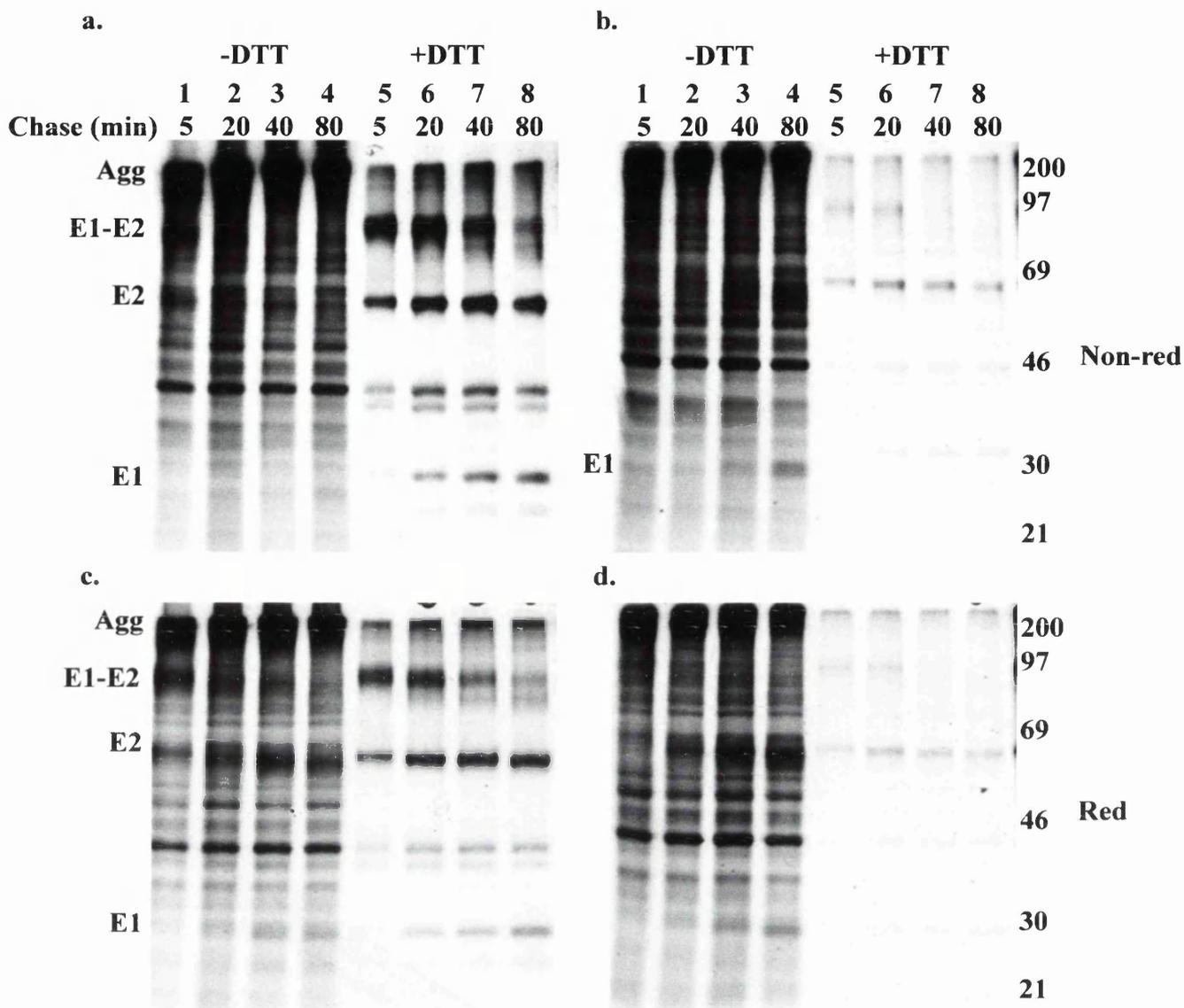


Fig 4.17. Complexes formed by strain H77 in the presence and absence of DTT. Cells were electroporated with RNA from pSFV/CE1E2_{H77} and incubated at 37°C for 12 h. Subsequently, cells were pulse labelled for 2 min and chased for the indicated times either in the presence (+DTT) or absence of DTT (-DTT). Immunoprecipitations on crude cell extracts were performed using E2-specific MAbs ALP98 and H53. Precipitates were divided and electrophoresed on 10% polyacrylamide gels under either non-reducing (Non-red, panels a and b) or reducing conditions (Red, panels c and d). Panels a and c show proteins precipitated by ALP98 while b and d are those precipitated by H53. For a and b, the positions of the oxidised and reduced forms of E1, E2, aggregated material (Agg) and E1-E2 precursor molecules are shown.

Chapter 5 – Characteristics of E1 and Their Relationship to Complex Formation with E2

5.1. Introduction

Despite a significant number of studies directed towards understanding the characteristics of HCV glycoprotein complexes, little is known about the properties of E1. Studies by both Ralston *et al.* (1993) and Fournillier-Jacob *et al.* (1996) indicated that the migratory properties and antigenicity of E1 were altered when it is expressed in the absence of E2. Moreover, the oxidation state of E1 also is dependent on E2 (Michalak *et al.*, 1997). Therefore, the influence of E2 on the behaviour of E1 was examined further in this chapter. From Chapter 4, covalent and non-covalent bonds are necessary for the stabilisation of aggregated complexes while intermolecular non-covalent interactions occur in native complexes. In an attempt to distinguish regions which may be important for formation of either type of complexes, internal deletions within E1 were generated and examined for their ability to associate with E2. Given the controversy over the role of disulphide bonds in native complex formation, and as a development of the DTT studies presented in Chapter 4, cysteine point mutants in E1 were examined for their interaction with E2 and effect on the properties of E1.

5.2.1. Construction of pSFV Recombinant Vectors Expressing Glasgow and H77 E1 in the Absence of E2

Plasmids expressing only the E1 glycoprotein from strains Glasgow and H77 were made by manipulating the parent constructs pGEM/E1_{his}E2 and pGEM/E1E2_{H77} (described in Section 3.2 and 4.6 respectively). The coding region of E2 in these clones was replaced with oligonucleotides, which encoded sequences from residues 382-386 and thus reconstituted the proteolytic cleavage site between E1 and E2 at amino acid residue 383. Following the sequences for residue 386, an artificial stop codon was also introduced (Fig 5.1 and Fig 5.2). pGEM/E1_{his}E2 was digested with *Sall* and *HindIII*, whereas pGEM/E1E2_{H77} was treated with *Sall* and *EcoRI*. The *HindIII* and *EcoRI* sites in these constructs reside in the pGEM1 vector, downstream from the E2 coding region. The *Sall* restriction enzyme site is conserved at the position encoding codon 381 in E1 for both

strains (Appendices 1 and 3). Insertion of oligonucleotides yielded plasmids that were named pGEM/E1_{Gla} and pGEM/E1_{H77} which encoded residues 169-386 corresponding to the E1 signal sequence, E1 protein and E2 signal sequence. The only differences between these plasmids were the presence of histidine tag in pGEM/E1_{Gla} (Section 3.2) and the presence of 68 residues upstream of the E1 signal sequence in pGEM/E1_{H77} (Section 4.6). *Bgl*III fragments produced from the two recombinant pGEM plasmids were then cloned into the pSFV1 vector to create expression constructs called pSFV/E1_{Gla} and pSFV/E1_{H77}.

5.2.2. Behaviour of E1 on Expression in the Absence of E2

As stated in the introduction to this chapter, there is evidence that E2 affects the oxidation state of E1 (Michalak *et al.*, 1997). To compare the behaviour of E1 in the presence and absence of E2, cells were transfected with RNA from strain H77 constructs (pSFV/E1E2_{H77} and pSFV/E1_{H77}) and strain Glasgow plasmids (pSFV/E1_{his}E2 and pSFV/E1_{Gla}). After radiolabelling, E1 was isolated from extracts by either immunoprecipitation (for strain H77) or Ni-NTA agarose purification (for strain Glasgow).

For strain Glasgow, three E1/E2-related products were isolated by Ni-NTA affinity chromatography from cells electroporated with pSFV/E1_{his}E2 RNA (Fig 5.3, lane 1). These corresponded to E1 (~32 kDa), E2 (diffuse material migrating between 60-65 kDa) and E1E2 polyprotein which had not been cleaved but was glycosylated (~97 kDa). As described in Chapter 3 (Section 3.7; Fig 3.16), E2 was isolated by virtue of complex formation with histidine-tagged E1. Ni-NTA chromatography on extracts from pSFV/E1_{Gla}-electroporated cells revealed not only a species corresponding to E1 (~32 kDa) but also a series of bands with apparent molecular weights of approximately 65, 100 and 130 kDa. Given that E2 is not produced by pSFV/E1_{Gla}, these higher molecular weight species were likely to be derived from expression of E1 alone. Indeed, from their sizes it is predicted that they correspond to dimeric (64 kDa), trimeric (96 kDa) and tetrameric (128 kDa) forms of E1. To confirm the E1-relatedness of these proteins, Western blot analysis was performed with the Penta-His antibody. Results revealed that the various high molecular weight species in the Ni-NTA purified material from pSFV/E1_{Gla}-electroporated cells were detected by the antibody (Fig 5.3, panel b, lane 2). The same species were recognised in the purified material expressed by pSFV/E1_{his}E2 but to a lesser extent (lane 1). These data indicate that E1 can homo-oligomerise but oligomeric products are formed less efficiently in the presence of E2.

Similar examination of strain H77 E1 expressed in the absence of E2 produced essentially identical results (Fig 5.4). In this case, immunoprecipitation assays were performed with cell lysates using two E1-specific MAbs AP21.010 (panel a) and AP497 which is conformation-specific (panel b; Appendix 5). The AP21.010-eluted samples, when analysed under reducing conditions, showed obvious differences between material precipitated from the extract containing only E1 expressed by pSFV/E1_{H77} (panel a, lane 1) to that containing E1 and E2 products made by pSFV/E1E2_{H77} (panel a, lane 2). Once again bands corresponding to E1 (~30 kDa), E2 (~69 kDa) and to the unprocessed E1-E2 polyprotein (~99 kDa) were precipitated from cells electroporated with pSFV/E1E2_{H77} RNA. However, high molecular weight bands of approximately 60, 90, 120 and 150 kDa were detected in crude extracts derived from pSFV/E1_{H77}-electroporated cells (panel a, lane 1). Immunoprecipitation data gained using AP497 revealed a similar pattern in precipitated material derived from cells expressing E1 in the absence and presence of E2 (compare panels a and b, lanes 1 and 2). In light of this, Western blot analysis was performed to confirm the identity of these additional bands using an anti-E1 specific antiserum (R528) (Fig 5.4, panel c). Data showed that the higher molecular weight bands as well as monomeric E1 were indeed recognised by the anti-E1 antiserum (panel c, lane 1). By comparison, such high molecular weight species were not detected to the same degree in samples derived from pSFV/E1E2_{H77}-electroporated cells expressing E1 in the presence of E2 (panel c, lane 2) although the E1-E2 uncleaved product was recognised. This shows that E1 expressed in the absence of E2, from both strains Glasgow and H77 behave in a comparable manner. From data presented in Chapter 4 (Section 4.7, Fig 4.12), strain Glasgow E1 protein migrated more slowly than strain H77 E1, while the converse was true for E2 protein from the two strains. As a result, the dimeric form of strain Glasgow E1 co-migrates with E2 while the E1 dimer for strain H77 has a lower apparent molecular weight than E2 for this strain. Overall, these data reveal the tendency of E1 to oligomerise in the absence of E2 and this behaviour is not strain-specific.

The second aspect regarding changes in the properties of E1 was discovered by analysis of the protein expressed in the absence or presence of E2 under non-reducing SDS-PAGE conditions. As described earlier, electrophoresis under these conditions does not interrupt intramolecular disulphide bonds within a protein, and thus the protein may migrate faster than its reduced counterpart since it can adopt a more compact structure. As strain

Glasgow proteins were not efficiently detectable under non-reducing conditions (Section 4.2, Fig 4.1), efforts to assess the contribution of E2 on E1 oxidation state were concentrated on the glycoproteins produced by strain H77. Cells were electroporated with RNA from strain H77 constructs pSFV/E1E2_{H77} and pSFV/E1_{H77} followed by radiolabelling for 4 to 12 h after electroporation. Proteins from these cells were isolated by immunoprecipitation using the E1-specific monoclonal antibody, AP21.010 and examined by non-reducing (Fig 5.5, panel a) and reducing (panel b) polyacrylamide gel electrophoresis. Non-aggregated E1 produced from both constructs was observed under non-reducing conditions (panel a). However, it was apparent that multiple monomeric forms of E1 were isolated from pSFV/E1_{H77}-electroporated cells (lane 1), whereas a single predominant E1 species was precipitated from cells expressing pSFV/E1E2_{H77} (lane 2). Moreover, none of the monomeric E1 species produced in the absence of E2 co-migrated with the major E1 species synthesised along with E2. Under reducing conditions however, these species did co-migrate (panel b, lanes 1 and 2). In panel a, samples were not reduced with DTT prior to electrophoresis and thus protein species would retain intramolecular disulphide bonds whereas in panel b, these bonds would be disrupted. Hence, the differential electrophoretic behaviour of the E1 species under non-reducing conditions is due to an altered oxidation state of the protein that is dependent on the presence of E2. It was concluded that interactions with E2 influence intramolecular disulphide bond formation in E1.

MAB AP21.010 isolated multiple monomeric E1 species (Figs 5.4 and 5.5, panel a, labelled –, =, + and ×), in addition to oligomeric forms from pSFV/E1_{H77}-electroporated cells. These multiple monomeric forms were less apparent on analysing pSFV/E1E2_{H77}-electroporated cells using this antibody. Further characterisation of these multiple monomeric forms of E1 was achieved by examination with endo H_f glycosidase treatment, which removes N-linked high-mannose oligosaccharides. Crude extracts, derived from cells electroporated with the pSFV/E1_{H77} RNA, were incubated with this enzyme and analysed by Western blot analysis using R528 antiserum. Data revealed that the multiple monomeric forms of E1 (Fig 5.6, panel a, lane 1), were reduced in size to a single species by endo H_f treatment (panel a, lane 2). This indicates that multiple monomeric bands correspond to differentially glycosylated forms of E1. Since these differentially glycosylated forms of E1 were more apparent in samples containing E1 in the absence of E2 co-expression (Fig 5.4, panel a, compare lanes 1 and 2), E2 may also influence the glycosylation pattern of E1. It is interesting to note that higher molecular bands recognised

by the R528 antiserum that represent the oligomeric forms of E1 were also reduced in size by deglycosylation with endo H_f (Fig 5.6, panel b). This reinforces the E1-relatedness of these bands.

From the above series of observations, three criteria were established by which E1 behaviour differed in the presence and absence of E2. These were oligomerisation, differential glycosylation, and oxidation state. Of these, differential glycosylation was occasionally less apparent since more than one monomeric species could be observed both in the presence and absence of E2 but there was a consistent difference in their relative abundance. These properties were used to further characterise E1 and to assess the interactions that occur between E1 and E2.

5.3.1. Construction of E1 Internal Deletion Mutants

In order to examine regions of E1 involved in association with E2, a panel of E1 internal deletion mutants was constructed in both Glasgow and H77 strains, using the parent plasmids pGEM/E1_{his}E2 and pGEM/E1E2_{H77}, respectively. The panel of constructs was designed to encode E1 mutants lacking peptide regions in a sequential manner, which were defined by suitable restriction enzyme sites in the nucleotide sequence. For strain Glasgow, a N-terminal E1 deletion mutant was created by replacing the coding region between amino acid residues 196 and 235 with the oligonucleotide indicated in Fig 5.7 (panel a). This oligonucleotide re-constructed the flanking restriction sites (*FspI* [at codon 195] and *XhoI* [at codon 236]) and the sequences coding for the histidine tag. The resultant plasmid was termed pGEM/E1_{Δ196-235}E2_{Gla}. Subsequent deletions in the glycoprotein were created by employing restriction enzyme sites *XhoI* (at codon 209), *Clai* (at codon 262), *MunI* (at codon 305), and *BamHI* (at codon 339) in the nucleotide sequence (see Appendices 1 and 3). These sites were bridged using oligonucleotides shown in Fig 5.7 (panels b, c, and d), which encoded a few amino acid residues on either side of the region to be deleted and in-frame nucleotide sequences to re-introduce the enzyme sites. These plasmids were named pGEM/E1_{Δ213-260}E2_{Gla}, pGEM/E1_{Δ266-303}E2_{Gla}, and pGEM/E1_{Δ309-338}E2_{Gla}. A final strain Glasgow E1 mutant resulted in a deletion between residues 341 and 366, where restriction sites *BamHI* and *HincII* (at codon 381) were connected by an oligonucleotide. As depicted in Fig 5.7 (panel e), this oligonucleotide re-constructed almost the entire coding sequence for the E1 transmembrane/E2 signal peptide domain (codons 366-377); codons 363-365 for this domain were omitted. This ensured

that the sequences required for E1 anchorage to lipid membranes, E2 translocation to the ER and the E1-E2 protease cleavage site were maintained. This construct was termed pGEM/E1 $_{\Delta 341-366}$ E2 $_{Gla}$. After the transfer of coding regions containing HCV sequences from the above E1 deletion mutant plasmids into the pSFV1 expression vector, the constructs were named pSFV/E1 $_{\Delta 196-235}$ E2 $_{Gla}$, pSFV/E1 $_{\Delta 213-260}$ E2 $_{Gla}$, pSFV/E1 $_{\Delta 266-303}$ E2 $_{Gla}$, pSFV/E1 $_{\Delta 309-338}$ E2 $_{Gla}$, and pSFV/E1 $_{\Delta 341-366}$ E2 $_{Gla}$, respectively.

For strain H77 E1, a histidine tag was not inserted at the N-terminus, and therefore, an alternative oligonucleotide (shown in Fig 5.8) was employed compared with that used to make pGEM/E1 $_{\Delta 196-235}$ E2 $_{Gla}$. A slightly smaller region was deleted since the oligonucleotide coded for additional in-frame sequences. Thereafter the construct was called pGEM/E1 $_{\Delta 199-233}$ E2 $_{H77}$. Internal deletions between residues 213-260, 266-303 and 309-338 were produced using strategies and oligonucleotides identical to those applied to construct similar deletion mutants in strain Glasgow (illustrated in Fig 5.7, panels b, c, and d). These constructs were termed pGEM/E1 $_{\Delta 213-260}$ E2 $_{H77}$, pGEM/E1 $_{\Delta 266-303}$ E2 $_{H77}$, and pGEM/E1 $_{\Delta 309-338}$ E2 $_{H77}$. For the last mutant in strain H77, called pGEM/E1 $_{\Delta 341-366}$ E2 $_{H77}$ (Fig 5.9), a different oligonucleotide from that used for strain Glasgow was employed to prevent incorporation of a naturally occurring amino acid change (aa 372) between the two strains. The introduction of strain H77 E1 internal deletion mutant sequences into pSFV1 resulted in the following plasmids: pSFV/E1 $_{\Delta 199-233}$ E2 $_{H77}$, pSFV/E1 $_{\Delta 213-260}$ E2 $_{H77}$, pSFV/E1 $_{\Delta 266-303}$ E2 $_{H77}$, pSFV/E1 $_{\Delta 309-338}$ E2 $_{H77}$, and pSFV/E1 $_{\Delta 341-366}$ E2 $_{H77}$. For simplicity, a schematic representing the deleted regions in E1 is shown in Fig 5.10. Due to the cloning strategy, E1 mutants pSFV/E1 $_{\Delta 196-235}$ E2 $_{Gla}$ ($\Delta 199-233$ in the case of H77) and pSFV/E1 $_{\Delta 213-260}$ E2 have overlapping deleted regions between amino acid residues 213-235 and 213-233 for strains Glasgow and H77 respectively. For all the above constructs, nucleotide sequences in the sense and anti-sense directions were determined to ensure the inserted oligonucleotides resulted in in-frame deletions.

5.3.2. Examination of the Properties of E1 Deletion Mutants

Initially, characterisation of the E1 mutants was conducted to establish their behaviour on reducing SDS-PAGE and their glycosylation patterns. Cells were electroporated with RNA from the constructs containing wild type and mutant forms of E1 from strains Glasgow and H77. A portion of crude extract obtained from these cells was separated

under reducing SDS-PAGE conditions and examined by Western blot analysis, probing for E1 using R528 (Fig 5.11). One of the features observed was the presence of multiple monomeric forms generated by strain H77 mutant proteins (Fig 5.11, panel a, lanes 4, 6, 8, and 12), which were detected also for wild type E1 expressed in the presence and absence of E2 (lanes 2 and 13). This was particularly evident when using R528 antiserum. The only mutant for which multiple monomeric forms were less apparent was pSFV/E1 $_{\Delta 309-338}$ E2 $_{H77}$ (lane 10). In addition, multiple species were discernible in samples obtained from three strain Glasgow E1 mutants, E1 $_{\Delta 196-235}$ Gla (Fig 5.11, panel a, lane 3), E1 $_{\Delta 341-366}$ Gla (lane 11) and to a limited extent E1 $_{\Delta 213-260}$ Gla (lane 5). Multiple monomeric forms were not detected for wild type strain Glasgow E1 and two mutants E1 $_{\Delta 266-303}$ Gla and E1 $_{\Delta 309-338}$ Gla (lanes 1, 7 and 9). Previously, multiple monomeric species were demonstrated to correspond to differentially glycosylated forms of monomeric E1 by endo H_f analysis (Section 5.3, Fig 5.6). To establish the nature of the multiple forms, crude extracts were subjected to deglycosylation by endo H_f digestion and again examined by Western blot analysis using the anti-E1 antiserum, R528. The counterparts in strains Glasgow and H77 were analysed in parallel (Fig 5.11, panel b) and revealed that deglycosylated mutants were represented by single monomeric species, thereby confirming that the multiple monomeric forms detected in the untreated samples indeed represented differentially glycosylated species.

Endo H_f digestion also confirmed that the different mobilities of the processed proteins made by two corresponding mutants, E1 $_{\Delta 196-235}$ E2 $_{Gla}$ and E1 $_{\Delta 199-233}$ E2 $_{H77}$ (compare panels a and b, lanes 3 and 4), were a consequence of differential glycosylation. In E1 $_{\Delta 196-235}$ E2 $_{Gla}$, three predicted glycosylation sites were removed as compared to one predicted site in E1 $_{\Delta 199-233}$ E2 $_{H77}$. Accordingly, the glycosylated mutant protein from strain Glasgow has greater mobility than its H77 counterpart. Upon endo H_f treatment however, the deglycosylated forms of the proteins made by these mutants have similar mobilities, and agree with the predicted molecular weights of the nascent polypeptides (Table 5.1). E1 mutants in strain Glasgow and H77 which lack a single glycosylation site [E1 $_{\Delta 213-260}$ E2 $_{(H77/Gla)}$ (lanes 5 and 6)] gave proteins with mobilities that were consistent with the loss of both amino acid residues as well as an oligosaccharide moiety. However, two mutants in each strain, E1 $_{\Delta 266-303}$ E2 $_{(H77/Gla)}$ (lanes 7 and 8) and E1 $_{\Delta 341-366}$ E2 $_{(H77/Gla)}$ (lanes 11 and 12) produced proteins with higher apparent molecular weights than would be predicted (Table 5.1). For all four mutant proteins, the highest molecular weight glycosylated form

in each case co-migrated with or close to the wild type form of the E1 proteins. A similar picture emerged from examination of the endo H_f-treated samples. Thus, the discrepancy in the apparent molecular weights of proteins made by these mutants is not solely a consequence of glycosylation. Another feature observed from analysis of the mutants was the appearance of novel higher molecular weight proteins (indicated by *) made by E1_{Δ266-303}E2_(H77/G1a) (50 kDa) and E1_{Δ309-338}E2_{H77} (45 kDa). For the latter mutant, a novel species was not observed for the corresponding strain Glasgow mutant, E1_{Δ303-338}E2_{G1a}; this may be a consequence of the lower levels of E1 protein that were detected for this particular mutant. This feature is discussed further in section 5.3.4.

Finally, there was marked increase in the amount of uncleaved E1-E2 polyprotein detected for mutant E1_{Δ341-366}E2 for both strains (Fig 5.11, panels a and b, lanes 11 and 12). This suggests that, while the cleavage between E1 and E2 in the polyprotein is predicted to occur at residue 384 (Grakoui *et al.*, 1993b; Ralston *et al.*, 1993; Selby *et al.*, 1994), amino acid sequences upstream of residue 366 in E1_{Δ341-366} mutants may be necessary to achieve efficient processing of the polyprotein.

5.3.3. Analysis of Complex Formation between E2 and E1 Deletion Mutants

The purpose for creating E1 internal deletion mutants was to enable the identification of regions in E1 which may associate with E2 in heteromeric complexes. Removal of regions in E1 in a systematic fashion was anticipated to abolish interaction of deletion mutant E1 proteins with co-expressed wild type E2. Cells were electroporated with RNA from strain H77 constructs expressing wild type proteins and mutant E1 species along with E2. After ³⁵S-labelling for 12 h, immunoprecipitations were performed on crude extracts obtained from these cells with anti-E1 antiserum, R528, and precipitates analysed by both non-reducing and reducing polyacrylamide gel electrophoresis (Fig 5.12).

Under reducing conditions, it was apparent that all the E1 internal deletion mutants continued to co-precipitate E2, suggesting that interaction between mutant E1 and E2 was not abolished by deletion of any region of E1 (Fig 5.12, panel b). However, the glycoproteins detected under these conditions may represent disrupted aggregated material. To determine whether E1 mutants could give rise to native complexes, the

immunoprecipitates were analysed by non-reducing gel electrophoresis. Results showed that some non-aggregated monomeric forms of mutant E1 may be precipitated from samples derived from constructs pSFV/E1 $_{\Delta 199-233}$ E2 $_{H77}$, pSFV/E1 $_{\Delta 213-260}$ E2 $_{H77}$, and pSFV/E1 $_{\Delta 266-303}$ E2 $_{H77}$ (panel a, lanes 2, 3, and 4), but not apparently from those produced by constructs pSFV/E1 $_{\Delta 309-338}$ E2 $_{H77}$ and pSFV/E1 $_{\Delta 341-366}$ E2 $_{H77}$ (lanes 5 and 6). Further, the co-elution of E2 with mutant E1 proteins was not readily seen in the radiolabelled material, implying that the non-aggregated monomeric E1 detected may correspond to E1 molecules which are not associated with E2. For confirmation purposes, Western blot analysis was performed to detect E2 on duplicate precipitates separated under the same electrophoresis conditions. From these data E2, migrating as a diffuse band, co-precipitated with wild type E1 (Fig 5.13, lane 1). However, significantly reduced amounts of E2 eluted with mutants E1 $_{\Delta 199-233}$, E1 $_{\Delta 213-260}$, and E1 $_{\Delta 266-303}$ (lanes 2, 3, and 4, respectively). The failure to detect E2 in samples from pSFV/E1 $_{\Delta 309-338}$ E2 $_{H77}$ - and pSFV/E1 $_{\Delta 341-366}$ E2 $_{H77}$ -electroporated cells (lanes 5 and 6), reflects the fact that no monomeric forms were detected for these mutant E1 proteins under non-reducing conditions (Fig 5.12, panel a, lanes 5 and 6).

A reciprocal experiment was conducted to verify the above findings, in which immunoprecipitations were performed using the E2-specific MAb, ALP98. Analysis of the samples under reducing conditions indicated that E2 precipitated all the E1 deletion mutants (Fig 5.14, panel b). By contrast, from non-reducing electrophoresis data, a monomeric species was readily detected only for wild type E1 (panel a, lanes 1-6). It was concluded that the E1 mutants from strain H77 do not efficiently form native complexes with E2 and that those complexes detected represented aggregates of E1 and E2.

Examination of crude extracts of cells electroporated with strain Glasgow E1 mutant constructs, using R528 antiserum, revealed a similar pattern under reducing conditions, whereby all the mutant E1 proteins were found to associate with E2 (Fig 5.15, panel b). For this strain, some non-aggregated E1 could be detected from pSFV/E1 $_{\Delta 196-235}$ E2 $_{Gla}$, pSFV/E1 $_{\Delta 213-260}$ E2 $_{Gla}$, and pSFV/E1 $_{\Delta 266-303}$ E2 $_{Gla}$ constructs (panel a, lanes 2, 3, and 4) in agreement with strain H77 E1 mutants, but again co-precipitation of non-covalently linked E2 was not detected. Together, these data reveal that manipulation of internal regions of E1 increase the tendency for aggregate formation, and in turn, that the native complex is highly sensitive to changes in E1.

5.3.4. Further Characterisation of Deletion Mutants E1_{Δ266-303} and E1_{Δ309-338}

With the exception of mutants E1_{Δ266-303} and E1_{Δ341-366} from both strains all the remaining E1 deletion mutants migrated according to their predicted molecular weights by SDS-PAGE (see Section 5.3.2). Due to time constraints, only the behaviour of E1_{Δ266-303} mutant was further examined. Crude extracts derived from cells electroporated with constructs pSFV/E1E2_{H77}, pSFV/E1_{Δ266-303}E2_{H77}, and pSFV/E1_{Δ309-338}E2_{H77} were analysed in parallel by Western blot analysis using R528. Compared to wild type E1, mutant E1_{Δ266-303} lacks 38 amino acid residues, but removal of these sequences should not affect the N-glycosylation sites. Therefore, the deletion should have resulted in a reduction of approximately 4.0 kDa (as predicted by ExPASy-PeptideMass program) as compared to wild type protein. However, wild type E1 and E1_{Δ266-303} essentially co-migrated (Fig 5.16, lanes 1 and 2) while mutant E1_{Δ303-338} protein migrated more closely to its apparent molecular weight (lane 3). To assess whether glycosylation played a role in the unpredicted increase in molecular weight of this mutant, crude extracts were treated with endo H_f and PNGase F endoglycosidases. Western blot analysis on these samples revealed that deglycosylated E1_{Δ266-303}, following treatment by either enzyme, continued to co-migrate with wild type E1 (Fig 5.16, lanes 4, 5, 7, and 8). *In vitro* transcription and translation of wild type E1 and mutant E1 proteins, where no post-translational modifications occur, suggested that E1_{Δ266-303} mutant protein did migrate according to its predicted molecular weight (Fig 5.17, lane 3). In this case, pGEM1 constructs encoding wild type E1 and mutant E1 proteins were linearised at codon position 381 using *Sall* restriction enzyme. This codon is situated two residues upstream of the authentic cleavage site between E1 and E2 in the polyprotein. The T7 promoter was utilised in the pGEM1 vector to produce run-off transcripts coding for wild type E1 and mutant E1 proteins alone. The coupled transcription and translation reaction was performed in the presence of ³⁵S-methionine and analysed by SDS-PAGE. The data from this experiment and those in Fig 5.16 indicated that the unusual migratory characteristics of mutant E1_{Δ266-303} were not a consequence of alterations in glycosylation in tissue culture cells.

Further examination of the Western blot and immunoprecipitation data presented in Figs 5.11. (panel a and b, lanes 8 and 10) and 5.12 (panel b, lane 4) revealed the presence of higher molecular weight species made by strain H77 E1 $_{\Delta 266-303}$ and E1 $_{\Delta 309-338}$ constructs (indicated by *), in addition to monomeric forms. These high molecular weight forms (~50 kDa and ~46 kDa) may represent dimers for each of these mutant E1 proteins based on their predicted molecular weights. The predicted molecular weights of fully glycosylated monomeric E1 $_{\Delta 266-303}$ and E1 $_{\Delta 309-338}$ of strain H77 were ~27 kDa and ~26 kDa respectively, and thus their corresponding dimers could be expected to be 54 kDa and 52 kDa. To examine this hypothesis these mutant E1 proteins were expressed in the absence of E2; as shown earlier, under such circumstances, increased amounts of wild type E1 oligomers are detected (see Fig 5.4, panel a, compare lanes 1 and 2).

The constructs expressing these two mutant E1 proteins alone were called pSFV/E1 $_{\Delta 266-303H77}$ and pSFV/E1 $_{\Delta 309-338H77}$ (described in Fig 5.18). They were constructed by obtaining *NsiI-BamHI* fragments from their parent plasmids pSFV/E1 $_{\Delta 266-303H77}E2$ and pSFV/E1 $_{\Delta 309-338H77}E2$, which were inserted into the pSFV/E1 $_{H77}$ construct (see Section 5.2.1 and Fig 5.2) employing the same sites. Subsequently, cells were electroporated with pSFV/E1 $_{H77}$, pSFV/E1 $_{\Delta 266-303H77}$, and pSFV/E1 $_{\Delta 309-338H77}$ RNAs. For control purposes RNAs from pSFV/E1 $E2_{H77}$, pSFV/E1 $_{\Delta 266-303}E2_{H77}$, and pSFV/E1 $_{\Delta 309-338}E2_{H77}$ were also electroporated into cells. Immunoprecipitation assays using rabbit anti-sera R528 performed on crude extracts from these cells were analysed under reducing conditions (Fig 5.19). Data revealed that the ~50 kDa band detected for the pSFV/E1 $_{\Delta 266-303}E2_{H77}$ construct (lane 2) was also detected when E1 $_{\Delta 266-303}$ was expressed in the absence of E2 (lane 5). This suggested that the ~50 kDa species did indeed represent a dimer of the mutant protein. In the case of mutant E1 $_{\Delta 309-338}$, a ~46 kDa band, detected previously only by Western blot analysis (Fig 5.11, panel a, lane 10), was apparent upon expression of this mutant in the absence of E2; this species co-migrated with a 45 kDa background protein (Fig 5.19, lane 6). Based on its apparent molecular weight, it is likely to be a dimer of E1 $_{\Delta 309-338}$ protein. The ability to detect such forms of E1, even in the presence of E2, again highlights the sensitivity of E1 behaviour to removal of certain sequences.

5.4.1. Construction of Cysteine Substitution Mutants

The detection of two forms of E1E2 complexes, the aggregate as well as the native complex, suggests that disulphide bonds play a role in the formation of the covalently aggregated material. However, sulphide groups on cysteine residues in the glycoproteins might also be involved in intramolecular disulphide bond formation in addition to the intermolecular links. Intramolecular disulphide bonds could indirectly influence E1E2 complex formation by affecting the tertiary structure of the proteins. Therefore, in addition to internal deletion mutagenesis, another strategy followed to study the properties of E1 important for E1E2 complex formation involved the creation of cysteine substitution mutant E1 proteins. There are eight cysteine residues in E1 (Fig 5.20, panel a), each of which is conserved throughout the HCV strain 1a genotypes.

The cysteine residues in strain H77 E1 were substituted by incorporation of point mutations at the cysteine-encoding codons via the introduction of oligonucleotides at appropriate restriction enzyme sites; this is illustrated schematically in Fig 5.20 (panel b-e). All the cysteine residues were substituted for serine residues, with the exception of the cysteine at residue 229, which was replaced by an alanine residue to avoid the creation of a novel glycosylation site (Fig 5.20, panel b). With the use of the unique *NsiI* site engineered at the N-terminus of the E1 signal sequence coding region and the *BamHI* site at codon 339 in E1, it was possible to construct all of these substitution mutants directly in the pSFV1 vector. The vector and E2 sequences were derived from pSFV/E1E2_{H77} linearised at *NsiI* and *BamHI*, while the fragments encoding E1 were obtained from pGEM/E1E2_{H77} using suitable restriction sites. The fragments were bridged using oligonucleotides carrying point mutations at cysteine codons. The construct containing the substitution at position 207 was named pSFV/E1_{ΔCys207}E2 (panel b), whereas the remainder of the intermediate cysteine substitution mutants, where up to four cysteine residues were substituted at a time, were named pSFV/E1_{ΔCys226,229,238}E2 (panel c), pSFV/E1_{ΔCys272,281,304}E2 (panel d) and pSFV/E1_{ΔCys272,281,304,306}E2 (panel e). The pSFV/E1_{ΔCys272,281,304,306}E2 mutant was constructed using pSFV/E1_{ΔCys272,281,304}E2 as the parent plasmid vector.

A mutant E1 protein in which all eight cysteine residues were substituted is shown in Fig 5.21. Coding fragments derived from pSFV/E1_{ΔCys207}E2 and pSFV/E1_{ΔCys272,281,304,306}E2

plasmids were ligated with the oligonucleotides originally used to make pSFV/E1 $_{\Delta\text{Cys}226,229,238}$ E2 (Fig 5.21). An *EagI-XbaI* fragment was obtained from pSFV/E1 $_{\Delta\text{Cys}207}$ E2 while pSFV/E1 $_{\Delta\text{Cys}272,281,304,306}$ E2 provided a *BstEII-XbaI* fragment, in which the common *XbaI* site is situated in the pSFV1 vector (Fig 3.3, panel b) and *EagI* and *BstEII* sites are located within the coding sequence of E1 at codon positions 216 and 241, respectively. This construct was called pSFV/E1 $_{\Delta\text{-cys}}$ E2, and the construct expressing E1 alone was called pSFV/E1 $_{\Delta\text{-cys}}$. The latter plasmid was made by excising a *NsiI-BamHI* fragment from pSFV/E1 $_{\Delta\text{-cys}}$ E2 and introducing it into the vector generated from the pSFV/E1 $_{\text{H77}}$ construct (see Fig 5.2) using the same sites; a similar strategy is described in Fig 5.18.

5.4.2. Analysis of Cysteine Substitution Mutants

From the studies described in section 5.2.2, it was established that E1, when expressed in the absence of E2, tends to oligomerise. To address whether intramolecular disulphide bonds alone are involved in oligomerisation of E1, the properties of the E1 mutant in which all of the cysteine residues had been substituted were examined in the presence and absence of E2. Cells were electroporated with pSFV/E1 $_{\Delta\text{-cys}}$ E2 and pSFV/E1 $_{\Delta\text{-cys}}$; for comparative purposes, pSFV/E1E2 $_{\text{H77}}$ and pSFV/E1 $_{\text{H77}}$ were also used.

Immunoprecipitations were performed on crude extracts from these cells with the anti-E1 antiserum, R528, and the precipitates analysed under reducing and non-reducing electrophoretic conditions (Fig 5.22). Data from the non-reducing gel revealed that a markedly reduced level of non-aggregated E2 co-precipitated with E1 $_{\Delta\text{-cys}}$ (Fig 5.22, panel a, lane 2) as compared with the amount precipitated with wild type E1 (panel a, lane 1). By contrast, from the reducing gel, it was apparent that approximately equivalent amounts of E2 were precipitated with both forms of E1 (panel b, compare lanes 1 and 2). This indicates that removal of all of the cysteine residues in E1 does not abolish complex formation with E2. However, the lack of monomeric E2 observed under non-reducing conditions indicates that the complexes formed by E1 lacking cysteine residues are aggregated and non-native complexes.

Another feature apparent under non-reducing electrophoresis conditions was the appearance of multiple monomeric forms of E1 $_{\Delta\text{-cys}}$ expressed in the presence and absence of E2 (Fig 5.22, panel a, lanes 2 and 3). Given that E1 $_{\Delta\text{-cys}}$ protein cannot form either inter-

or intramolecular covalent bonds, reducing agents such as DTT or β -mercaptoethanol would have no effect on its migration properties. Previous data suggested that these monomeric forms were found in increased amounts when E1 was expressed alone (refer to Section 5.2.2) and, from endo H_f deglycosylation analysis, represented differentially glycosylated species. To confirm that this was the case, crude extracts derived from cells electroporated with the four constructs stated above were subjected to deglycosylation by endo H_f. Western blot analysis on these samples did indeed show that all the multiple species were represented by a single non-glycosylated form (Fig 5.23, lanes 1-4).

Non-reducing electrophoresis data from Fig 5.22 showed also that wild type E1 made in the presence of E2 (panel a, lane 1) had a greater mobility than the highest molecular weight species of E1 made by pSFV/E1_{H77}, pSFV/E1 _{Δ -cys}E2 and pSFV/E1 _{Δ -cys} (panel a, lanes 2-4). Results presented earlier in this chapter (Section 5.2.2) had indicated that E2 influenced the oxidation state of E1, whereby the oxidised form of E1 was not observed when synthesised in the absence of E2. In this instance, E1 cysteine substitution mutants would not be expected to reach an oxidised state due to the absence of intramolecular disulphide links. It is therefore significant that the species made by E1 _{Δ -cys} co-migrated with E1 protein produced in the absence of E2 but not with E1 that was made in the presence of E2. This provides conclusive evidence that wild type E1 is indeed unable to reach an appropriate oxidised state in the absence of E2.

Finally, examination of wild type E1 and E1 _{Δ -cys}, when expressed in the absence of E2, under reducing SDS-PAGE conditions (Fig 5.22, panel b, lanes 3 and 4), revealed that E1 _{Δ -cys} is able to oligomerise, but to a lesser extent than wild-type protein. As would be predicted, these oligomers can be observed also under non-reducing conditions for E1 _{Δ -cys} protein (panel a, lane 3). This suggests that oligomer formation is not entirely dependent on disulphide bond formation and can be stabilised by non-covalent association. This observation is also reproduced when immunoprecipitations were performed with E1-specific antibody AP21.010 (Fig 5.24, panels a and b, lanes 7 and 8). There is a small but noticeable difference in mobilities observed between oligomeric forms of wild type E1 and E1 _{Δ -cys}. ExpASY-Peptide Mass program did not predict a significant difference in the molecular weight of wild type E1 and E1 _{Δ -cys} (20.9 and 20.8 kDa respectively), due to the cysteine substitutions. The possible basis for the altered mobilities of the oligomers is discussed in Section 5.5.

5.4.3. Analysis of Partial Cysteine Substitution Mutants

To further study the effect of cysteine residue substitutions in E1 on native complex formation, cells were electroporated with RNA from constructs encoding E1 mutants carrying up to 4 cysteine residue substitutions (pSFV/E1 $_{\Delta\text{Cys}207}$ E2, pSFV/E1 $_{\Delta\text{Cys}226,229,238}$ E2, pSFV/E1 $_{\Delta\text{Cys}272,281,304}$ E2, and pSFV/E1 $_{\Delta\text{Cys}272,281,304,306}$ E2), as well as from constructs encoding E1 proteins with all the cysteine amino acids replaced (pSFV/E1 $_{\Delta\text{-cys}}$ E2 and pSFV/E1 $_{\Delta\text{-cys}}$). For comparative purposes cells were also electroporated with pSFV/E1E2 $_{\text{H77}}$, pSFV/E1 $_{\text{H77}}$ and pSFV1. Immunoprecipitations were performed on the crude extracts obtained from these cells using the E1-specific MAbs, AP21.010 and AP497, in addition to the E2-specific MAb, ALP98; the precipitates were examined under both reducing and non-reducing conditions. The non-reducing data for the E1-specific antibody AP21.010 revealed that non-aggregated E1 was recognised for all four intermediate mutants (Fig 5.24, panel a, lanes 2-5), although the relative amount of non-aggregated E1 $_{\Delta\text{Cys}207}$ was somewhat reduced as compared to the other E1 variants. Co-precipitation of non-covalently linked E2 was not as obvious as in the case of wild type E1 (panel a, compare lane 1 with lanes 2-5). However, under reducing conditions, where both the native and aggregated E1E2 complexes would be represented, relatively equal levels of E2 were co-precipitated by wild type E1 (panel b, lane 1) and by the cysteine substitution mutants E1 $_{\Delta\text{Cys}207}$, E1 $_{\Delta\text{Cys}226,229,238}$ and E1 $_{\Delta\text{-cys}}$ E2 (panel b, lanes 2, 3 and 6 respectively). There was a slight reduction in the amount of E2 co-immunoprecipitated by AP21.010 from extracts containing E1 $_{\Delta\text{Cys}272,281,304}$ and E1 $_{\Delta\text{Cys}272,281,304,306}$ (panel b, lanes 4 and 5). Nonetheless, these data indicated that E1 cysteine substitution mutants predominately formed aggregated but not native complexes with E2.

Closer examination of the relative mobilities of E1 precipitated by AP21.010 under non-reducing conditions revealed that the removal of a single cysteine residue at position 207 abolished the formation of oxidised E1 despite co-expression and complex formation with E2 (Fig 5.24, panel a, compare lanes 1 and 2). In addition, oxidised E1 was not detected for any of the other cysteine substitution mutants (Fig 5.24, panel a, lanes 3-7). This demonstrates the sensitivity of E1 folding and oxidation to alterations at cysteine residues in the protein.

As mentioned previously (Section 4.7), the ability of AP497 to recognise E1 is sensitive to reduction of the protein by DTT. The antibody did not detect E1_{ΔCys226,229,238} protein under either reducing or non-reducing conditions (Fig 5.25, panels a and b, lanes 3). Thus neither the native nor the aggregated complexes formed by this mutant were recognised (Fig 5.25, panels a and b, lanes 3). As a consequence AP497 also did not precipitate E1 protein in which all of the cysteine residues were substituted (panels a and b, lanes 6 and 7). However, immunoprecipitation analysis with this antibody did confirm that non-aggregated forms of the remaining E1 cysteine substitution mutants were detected readily under non-reducing conditions (Fig 5.25, panel a, lanes 4 and 5) with the exception of E1_{ΔCys207} protein which was significantly reduced in abundance (Fig 5.25, panel a, lane 2). As found with AP21.010, the level of this mutant in the reducing gel was increased (panel b, lane 2) indicating that most of this protein is linked to other E1 molecules or E2 by covalent intermolecular bonds. Once again only the reduced form of the E1 cysteine mutant proteins was detected under non-reducing conditions (panel a, lanes 4 and 5).

Finally, analysis of material precipitated by the E2-specific MAb, ALP98, under non-reducing conditions showed that reduced amounts of non-aggregated E2 expressed in the presence of E1 cysteine substitution mutants were detected as compared to the wild type proteins (Fig 5.26, panel a, compare lane 1 with lanes 2-6). However, under reducing conditions, equivalent amounts of E2 were clearly observed with wild type as well as cysteine mutants in E1 (panel b, lanes 1-6). This result suggests that alterations at individual amino acids in E1 may significantly affect the formation of non-aggregated E2. Examination of ALP98-precipitated material also showed that non-aggregated E1 could be detected for the cysteine mutants under non-reducing conditions (panel a, lanes 3-6) with the exception of E1_{ΔCys207} protein (lane 2). The protein was present under reducing conditions indicating that it did associate with E2. Thus, the reduced amount of the non-aggregated form of this protein was consistently found with three separate antibodies.

5.3.5. Discussion

Results presented in this and the previous chapter indicate that E1 behaves unpredictably on examination in reducing polyacrylamide gels. Firstly, a difference in the mobilities of strains Glasgow and H77 was observed in Chapter 4, despite the prediction that both nascent polypeptides are of identical length and very similar molecular mass. In this

chapter, deglycosylation of E1 from both strains with endo H_f did not yield nascent proteins with similar mobilities, thus indicating that N-linked glycosylation is not responsible for the observed difference. The difference in apparent molecular weights for E1 between the strains is approximately 1-2 kDa. There are 9 amino acid changes between the strains and it is possible that this has introduced an additional post-translational processing site into E1 from strain Glasgow. Apart from N-glycosylation, the forms of covalent modifications that take place in the cell include acetylation, O-linked glycosylation, hydroxylation, methylation, nucleotidylation, phosphorylation, ADP-ribosylation and ubiquitination. Further comparative studies would be necessary to examine whether additional modifications are present on strain Glasgow E1. This could be attempted through analysis of peptides derived from peptidase digestion of E1 and accurate molecular weight derivation by mass spectrometry.

A second possible explanation for the apparent molecular weight difference in E1 between the strains could be aberrant mobility on polyacrylamide gels. Such a phenomenon was observed for two E1 mutant proteins (E1_{Δ266-303} and E1_{Δ341-366}) and this was found for mutants in both strains. The apparent molecular weights for these proteins were greater than predicted and was not a result of alterations to the predicted patterns of N-glycosylation. Interestingly, both of the regions removed are contained within hydrophobic segments of E1. Aberrant mobility for E1 from strain H occurs also on removal of the E2 signal peptide sequence from the C-terminal end of E1 (residues 371-383; Fournillier-Jacob *et al.*, 1996). These authors concluded that this hydrophobic sequence was responsible for influencing the migratory characteristics of E1 such that their removal decreased mobility of the protein on gels. A similar phenomenon may account for the behaviour of E1_{Δ266-303} and E1_{Δ341-366} mutant proteins. It is possible that the hydrophobic segments may retain or adopt a local compact configuration following denaturation which is sufficient to increase the mobility of E1 on gels. Returning to the apparent difference in molecular weight of E1 between the two strains, no obvious difference was observed in the hydrophobicity plots for this protein from both strains.

Under non-reducing conditions E1 synthesised on its own migrated differently from E1 synthesised in the presence of E2; E1 expressed along with E2 in *cis* attained an oxidised state while the majority of E1 expressed in the absence of E2 remained in a reduced state. This finding confirms observations made by Michalak *et al.* (1997) which indicated that the behaviour of E1 is altered in the absence of E2. From their analysis, an oxidised form

of vaccinia virus-expressed strain H E1 protein was detected after a 2 h chase period from a C-E1-E2₁₋₇₄₆ construct but no such species was observed from a construct expressing E1 residues from 1-383. In addition to this observation, two more criteria have been established in this study to differentiate between E1 synthesised in the presence and absence of E2 which are based on the behaviour of E1-specific species under reducing conditions. In the absence of E2, E1 tends to form glycosylated oligomers and E1 monomers exist as a series of differentially glycosylated species.

The existence of oligomeric forms of E1 has not been reported previously. Oligomers may form through exposed regions of E1, not normally available when associated with E2, which interact with neighbouring E1 molecules in an identical state. The interactions between E1-E1 proteins may be covalent or non-covalent or indeed a combination of both. The observation that E1-specific oligomers exist under reducing conditions suggests that they are stabilised by non-covalent interactions. However, the level of E1 oligomerisation detected between E1 cysteine deletion mutants is reduced, which implies covalent as well as non-covalent interactions contribute to E1-E1 association.

Evidence obtained from endo H_f deglycosylation implied that E1 formed significantly more differentially glycosylated monomeric forms when expressed in the absence of E2. Four distinct monomeric E1-specific bands were detected, all of which were higher in molecular weight than the endo H_f deglycosylated form. The occurrence of four species implies the addition of four oligosaccharide moieties to E1, thus leaving one potential N-linked glycosylation site unmodified. Studies conducted by Meunier *et al.* (1999) examining the utilisation of N-linked glycosylation sites with E1 from strain H indicated that the C-terminal site at aa 325 is not modified. This was due to the presence of a proline residue at aa 328, since its substitution with a glycine residue led to the addition of N-linked oligosaccharide at position 325. In the absence of more direct evidence, as presented by Meunier *et al.* (1999), it was assumed that glycosylation did not occur at residue 325 in strains H77 and Glasgow since their sequences also contain the proline residue found in strain H.

Differential glycosylation was observed for strain H77 but not for strain Glasgow E1 in the absence of E2. This difference in behaviour of strains Glasgow and H77 E1 may be due to the differences in their amino acid sequences (see Appendix 2). Furthermore, these forms are not recognised as efficiently using the conformation specific MAAb AP497 (see Fig 5.4),

implying that the epitope identified by this antibody is dependent on proper glycosylation of E1 protein. Fournillier-Jacob *et al.* (1996) also observed the appearance of minor products, from E1₁₇₄₋₃₇₀ and E1₁₇₄₋₃₈₃ expressing constructs, which migrated faster than the major E1 product. These were assumed to be incompletely glycosylated forms of E1. The efficient glycosylation of protein is dependent on a sufficient pool of precursor oligosaccharide structures attached to dolichol lipid, the efficiency of oligosaccharyltransferase, and a properly oriented and accessible Asn-X-Ser/Thr sequence on a nascent polypeptide (Kornfeld & Kornfeld, 1985). The appearance of differentially glycosylated forms of E1 only in the absence of E2 is unlikely to be a result of depletion in oligosaccharide-dolichol lipid structure or the inefficiency of oligosaccharyltransferase. However, in the absence of E2 in the ER lumen, glycosylation sites in some nascent E1 molecules may not be accessible to the glycosylation machinery, for example due to misfolding. Evidence from examination of the oligomeric forms of E1 may support this proposal. The apparent molecular weight of the oligomers are smaller than predicted. In addition, the oligomers for the E1 mutant protein lacking all of the cysteine residues have slightly less mobility than those for the wild type form of E1 (see Fig 5.22, lanes 3 and 4; Fig 5.24, panel b, lanes 7 and 8). However, the endo H_f-digested products for E1 and E1_{Δ-cys} oligomers are almost identical (Fig 5.23, lanes 1-4). This suggests that the oligomers of E1_{Δ-cys} are glycosylated to a greater extent than those of wild type E1. Hence, disulphide bond formation in aggregates of E1 may render certain sites inaccessible to N-linked glycosylation. Alternatively, the various glycosylated forms of E1 may represent intermediate deglycosylated species. Degradation of incorrectly folded glycoproteins in the ER follows a pathway which involves deglycosylation and targeting to the proteasome (Suzuki *et al.*, 1998). Thus, E1 in the absence of E2 may encounter such deglycosylation processes but since the intermediate forms are relatively abundant, they appear capable of circumventing degradation.

Studies presented in Chapter 4 revealed that covalent as well as non-covalent bonds participated in stabilising the aggregated complexes between E1 and E2 proteins. In order to identify the regions involved in covalent and non-covalent interactions, a number of E1 internal deletion mutants were generated. Results indicated that native complexes did not form between mutant E1 and wild type E2, although E2 protein was co-precipitated from aggregated complexes formed by mutant E1. Therefore, it has been difficult to separate the two sets of interactions. Alterations to E1 appears to significantly affect the formation of a native complex, while aggregate formation may not be stringently dependent on

protein sequence. This data suggests that caution should be exercised on the authenticity of regions which are identified as necessary for E1E2 association. This is particularly the case where denatured proteins have been used to map sequences for E1E2 binding such as studies employing Far-Western blot analysis (Yi *et al.*, 1997a). While segments identified by such analysis may contribute to E1E2 association, their role in native complex formation as opposed to aggregation cannot be deduced.

Results with the E1 cysteine substitution mutants indicated that mutation of these residues abolished the formation of native complex, emphasising the sensitivity of E1 to alterations to its primary sequence. Under non-reducing conditions, E1 cysteine substitution mutants co-precipitated with E2, although, native E2 was not detected. Native complex formation was sensitive even to alteration at a single residue (Cys 207) thus reflecting the importance of cysteine residues in E1. In a study examining the role played by glycans on E1E2 native complex formation (Meunier *et al.*, 1999) also confirmed the sensitivity of this complex to changes. In that instance, native complex between strain H E1 and E2 was abolished in the absence of one glycosylation site at residue 305 (Meunier *et al.*, 1999).

The mutant in which all of the cysteine residues in E1 had been mutated also continued to interact with E2. It is presumed that this occurred through non-covalent association of E1 with aggregates of E2. The cysteine residues in E1 are well conserved in all genotypes and the results presented in this chapter indicate their importance to native complex formation. The data also provide firm evidence that not only do disulphide bonds exist between E1 and E2 in an aggregated complex but also non-covalent interactions can stabilise association between the glycoproteins in such complexes. This is in agreement with results presented in Chapter 4 where DTT treatment *in situ* could not disrupt E1E2 aggregate complexes.

As mentioned above, the inefficiency of MAb AP497 to immunoprecipitate the differentially glycosylated monomeric E1 species suggested that the epitope recognised by this antibody was dependent on glycosylation of E1. Additional to this, substitution of cysteine residues at 226, 229 and 238 in E1 abolished recognition of these mutants by MAb AP497, revealing that these residues also play an essential role in the formation of the AP497-specific conformation epitope. Interestingly, substitution of a single cysteine residue at aa 207 eliminated the formation of oxidised E1 and a reduced level of non-aggregated E1 for this mutant was precipitated by AP497. The involvement of cysteine

207 in E1 oxidation and the ability of AP497 to recognise a particular oxidised and glycosylated form of E1 remain to be defined.

This chapter has highlighted the requirement for E2 for the appropriate folding of E1, which in turn affect the formation of native complex. In the following chapter the properties of E2 important for E1E2 complex formation and E1 folding were further examined.

Strain	Protein	No. of E1 aa deleted	No. of potential N-glycan sites lost (aa) [#]	Predicted MW of nascent protein (kDa)	Predicted MW of modified protein (~kDa) [*]
Glasgow	E1 _{his}	-	-	21.9	29.9
	E1 _{Δ196-235} [‡]	38	3 (196, 209, 234)	17.7	19.7
	E1 _{Δ213-260}	46	1 (234)	16.9	22.9
	E1 _{Δ266-303}	36	0	18.0	26.0
	E1 _{Δ309-338}	28	0	18.8	26.8
	E1 _{Δ341-366}	24	0	19.4	27.4
H77	E1 _{WT}	-	-	20.9	28.9
	E1 _{Δ199-233}	33	1 (209)	17.3	23.3
	E1 _{Δ213-260}	46	1 (234)	16.0	22.0
	E1 _{Δ266-303}	36	0	17.0	25.0
	E1 _{Δ309-338}	28	0	17.8	25.8
	E1 _{Δ341-366}	24	0	18.4	26.4

Table 5.1. Predicted molecular weights (MW) of E1 internal deletion mutant proteins of strains Glasgow and H77. * It is assumed that addition of a high mannose oligosaccharide increased the molecular weight of nascent protein by ~2.0 kDa per glycosylation site (Kornfeld & Kornfeld, 1985). [#] It is assumed that glycosylation at amino acid residue 325 in E1 does not occur as reported by Meunier *et al.* (1999). [‡] Histidine residues corresponding to the tag replaced the deleted region within this mutant.

GGC **GTC** **GAC** **GCG** **GAG** **ACC** **CAC** **TAA** **GAT** **CTA** **AGC** **TTG**
G **V** **D** **A** **E** **T** **H** *
380 386

Fig 5.1. Construction of a plasmid expressing strain Glasgow E1 in the absence of E2. pGEM/E1_{his}E2 plasmid was linearised by digesting with *Sall* (which lies in the 3' coding region of E1 position 1483; Appendix 1) and *HindIII* (that lies in the pGEM vector downstream of E2 coding region). These sites were bridged using an oligonucleotide (over-lined in red). The oligonucleotide was designed to regenerate the proteinase cleavage site between E1-E2 at residues 383-384. This construct was termed pGEM/E1_{Gla}. A stop codon, introduced by the oligonucleotide is indicated by *. E1 residues are in brown, while E2 residues are shown in green.

GGC **GTC** **GAC** **GCG** **GAA** **ACC** **CAC** **TAA** **GAT** **CTG** **AAT** **TCC**
G **V** **D** **A** **E** **T** **H** *
380 386

Fig 5.2. Construction of a plasmid expressing strain H77 E1 in the absence of E2. The strategy employed was similar to that used to obtain strain Glasgow E1, except a *Sall*-*EcoRI* linear fragment from pGEM/E1E2_{H77} encoding the E1 sequences was re-circularised by inserting an oligonucleotide (over-lined in red). This construct was called pGEM/E1_{H77}. The key stated for Fig 5.1 applies.

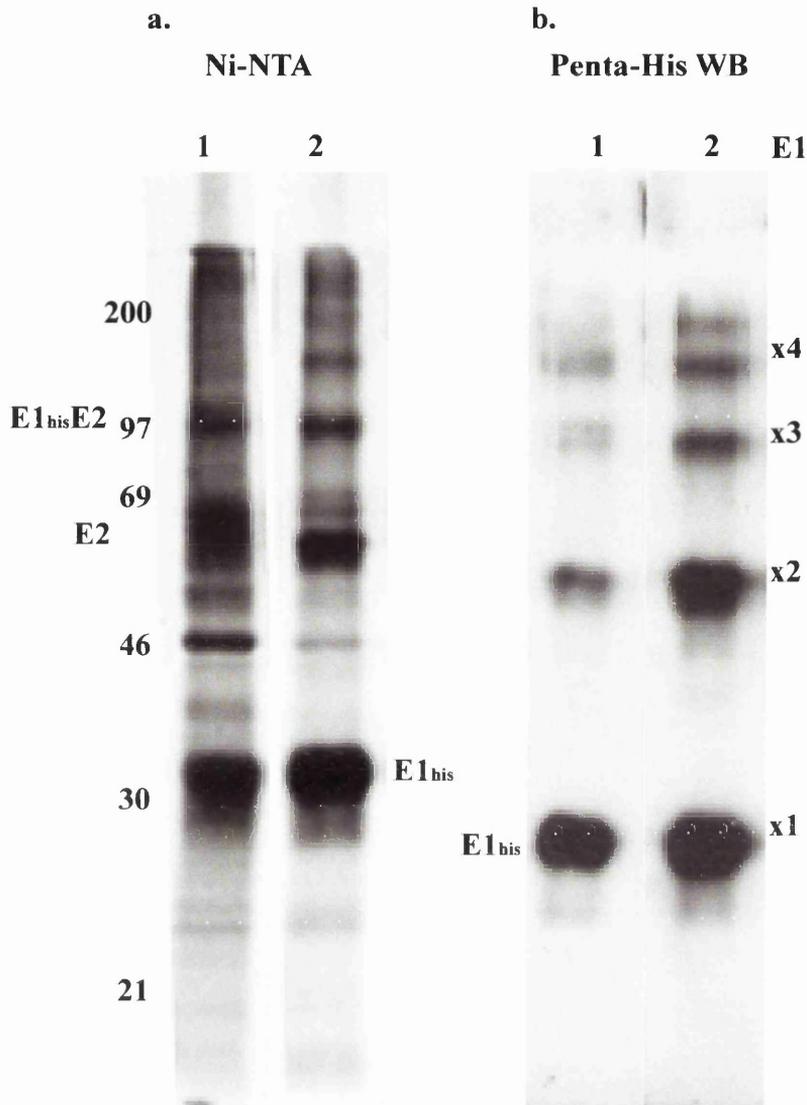


Fig 5.3. Behaviour of strain Glasgow E1 glycoprotein expressed in the absence of E2. Cells were electroporated with pSFV/E1_{his}E2 and pSFV/E1_{Gla} RNAs and radiolabelled from 4-12 h after electroporation. Crude cell extracts were prepared in binding buffer and applied to Ni-NTA agarose. The bound material was analysed under reducing conditions on a 10% polyacrylamide gel (panel a). Crude cell extracts from identical cell cultures, which were not radiolabelled, were examined by Western blot analysis using the Penta-his MAb to detect the histidine tag in E1 (panel b). Lanes 1 in both panels show material derived from crude extracts from pSFV/E1_{his}E2 RNA-electroporated cells and lanes 2 were from pSFV/E1_{Gla} RNA-electroporated cells. The positions of E1_{his}E2, E2 and the oligomers of E1_{his} are shown.

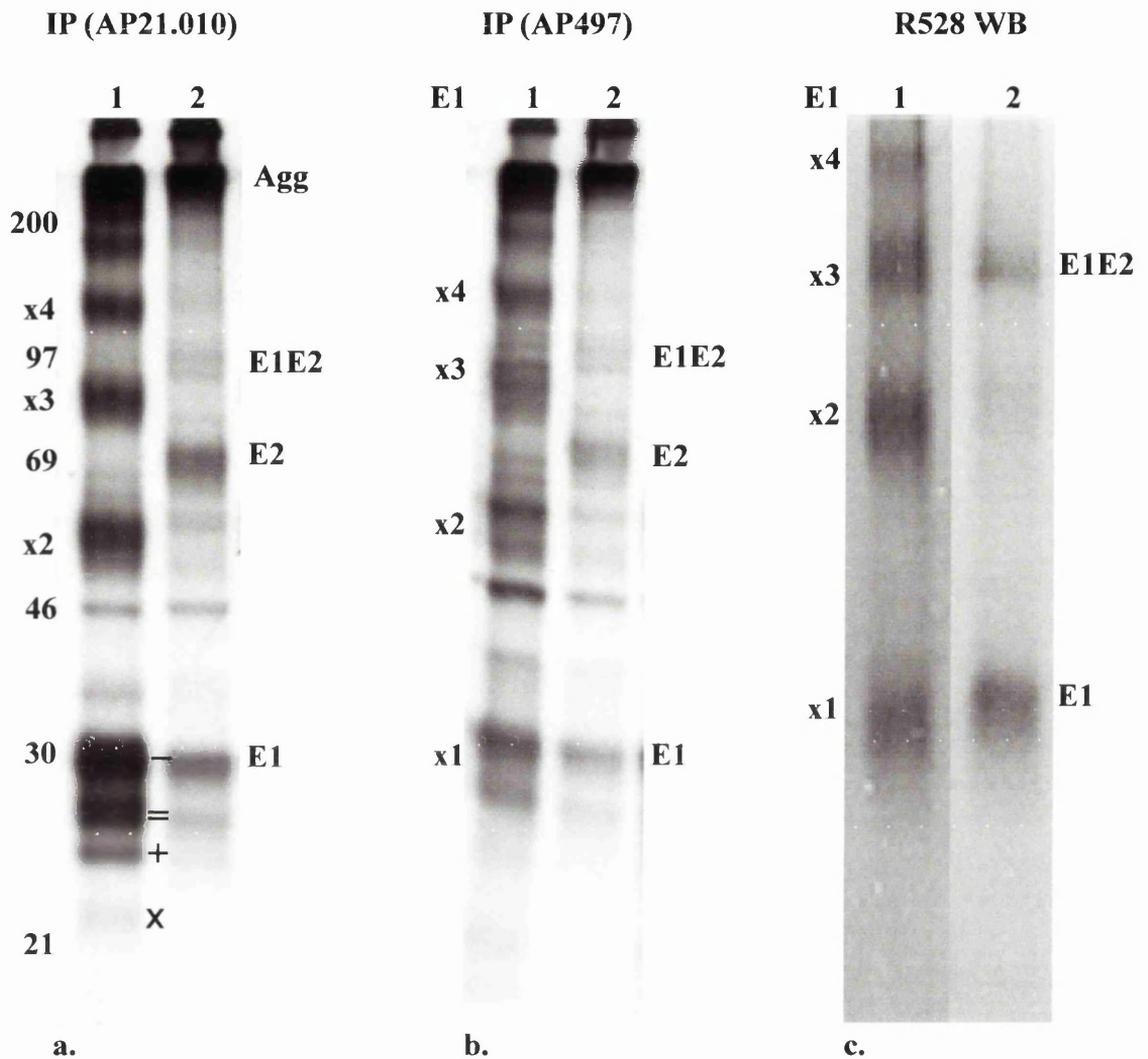


Fig 5.4. Behaviour of strain H77 E1 glycoprotein expressed in the absence of E2. Cells were electroporated with pSFV/E1E2_{H77} and pSFV/E1_{H77} RNAs and radiolabelled with ³⁵S-methionine between 4-12 h after electroporation. Immunoprecipitations were performed on the crude extracts obtained from these cells with E1-specific antibodies AP21.010 (panel a) and AP497 (panel b). Immunoprecipitates were analysed under reducing conditions on a 10% polyacrylamide gel. Western blot analysis, following transfer from a 10% polyacrylamide gel, was performed on crude extracts from duplicated cell cultures with an E1-specific antiserum, R528 (panel c). Lanes 2 show samples derived from cells electroporated with pSFV/E1E2_{H77} RNA and lanes 1 represent those obtained from cells electroporated with pSFV/E1_{H77} RNA. The positions of monomeric (x1), dimeric (x2), trimeric (x3), tetrameric (x4) E1 and E2 are indicated. Aggregated material (Agg) and E1-E2 precursor are also indicated. Bands marked by symbols are discussed in the text.

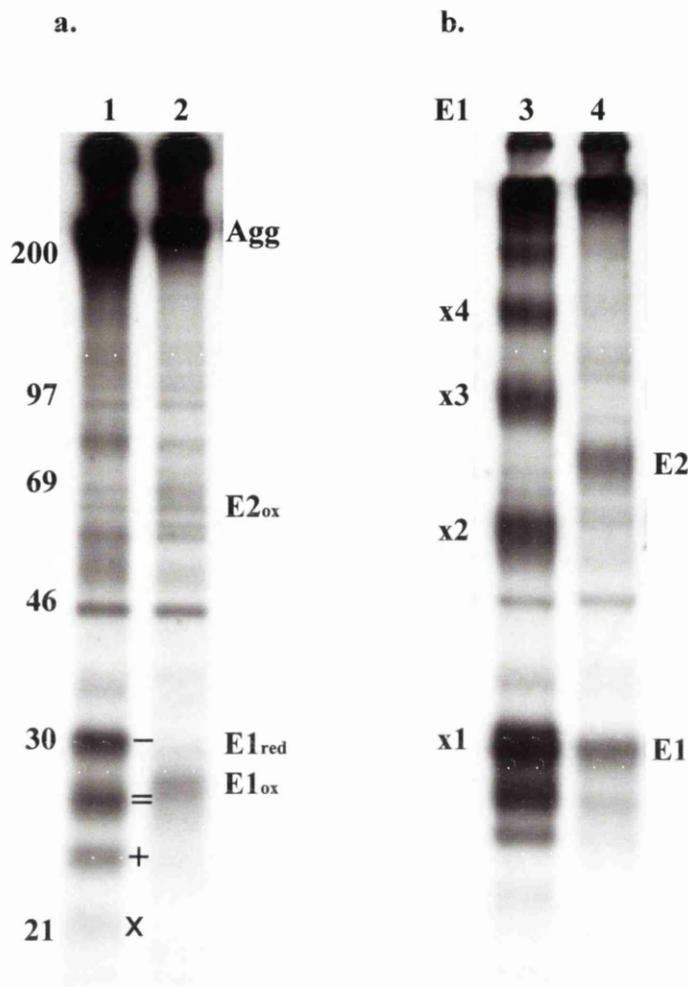


Fig 5.5. Oxidation state of strain H77 E1 glycoprotein expressed in the absence of E2. Experimental conditions were identical to those described in Fig 5.4, except immunoprecipitates obtained using AP21.010 were analysed under non-reducing (panel a) and reducing (panel b) electrophoresis conditions. Lanes 2 and 4 show precipitated material derived from cells electroporated with pSFV/E1E2_{H77} RNA and lanes 1 and 3 show precipitates from cells electroporated with pSFV/E1_{H77} RNA. The positions of oxidised (ox) and reduced (red) E1 and E2 are indicated. The aggregated material (Agg) is also shown. Bands marked by symbols are discussed in the text.

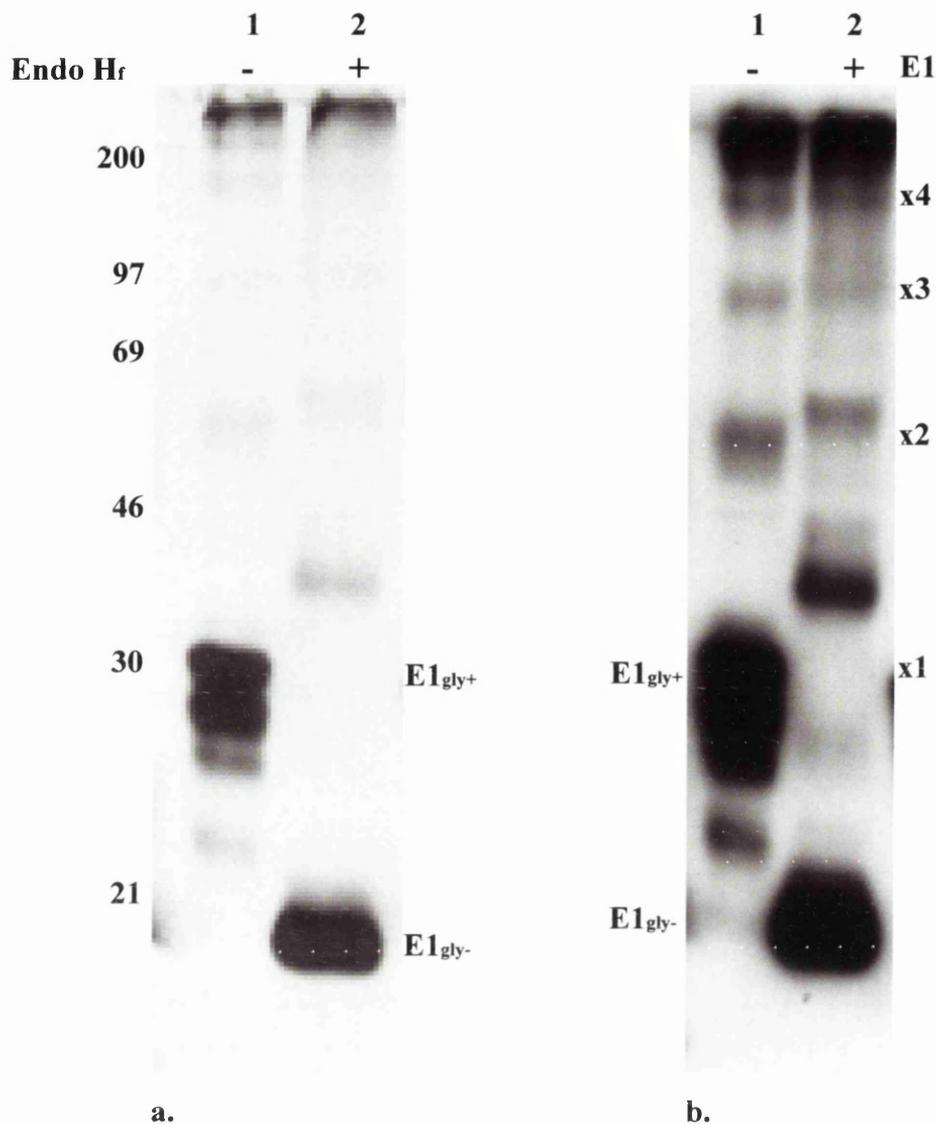


Fig 5.6. Analysis of strain H77 E1 glycoprotein by treatment with endo H_r. Cells were electroporated with pSFV/E1_{H77} RNA. Crude extract obtained from these cells 16 h after electroporation was treated with endo H_r before separation by electrophoresis on a 12% polyacrylamide gel (lane 2) alongside untreated crude extract (lane 1). Western blot analysis was performed, probing for E1 using the R258 antiserum. Panel a represents a lower exposure of the same blot shown in panel b. The positions of glycosylated (gly⁺) and de-glycosylated (gly⁻) E1 oligomeric forms are indicated.

Fig 5.7. Construction of strain Glasgow E1 deletion mutant constructs.

a. Two fragments generated by *HindIII*-*FspI*_{c195} (coding for the N-terminus of E1) and *XhoI*_{c236}-*HindIII* (coding for the remaining sequences of E1 and E2) derived from pGEM/E1_{his}E2 plasmid were ligated along with an oligonucleotide (over-lined in red) to bridge the *FspI* and *XhoI* sites. The *FspI* site is positioned at codon (c) 195 (nucleotide position 925; Appendix 1), *XhoI* is at 236 (nucleotide position 1049; Appendix 1) while the *HindIII* site is located within the pGEM1 vector downstream of E2 coding region. The oligonucleotide coded for sequences flanking the restriction sites as well as sequences coding for the histidine tag and generated a deletion between residues 196 and 235. The resultant construct was called pGEM/E1_{Δ196-235}E2_{Gla}.

b. *HindIII*-*XhoI*_{c209} and *ClaI*_{c262}-*HindIII* fragments from pGEM/E1_{his}E2 were ligated using an oligonucleotide to bridge the *XhoI* and *ClaI* sites (nucleotide positions 971 and 1127 respectively; Appendix 1). The oligonucleotide encoded for sequences flanking the restriction sites and generated a deletion between residues 213 and 260. The construct was called pGEM/E1_{Δ213-260}E2_{Gla}. The same oligonucleotide was employed to ligate *EcoRI*-*XhoI* and *ClaI*-*EcoRI* fragments derived from plasmid pGEM/E1E2_{H77}, where *EcoRI* was located downstream of E2 coding region. This construct was named pGEM/E1_{Δ213-260}E2_{H77}.

c. *HindIII*-*ClaI*_{c262} and *MunI*_{c305}-*HindIII* fragments from pGEM/E1_{his}E2 were ligated using an oligonucleotide to bridge between *ClaI* and *MunI* sites (nucleotide positions 1127 and 1256 respectively; Appendix 1). The oligonucleotide encoded for sequences flanking the restriction sites and introduced a deletion between residues 266 and 303. This construct was called pGEM/pSFV/E1_{Δ266-303}E2_{Gla}. The same oligonucleotide was used to generate an equivalent deletion in strain H77, in which case *EcoRI*-*ClaI* and *MunI*-*HindIII* fragments derived from pGEM/E1E2_{H77} were employed. This construct was called pGEM/E1_{Δ266-303}E2_{H77}.

d. *HindIII*-*MunI*_{c305} and *BamHI*_{c339}-*HindIII* fragments from pGEM/E1_{his}E2 were ligated using the oligonucleotide indicated, which generated a deletion between positions 309 and 338. The *MunI* and *BamHI* sites are located at nucleotide positions 1256 and 1358 respectively (Appendix 1). This construct was called pGEM/pSFV/E1_{Δ309-338}E2_{Gla}, and a construct containing a similar deletion in strain H77, using fragments *EcoRI*-*MunI* and *BamHI*-*EcoRI* derived from pGEM/E1E2_{H77}, was called pGEM/E1_{Δ309-338}E2_{H77}.

e. *HindIII*-*BamHI*_{c339} and *HincII*_{c381}-*HindIII* fragments from pGEM/E1_{his}E2 were joined by the oligonucleotide shown, which allowed the region between residues 341 and 366 to be deleted and re-introduced the E1 C-terminal sequences required for anchoring E1 to membrane and directing E2 sequences to the ER. This construct was called pGEM/E1_{Δ341-366}E2_{Gla}. The *BamHI* and *HincII* sites are located at nucleotide positions 1358 and 1485 respectively (Appendix 1).

Restrictions sites are shown in italic, histidine tag residues in pink, E1 amino acid sequences in brown and E2 signal sequence residues in purple.

Fig 5.7. a.

CAA GTG CGC AAC CAT CAT CAT CAT CAT CAT GCC TCG AGG TGT TGG
Q V R N H I H I H I H A S R C W
195 196 235 236

b.

OCT AAC TCG AGT ATT CGA CGT CAC ATC GAT CTG
P N S S I V R H I D L
209 213 260 262

c.

CGT CAC ATC GAT CTG CTG GTC GGG TGC AAT TGT TOC
R H I D L L V G C N C S
262 266 303 305

d.

GGC TGC AAT TGT TOC ATC TAT CTC CGG ATC CCA
G C N C S I Y L R I P
304 309 338 339

e.

CGG ATC CCC GGG AAC TGG GCG AAG GTC CTG GCA GTG CTG CTG CTA TTT GCC GGC GTC GAC
R I P G N W A K V L A V L L L F A G V D
339 341 366 381

GTG CGC AAT TCC TCG GGG GGT AAC GCG TCG AGG
 V R N S S G G N A S R
 194 199 233 237

Fig 5.8. Construction of a N-terminal deletion mutant for strain H77 E1. Two fragments *EcoRI-FspI*_{c195} and *XhoI*_{c236}-*EcoRI* obtained from pGEM/E1E2_{H77} were ligated using an oligonucleotide (over-lined in red) to bridge the *FspI* and *XhoI* sites. The *FspI* (nucleotide position 992) and *XhoI* (nucleotide position 1046) sites are shown in Appendix 3 and the *EcoRI* site was located in the pGEM vector backbone. The oligonucleotide encoded sequences flanking the restriction sites and introduced a deletion between residues 199 and 233. The construct was called pGEM/E1_{Δ199-233}E2_{H77}. Restriction sites are in italics and E1 residues are shown in brown.

CTC CGG ATC CCA GGG AAC TGG GCG AAG GTC CTG GTA GTG CTG CTG CTA TTT GGC GGC GTC GAC GCG
 L R I P G N W A K V L V V L L L F A G V D A
 338 341 366 381

Fig 5.9. Construction of a C-terminal deletion mutant for strain H77 E1. *EcoRI-BamHI*_{c339} and *Sall*_{c381}-*EcoRI* fragments obtained from pGEM/E1E2_{H77} were ligated using the oligonucleotide (over-lined in red) which bridged between *BamHI* and *Sall* sites. The *BamHI* (position 1357) and *Sall* sites (position 1482) are shown in Appendix 3 and the *EcoRI* site was located in the pGEM vector backbone. The oligonucleotide encoded sequences flanking the restriction sites and generated a deletion between residues 341 and 366. This construct was called pGEM/pSFV/E1_{Δ341-366}E2_{H77}. Restriction sites are in italics, E1 amino acid sequences in brown, and the E2 signal sequence in purple.

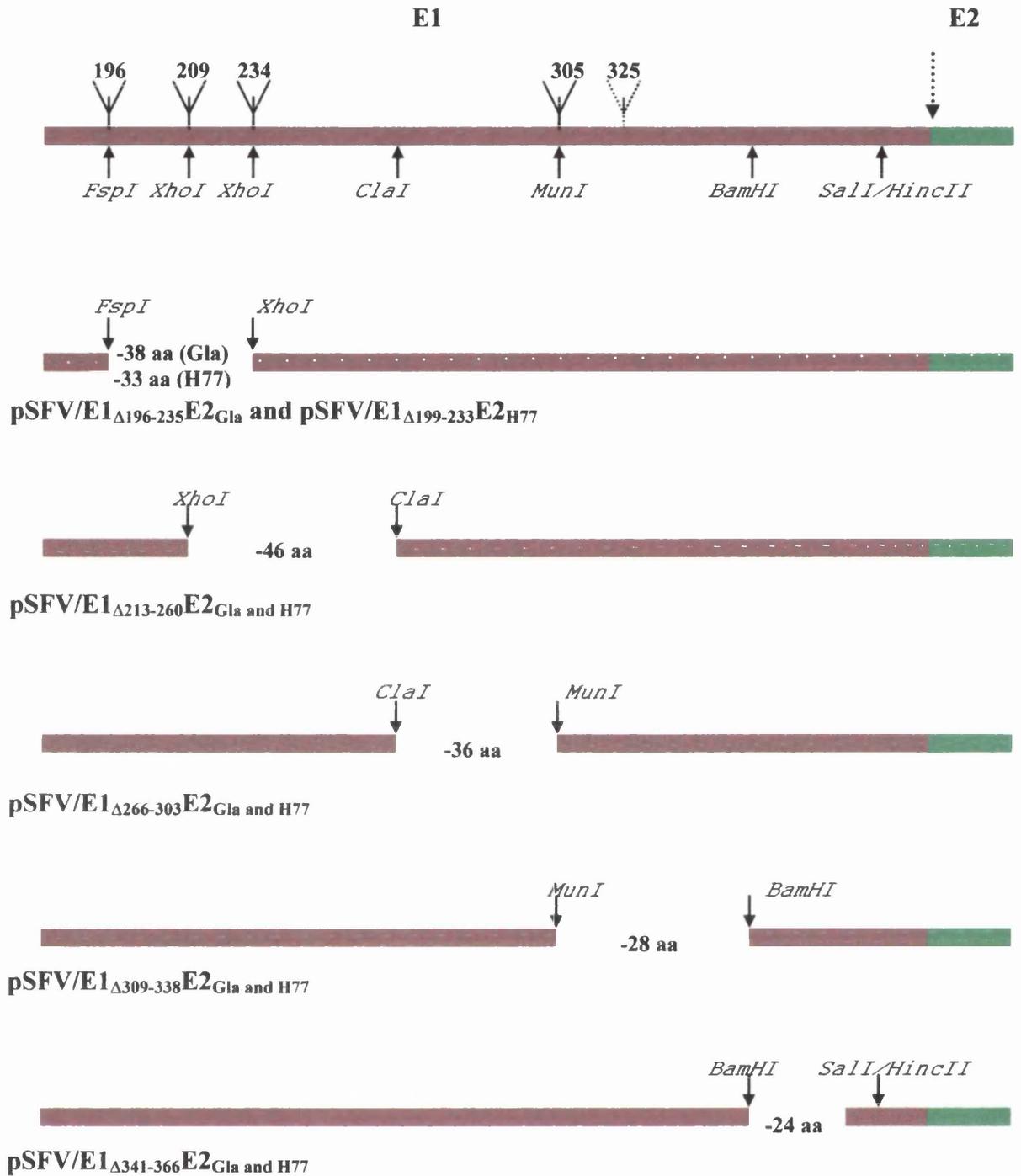


Fig 5.10. Schematic representation of the deletion mutants created in both strains Glasgow and H77 E1 glycoproteins. Restriction enzyme sites used in the construction of each mutant are shown, as are the numbers of residues removed in each case. The location of the 5 predicted glycosylation sites in E1 are indicated. The predicted site at residue 325 appears not to be utilised in tissue culture cells (Meunier *et al.*, 1999). A downward arrow indicates the predicted cleavage site between E1 (brown) and E2 (green).

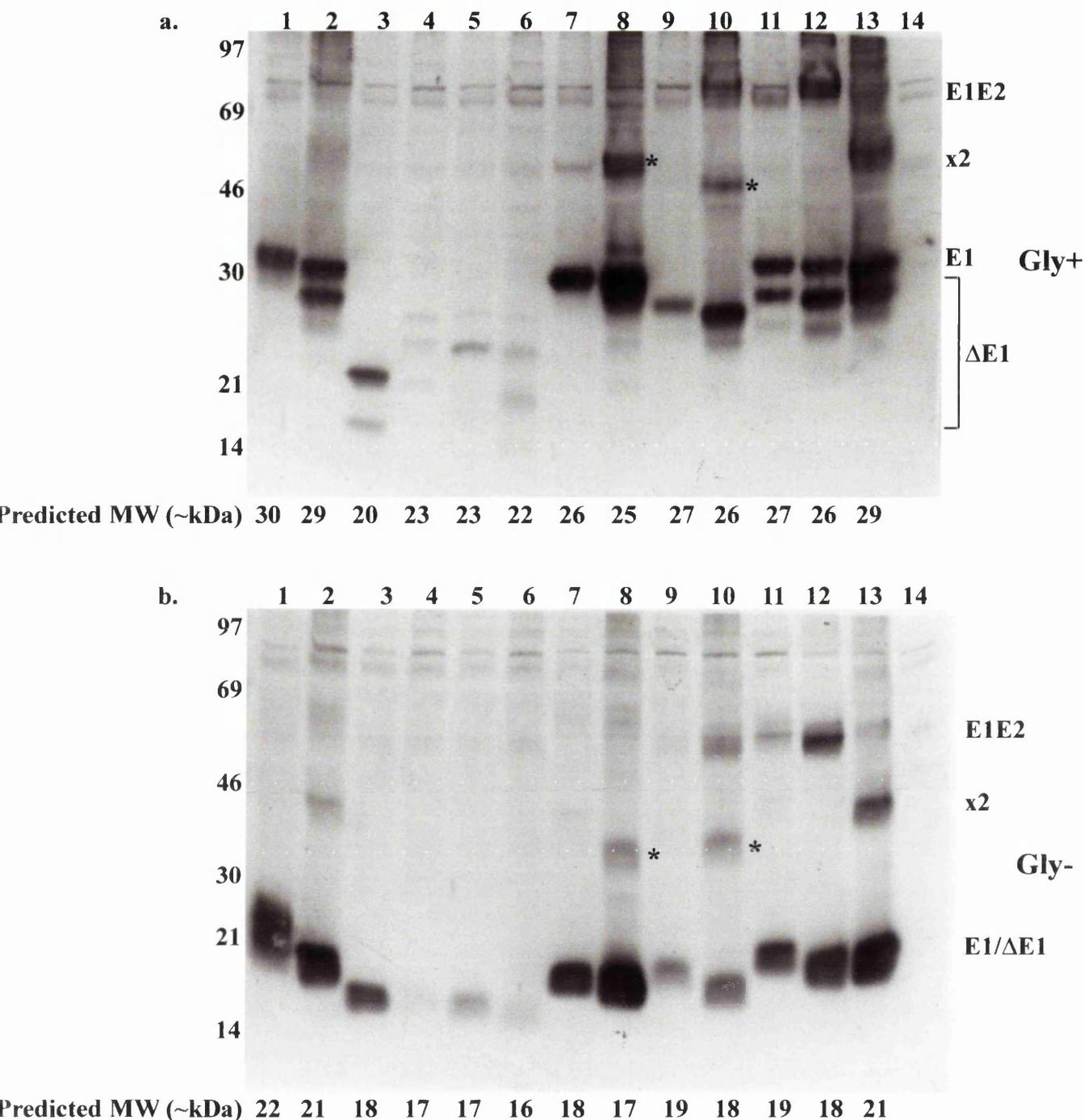


Fig 5.11. Endo H_f treatment of wild type and mutant E1 glycoproteins from strains Glasgow and H77. Cells were electroporated with RNA from SFV constructs. 12 h following electroporation, crude extracts were prepared and either treated or not treated with endo H_f before loading on 12% reducing polyacrylamide gels. The corresponding deletion mutants from strain Glasgow and H77 were analysed beside one another for comparative purposes. Proteins were transferred to membranes and examined by Western blot analysis using E1-specific antiserum, R528. Panels a and b show non-treated and treated crude extracts, respectively. Samples were pSFV/E1E2_{Gla} (lanes 1), pSFV/E1E2_{H77} (lanes 2), pSFV/E1_{Δ196-235}E2_{Gla} (lanes 3), pSFV/E1_{Δ199-233}E2_{H77} (lanes 4), pSFV/E1_{Δ213-260}E2_{Gla} (lanes 5), pSFV/E1_{Δ213-260}E2_{H77} (lanes 6), pSFV/E1_{Δ266-303}E2_{Gla} (lanes 7), pSFV/E1_{Δ266-303}E2_{H77} (lanes 8), pSFV/E1_{Δ309-338}E2_{Gla} (lanes 9), pSFV/E1_{Δ309-338}E2_{H77} (lanes 10), pSFV/E1_{Δ341-366}E2_{Gla} (lanes 11), pSFV/E1_{Δ341-366}E2_{H77} (lanes 12), pSFV/E1_{H77} (lanes 13), and pSFV1 (lanes 14). Positions of glycosylated (Gly⁺) and deglycosylated (Gly⁻) E1 molecules, as well as their predicted molecular weights (MW, below gel lanes) are shown. The bands indicated by * are examined further in Fig 5.16.

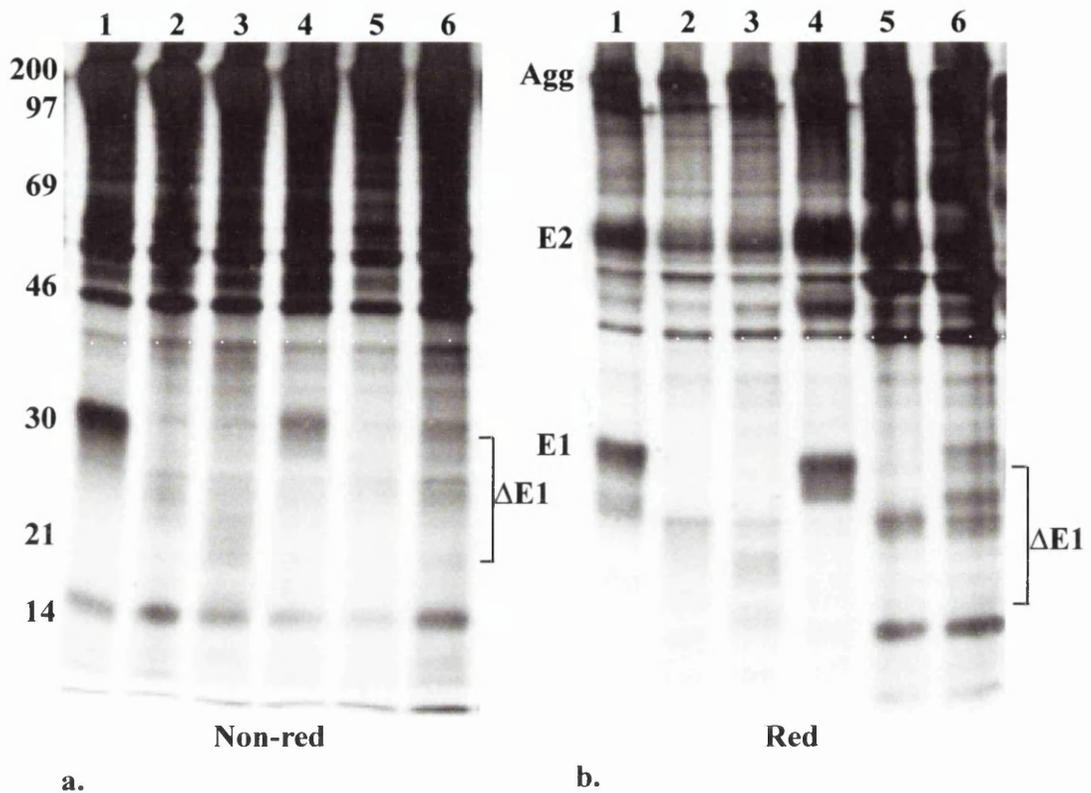


Fig 5.12. Examination of E1E2 complexes formed by H77 E1 deletion mutants using R528 anti-E1 antisera. Cells electroporated with RNA transcripts from pSFV wild type and mutant E1 constructs were radiolabelled from 4-12 h after electroporation with ^{35}S -methionine. Immunoprecipitations were performed on crude extracts from these cells using E1-specific antiserum R528. Precipitates were analysed on 12% polyacrylamide gels under non-reducing (panel a) and reducing (panel b) conditions. Samples were pSFV/E1E2_{H77} (lanes 1), pSFV/E1 Δ ₁₉₉₋₂₃₃E2_{H77} (lanes 2), pSFV/E1 Δ ₂₁₃₋₂₆₀E2_{H77} (lanes 3), pSFV/E1 Δ ₂₆₆₋₃₀₃E2_{H77} (lanes 4), pSFV/E1 Δ ₃₀₉₋₃₃₈E2_{H77} (lanes 5), and pSFV/E1 Δ ₃₄₁₋₃₆₆E2_{H77} (lanes 6). The positions of the aggregated material (Agg), E2, E1 and mutant E1 (Δ E1) proteins are indicated.

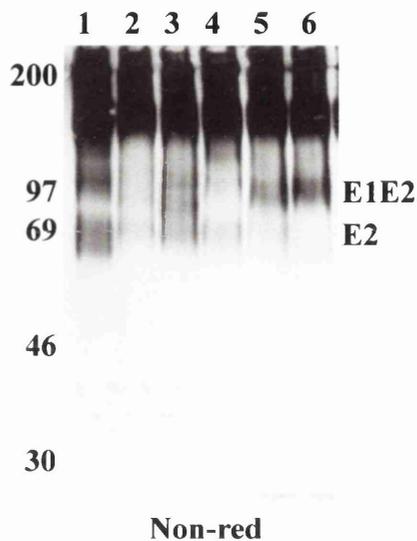


Fig 5.13. Examination of E2 present in E1E2 complexes formed by E1 deletion mutants by Western blot analysis. The experimental conditions were as described for Fig 5.12. with the exception that immunoprecipitations with R528 were performed on the crude extracts followed by Western blot analysis using E2-specific MAb ALP98. Samples were subjected to electrophoresis under non-reducing conditions on a 10% polyacrylamide gel. Precipitates are shown in the following order: pSFV/E1E2_{H77} (lane 1), pSFV/E1 Δ ₁₉₉₋₂₃₃E2 (lane 2), pSFV/E1 Δ ₂₁₃₋₂₆₀E2 (lane 3), pSFV/E1 Δ ₂₆₆₋₃₀₃E2 (lane 4), pSFV/E1 Δ ₃₀₉₋₃₃₈E2 (lane 5), and pSFV/E1 Δ ₃₄₁₋₃₆₆E2 (lane 6). Positions of E1E2 precursor and E2 are indicated.

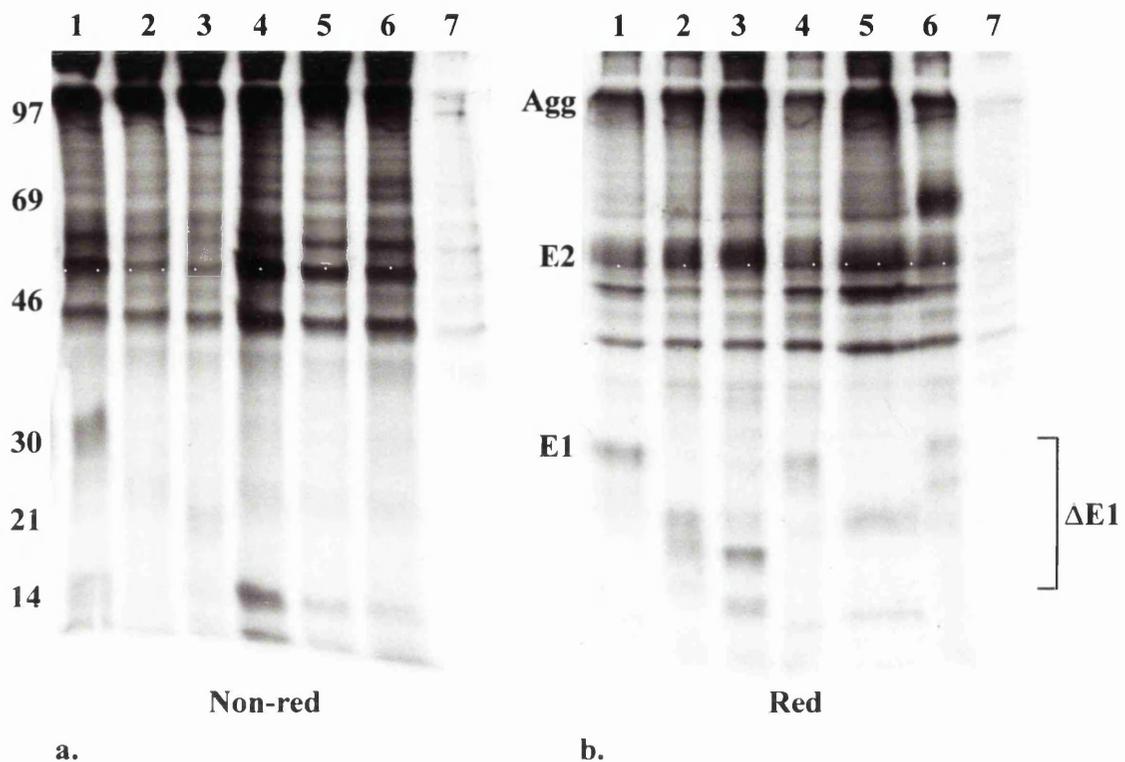


Fig 5.14. Examination of strain H77 E1 deletion mutants using E2-specific MAb ALP98. Cells were electroporated with RNA from pSFV wild type and mutant E1 constructs and pSFV1. Cells were radiolabelled between 4-12 h after electroporation with ^{35}S -methionine and lysed in immunoprecipitation buffer. The crude extracts were subjected to precipitation using E2-specific MAb ALP98. Precipitated material was analysed under non-reducing (panel a) and reducing (panel b) electrophoresis conditions on 10% polyacrylamide gels. Samples were pSFV/E1E2_{H77} (lanes 1), pSFV/E1 Δ ₁₉₉₋₂₃₃E2_{H77} (lanes 2), pSFV/E1 Δ ₂₁₃₋₂₆₀E2_{H77} (lanes 3), pSFV/E1 Δ ₂₆₆₋₃₀₃E2_{H77} (lanes 4), pSFV/E1 Δ ₃₀₉₋₃₃₈E2_{H77} (lanes 5), pSFV/E1 Δ ₃₄₁₋₃₆₆E2_{H77} (lanes 6), and pSFV1 (lanes 7). The positions of aggregated material (Agg), E2, E1 and mutant E1 (Δ E1) proteins are indicated.

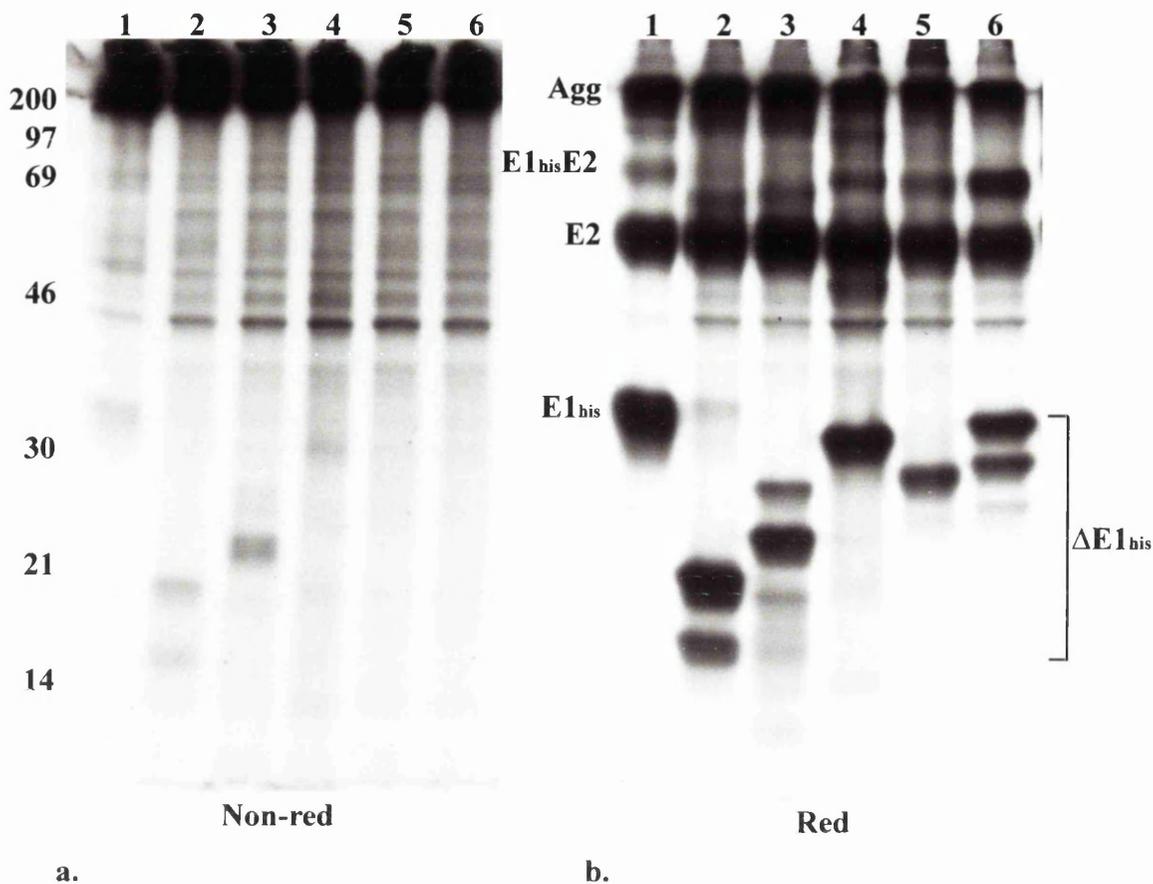


Fig 5.15. Examination of E1E2 complexes formed by strain Glasgow E1 deletion mutants using anti-E1 antiserum R528. Cells electroporated with RNA transcripts from pSFV wild type and mutant E1 constructs for strain Glasgow were radiolabelled between 4-12 h with ^{35}S -methionine. Immunoprecipitations were performed on crude extracts obtained from these cells using E1-specific antiserum, R528. Precipitates were derived from pSFV/E1_{his}E2 (lanes 1), pSFV/E1 $\Delta_{196-235}$ E2_{Gla} (lanes 2), pSFV/E1 $\Delta_{213-260}$ E2_{Gla} (lanes 3), pSFV/E1 $\Delta_{266-303}$ E2_{Gla} (lanes 4), pSFV/E1 $\Delta_{309-338}$ E2_{Gla} (lanes 5), and pSFV/E1 $\Delta_{341-366}$ E2_{Gla} (lanes 6), and analysed on 10% polyacrylamide gels under non-reducing (panel a) and reducing (panel b) conditions. Positions of the aggregated material (Agg), E1_{his}E2 precursor molecules, E2, E1_{his} and mutant E1_{his} ($\Delta\text{E1}_{\text{his}}$) proteins are indicated.

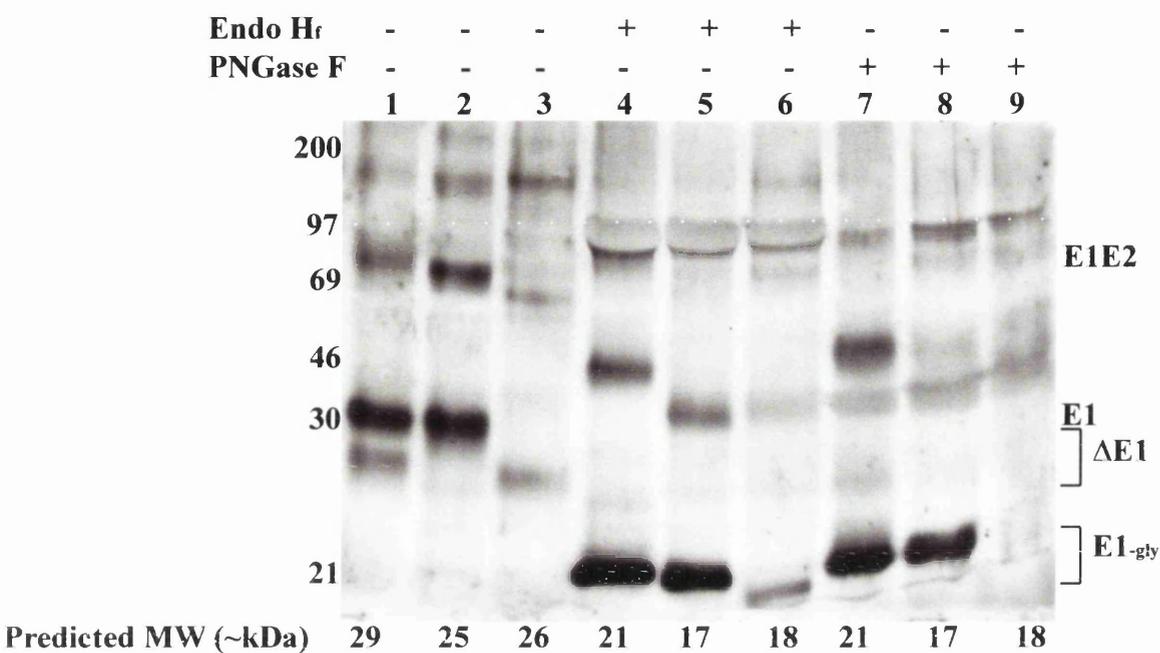


Fig 5.16. Endo H_f and PNGase F treatment of E1_{Δ266-303} mutant. Crude extracts derived from cells electroporated with RNA from pSFV/E1E2_{H77} (lanes 1, 4 and 7), pSFV/E1_{Δ266-303}E2_{H77} (lanes 2, 5 and 8), and pSFV/E1_{Δ309-338}E2_{H77} (lanes 3, 6 and 9) were subjected to deglycosylation by endo H_f (lanes 4-6) and PNGase F (lanes 7-9). Samples were analysed along with equivalent amounts of untreated material (lanes 1-3) under reducing conditions on a 12% polyacrylamide gel. Positions of E1E2, E1, mutant E1 (ΔE1) and deglycosylated E1 (-gly) proteins are shown.

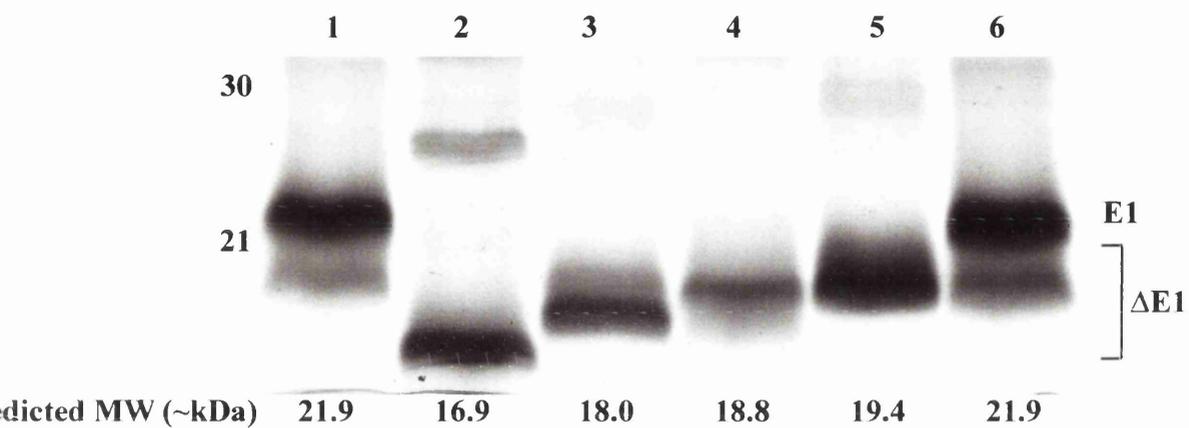


Fig 5.17. *In vitro* transcription and translation of E1 deletion mutants. *In vitro* transcription and translation reactions were performed from linearised pGEM vectors encoding strain Glasgow wild type and mutant E1 proteins. Reactions were performed in the presence of ^{35}S -methionine. Samples were analysed on a 12% polyacrylamide gel, and are presented in the following order: pGEM/E1_{Gla} (lanes 1 and 6), pGEM/E1 Δ ₂₁₃₋₂₆₀ (lane 2), pGEM/E1 Δ ₂₆₆₋₃₀₃ (lane 3), pGEM/E1 Δ ₃₀₉₋₃₃₈ (lane 4), and pGEM/E1 Δ ₃₄₁₋₃₆₆ (lane 5). Positions of E1 and mutant E1 proteins (Δ E1) are indicated.

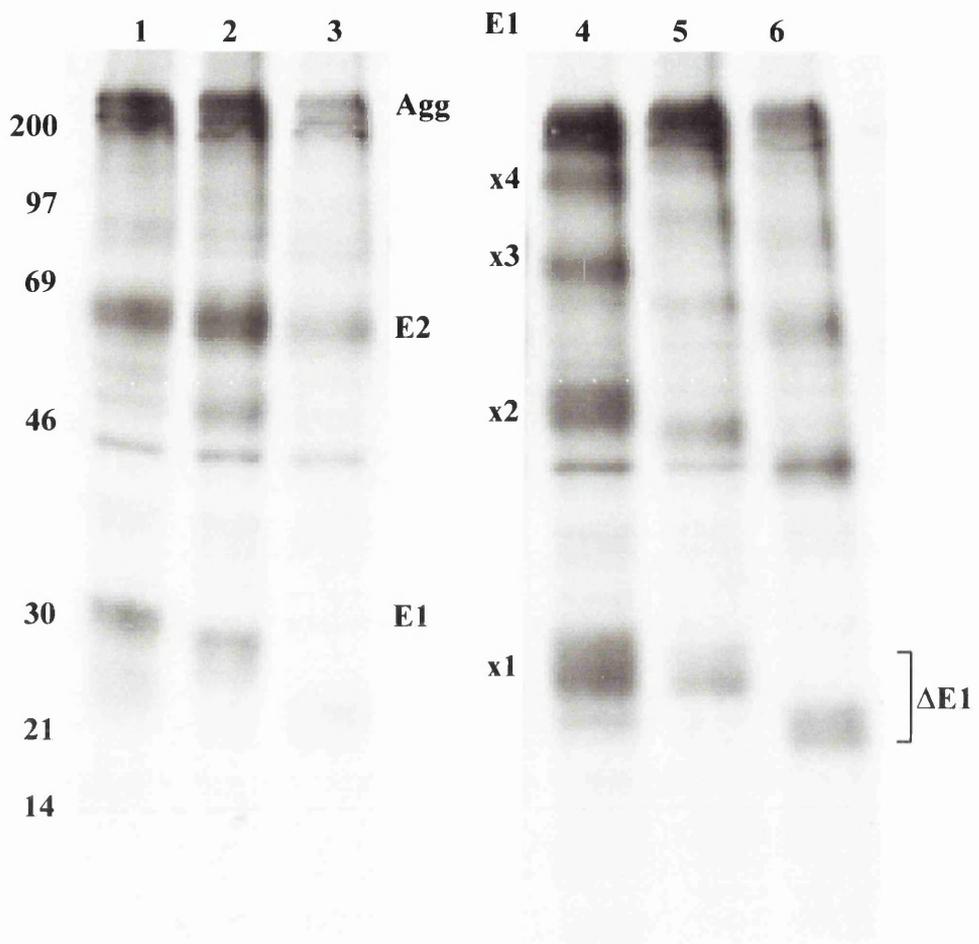


Fig 5.19. Oligomerisation pattern of mutant E1 molecules compared to oligomerisation of E1 expressed alone. Cells were electroporated and radiolabelled from 4-12 h after electroporation with ^{35}S -methionine before lysis in immunoprecipitation buffer. Precipitations were performed with crude cell extracts using E1-specific antiserum R528 and analysed under reducing conditions on a 12% polyacrylamide gel. Samples are pSFV/E1E2_{H77} (lane 1), pSFV/E1 Δ ₂₆₆₋₃₀₃E2_{H77} (lane 2), pSFV/E1 Δ ₃₀₉₋₃₃₈E2_{H77} (lane 3), pSFV/E1_{H77} (lane 4), pSFV/E1 Δ ₂₆₆₋₃₀₃H77 (lane 5), and pSFV/E1 Δ ₃₀₉₋₃₃₈H77 (lane 6). The positions of E2, E1, mutant E1 (Δ E1) and oligomeric E1 proteins are indicated.

Fig 5.20. Construction of cysteine substitution mutants in strain H77 E1. Cysteine substitution mutants were constructed by introducing point mutants via oligonucleotides inserted between appropriate restriction sites.

a. Schematic shows the positions of cysteine residues within E1 and the restriction enzyme sites used for cloning purposes.

b. Substitution of cysteine residue 1 (C1) (207). A plasmid backbone was derived from pSFV/E1E2_{H77} linearised with *NsiI* and *BamHI* (Fig 5.18). This fragment contained coding sequences for the E1 signal peptide upstream of the *NsiI* site (see Fig 4.10) and those for E2 downstream of the *BamHI* site. *NsiI-FspI*_{c195} and *EagI*_{c216}-*BamHI* fragments, derived from pGEM/E1E2_{H77}, were ligated into the backbone plasmid using an oligonucleotide which bridged the *FspI* and *EagI* sites. The oligonucleotide encoded all the residues between the *FspI* and *EagI* sites, except that a cysteine codon (207) was substituted with an alanine codon to prevent the generation of a novel N-linked glycosylation site. This construct was called pSFV/E1_{Δcys207}E2.

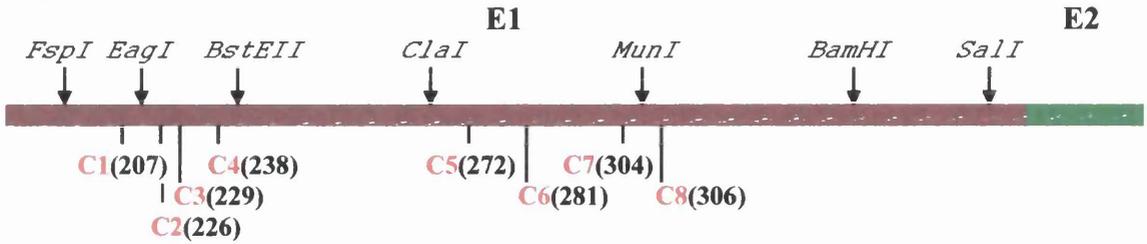
c. Substitution of residues C2(226), C3(229) and C4(238). *NsiI-EagI*_{c216} and *BstEII*_{c241}-*BamHI* fragments obtained from pGEM/E1E2_{H77} were ligated into the backbone plasmid described in part b using an oligonucleotide which bridged between *EagI* and *BstEII* sites. The oligonucleotide encoded the residues between *EagI* and *BstEII* sites except that cysteine codons (226, 229, and 238) were substituted for serine codons. This construct was named pSFV/E1_{Δcys226,229,238}E2.

d. Substitution of residues C5(272), C6(281) and C7(304). *NsiI-ClaI*_{c262} and *MunI*_{c304}-*BamHI* fragments obtained from pGEM/E1E2_{H77} were ligated into the backbone plasmid along with an oligonucleotide. This oligonucleotide bridged between *ClaI* and *MunI* sites and coded for the residues between *ClaI* and *MunI* except that cysteine codons (272, 281, and 304) were substituted for serine codons. This construct called pSFV/E1_{Δcys272,281,304}E2.

e. Substitution of C8(306) residue. A *NsiI-MunI*_{c304} fragment from pSFV/E1_{Δcys272,281,304}E2 was ligated into the backbone plasmid (part b) using an oligonucleotide which bridged between *MunI* and *BamHI*. The oligonucleotide coded for the residues between *MunI* and *BamHI* sites except that cysteine codon (306) was substituted for a serine codon. The plasmid was called pSFV/E1_{Δcys272,281,304,306}E2.

The restriction sites are shown in italics, amino acid residues in brown represent E1 and residues in red show the modifications.

Fig 5.20. a.



b.

```

OGC AAT TOC TOG GGG CTT TAC CAT GTC ACC AAT GAT GOC OCT AAC TOC AGT ATT GTG TAC GAG GOG
R  N  S  S  G  L  Y  H  V  T  N  D  A  P  N  S  S  I  V  Y  E  A
195                                     ↑
                                     C
                                     207
    
```

c.

```

GOG GOC GAT GOC ATC CTG CAC ACT CCG GGG TTT GTC OCT TOC GGT CGC GAG GGT AAC GOG TOG AGG AGT
A  A  D  A  I  L  H  T  P  G  S  V  P  S  V  R  E  G  N  A  S  R  S
216                                     ↑           ↑           ↑
                                     C           C           C
                                     226         229         238

TGG GTG GCG
W  V  A
    241
    
```

d.

```

ATC GAT CTG CTT GTC GGG AGC GOC ACC CTC AGC TOG GOC CTC TAC GTA GGG GAC CTG TIC GGG TCT GTC TTT
I  D  L  L  V  G  S  A  T  L  S  S  A  L  Y  V  G  D  L  S  G  S  V  F
262                                     ↑           ↑           ↑
                                     C           C           C
                                     272         281         345

CIT GGT GGT CAA CIG TTT ACC TTC TCT CCC AGG GGC CAC TGG ACG ACG CAA GAC AGC
L  V  G  Q  L  F  T  F  S  P  R  R  H  W  T  T  Q  D  S
346                                     ↑
                                     C
                                     304
    
```

e.

```

AAT TCT TCT ATC TAT CCC GGC CAT ATA ACG GGT CAT GGC ATG GCA TGG GAT ATG ATG ATG AAC TGG TOC OCT
N  S  S  I  Y  P  G  H  I  T  G  H  R  M  A  W  D  M  M  M  N  W  S  P
305 ↑
    C
    306

ACG GCA GGG TTG GTG GTA GCT CAG CTC CTC CG
T  A  A  L  V  V  A  Q  L  L  R
329                                     339
    
```

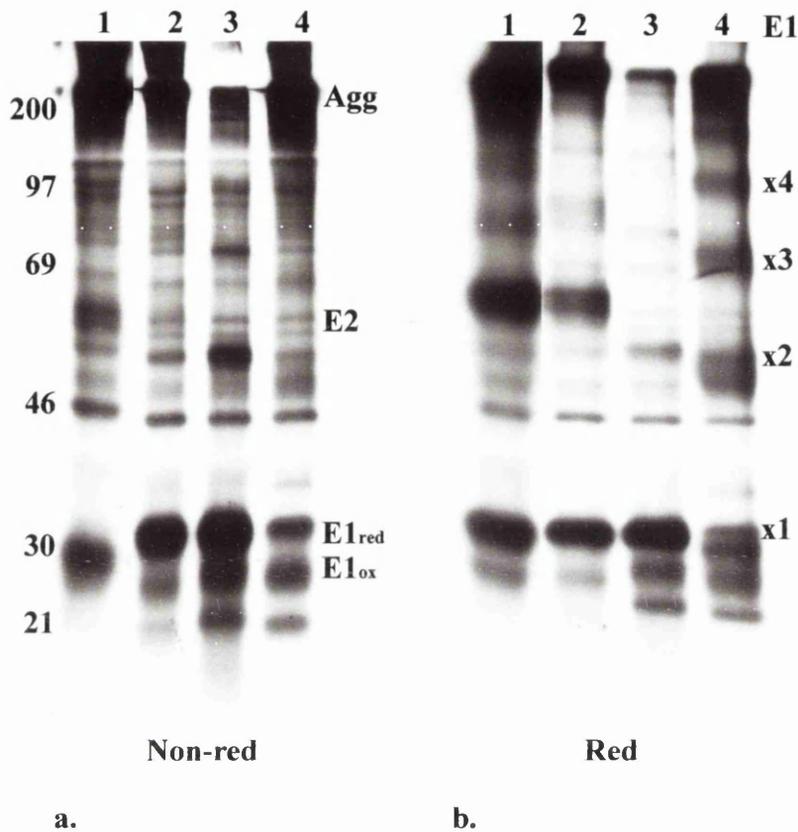



Fig 5.22. Examination of strain H77 E1 cysteine substitution mutants using anti-E1 antiserum, R528. Cells were electroporated with RNA followed by radiolabelling between 4-12 h after electroporation with ^{35}S -methionine. Crude extracts were subjected to immunoprecipitation with E1-specific antiserum, R528 and precipitates were analysed under non-reducing (panel a) and reducing (panel b) conditions on 10% polyacrylamide gels. Samples were from cells electroporated with RNA from the following constructs: pSFV/E1E2_{H77} (lanes 1), pSFV/E1 Δ -cysE2 (lanes 2), pSFV/E1 Δ -cys (lanes 3), and pSFV/E1_{H77} (lanes 4). Positions of the aggregated material (Agg), E2, oxidised (ox) and reduced (red) forms of E1, and oligomeric E1 proteins are shown.

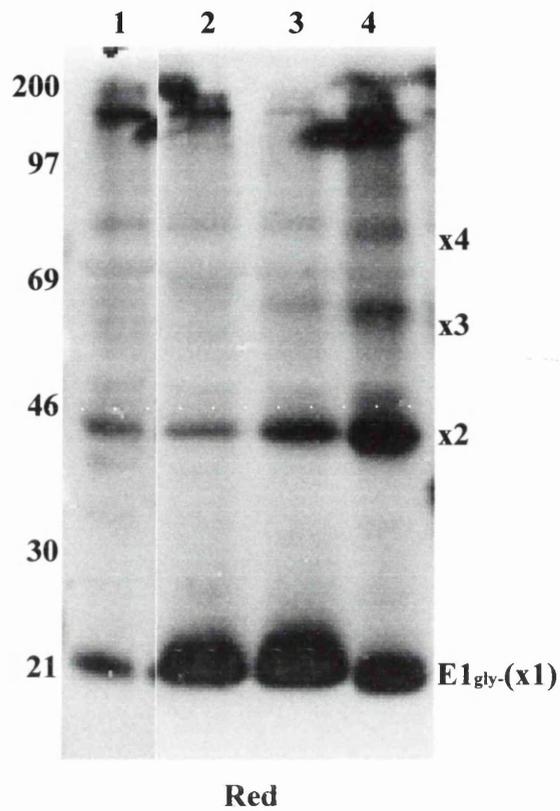


Fig 5.23. Endo H_f treatment of strain H77 E1 cysteine substitution mutants. Experimental conditions were as for those described for Fig 5.22 except the crude extracts obtained were treated with endo H_f, separated on a 12% reducing polyacrylamide gel and transferred to membrane. Western blot analysis was performed using anti-E1 antiserum, R528. Samples were derived from pSFV/E1E2_{H77} (lane 1), pSFV/E1_{Δ-cys}E2 (lane 2), pSFV/E1_{Δ-cys} (lane 3), and pSFV/E1_{H77} (lane 4). The positions of monomeric and oligomeric forms of deglycosylated (gly-) E1 are shown.

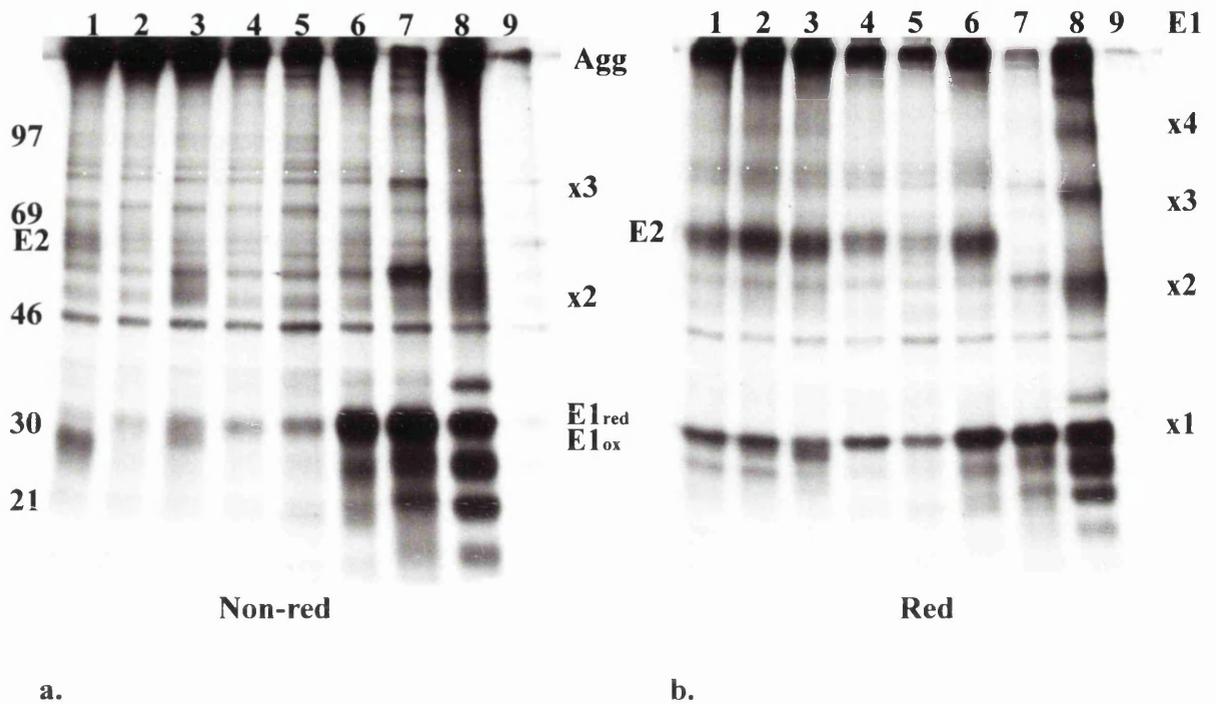


Fig 5.24. Examination of intermediate E1 cysteine substitution mutants using AP21.010. Cells, electroporated with RNA from pSFV constructs, were radiolabelled from 4-12 h after electroporation with ³⁵S-methionine and lysed in immunoprecipitation buffer. The crude extracts were subjected to immunoprecipitation using E1-specific MAb AP21.010 and precipitates were analysed under non-reducing (panel a) and reducing (panel b) conditions on 12% polyacrylamide gels. Samples were pSFV/E1E2_{H77} (lanes 1), pSFV/E1 Δ Cys₂₀₇E2 (lanes 2), pSFV/E1 Δ Cys_{226,229,238}E2 (lanes 3), pSFV/E1 Δ Cys_{272,281,304}E2 (lanes 4), pSFV/E1 Δ Cys_{272,281,304,306}E2 (lanes 5), pSFV/E1 Δ -cysE2 (lanes 6), pSFV/E1_{H77} (lanes 7), pSFV/E1_{H77} (lanes 8), and pSFV1 (lanes 9). The positions of aggregated material (Agg), E2, oxidised (ox), reduced (red) and oligomeric E1 proteins are shown.

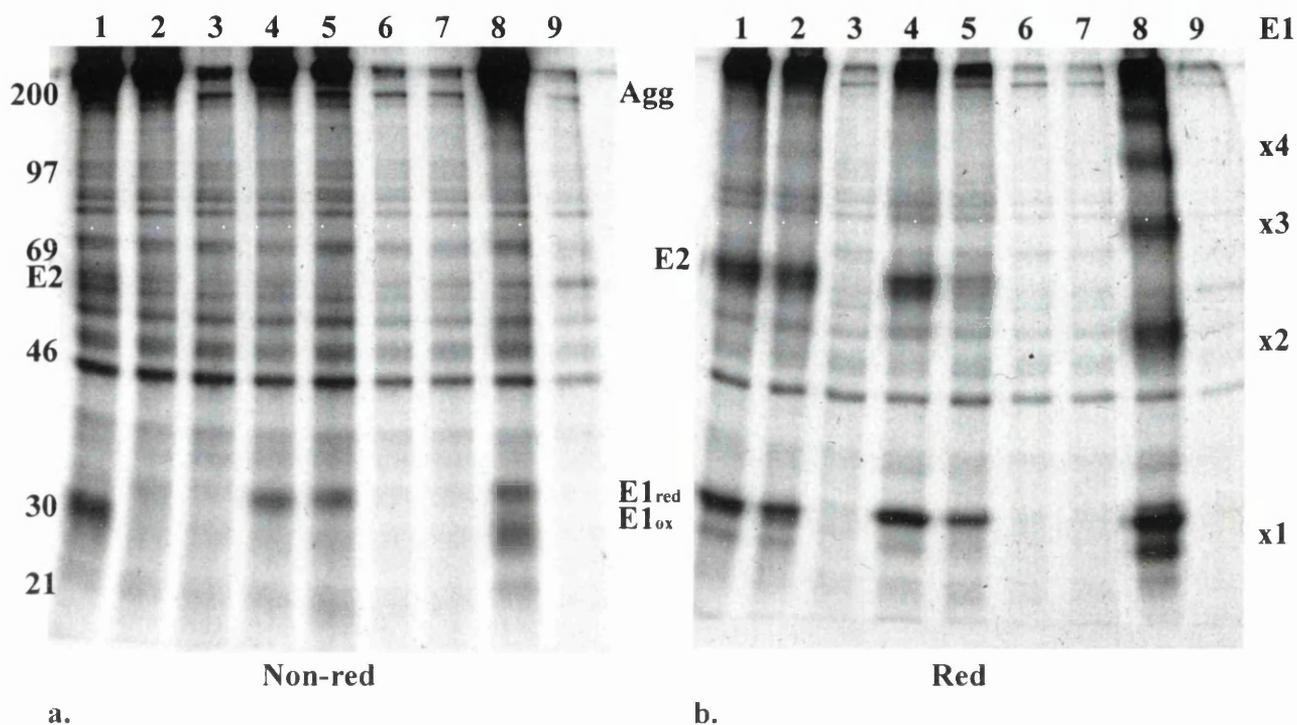


Fig 5.25. Examination of intermediate E1 cysteine substitution mutants using E1-specific MAb AP497. Experimental conditions were as for those described for Fig 5.24, except crude extracts were subjected to immunoprecipitation using E1-specific MAb AP497 and precipitates were analysed under non-reducing (panel a) and reducing (panel b) conditions on 10% polyacrylamide gels. Samples were pSFV/E1E2_{H77} (lanes 1), pSFV/E1 Δ Cys207E2 (lanes 2), pSFV/E1 Δ Cys226,229,238E2 (lanes 3), pSFV/E1 Δ Cys272,281,304E2 (lanes 4), pSFV/E1 Δ Cys272,281,304,306E2 (lanes 5), pSFV/E1 Δ -cysE2 (lanes 6), pSFV/E1 Δ -cys (lanes 7), pSFV/E1_{H77} (lanes 8), and pSFV1 (lanes 9). The positions of aggregated material (Agg), E2, oxidised (ox), reduced (red) and oligomeric E1 proteins are shown.

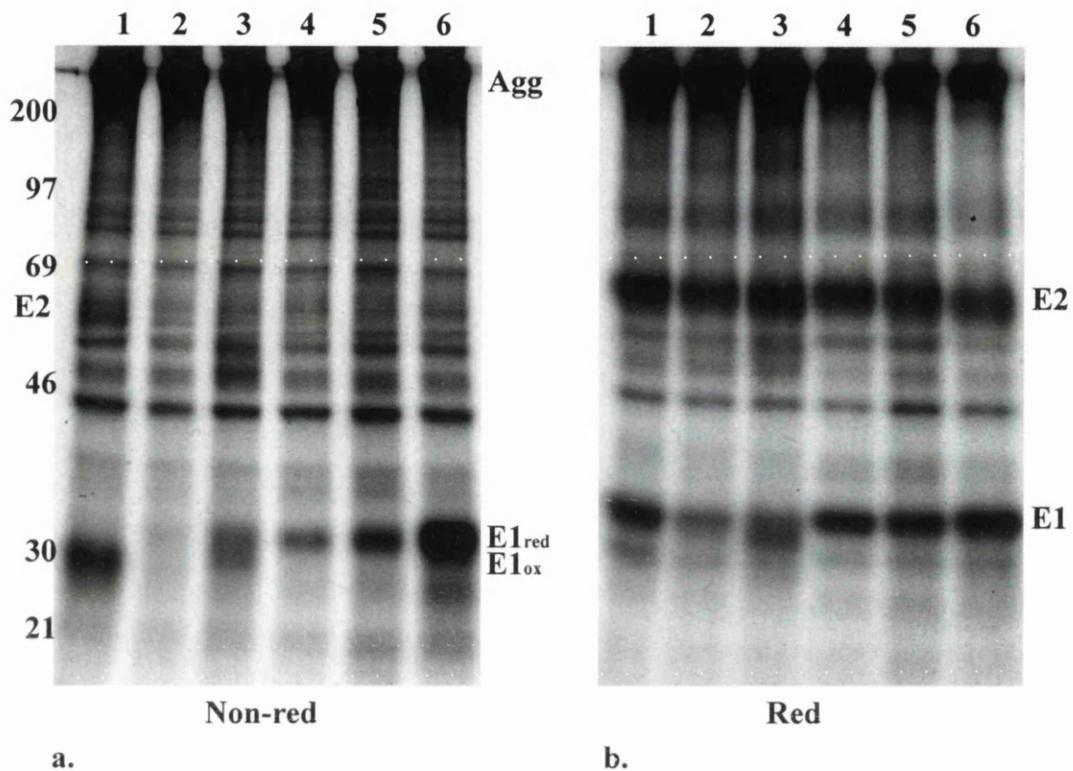


Fig 5.26. Examination of intermediate E1 cysteine substitution mutants using E2-specific ALP98. Experimental conditions were as for those described for Fig 5.24 except crude extracts were subjected to immunoprecipitation using E2-specific MAb ALP98. Precipitates were analysed under non-reducing (panel a) and reducing (panel b) conditions on 10% polyacrylamide gels. Samples were pSFV/E1E2_{H77} (lanes 1), pSFV/E1 Δ Cys₂₀₇E2 (lanes 2), pSFV/E1 Δ Cys_{226,229,238}E2 (lanes 3), pSFV/E1 Δ Cys_{272,281,304}E2 (lanes 4), pSFV/E1 Δ Cys_{272,281,304,306}E2 (lanes 5), and pSFV/E1 Δ -cysE2 (lanes 6). The positions of aggregated material (Agg), E2, oxidised (ox), and reduced (red) E1 proteins are shown.

Chapter 6 – Identification of a Region in E2 Required for the Folding of E1

6.1.1. Introduction

Results discussed in the preceding chapter had concentrated on identifying the properties in E1 that contribute towards the formation of a complex between E1 and E2. As a logical progression, the contribution made by E2 with respect to complex formation was analysed. The strategy employed was similar to that used for studying E1, where internal deletions were introduced to identify the regions of the polyprotein involved in interacting with E1. In this case, the internal deletion mutants were only created in strain Glasgow as this was the only strain available at the time during which the construction of these mutants was underway. The final mutant in this series consisted of a truncated version of the strain Glasgow E2 that lacked the transmembrane domain. This particular mutation was duplicated in strain H77 for corroboration purposes. As a consequence of the results obtained from the investigation of truncated E2, the ecto- and transmembrane domains of E2 were used to analyse their influence on the folding of E1 and the formation of native complex.

6.1.2. Construction of Internal Deletion Mutants in Strain Glasgow E2

To identify regions within E2 involved in association with E1, a panel of E2 deletion mutants in strain Glasgow was constructed in which regions were sequentially removed. A detailed schematic diagram illustrating the construction of these deletion mutants is presented in Fig 6.1. In a strategy similar to that used for creating E1 internal deletion mutants, plasmid pGEM/E1_{his}E2 (Section 3.2) was used as the parent construct. An N-terminal deletion in E2 was designed to remove sequences coding for residues 384 to 411. In this instance an oligonucleotide was inserted between *Sall* (at codon 381) and *BclI* (at codon 411) restriction enzyme sites, which reconstructed the predicted cleavage site between E1 and E2 at residues 383/384 (Fig 6.1, panel a). This construct was termed pGEM/E1_{his}E2_{Δ384-411}. Subsequent deletions in the E2 protein coding region were introduced using restriction sites *BclI*, *MscI* (at codon 457), *AscI* (at codon 523) and *HincII*

(at codon 610) in the nucleotide sequence (Appendix 1). Once again the oligonucleotides employed in each instance encoded a few residues on either side of the deleted sequences and regenerated the restriction sites. These constructs were named pGEM/E1_{his}E2 Δ ₄₁₅₋₄₅₄, pGEM/E1_{his}E2 Δ ₄₅₉₋₅₂₀, and pGEM/E1_{his}E2 Δ ₅₂₈₋₆₀₉ (Fig 6.1, panels b-d). The final internal deletion mutant was engineered to remove sequences between codon 610 and 699 and was created using an alternative strategy. In this case, a PCR fragment was amplified and inserted at suitable restriction sites (Fig 6.2). The 5'-primer was designed to anneal to sequences encoding residues 699 to 706 and carried sequences for codons 609 and 610, which included an in-frame *Sall* restriction enzyme site. The second primer was complementary to the SP6 promoter sequences in the pGEM1 vector (see Fig 3.3, panel a). This allowed amplification of the C-terminal coding region of E2 from pGEM/E1_{his}E2, thus ensuring that the stop codon, *BglII* and *HindIII* sites downstream of the E2 coding sequences were maintained. The PCR product was digested with *Sall* and *HindIII* and the resultant fragment was inserted into pGEM/E1_{his}E2 linearised with *Sall* (situated at codon 609) and *HindIII* enzymes. This construct was called pGEM/E1_{his}E2 Δ ₆₁₀₋₆₉₉.

Another construct made in this series encoded an E2 mutant truncated at residue 702, thereby removing the predicted C-terminal end of E2 and the p7/NS2 coding sequences. The truncated mutant was created using an oligonucleotide depicted in Fig 6.3 that was inserted between the *ScaI* site, at codon 701, and *HindIII* site situated in the pGEM1 vector of construct pGEM/E1_{his}E2 (Fig 3.4). This construct was called pGEM/E1_{his}E2_{384-702Gla}, which designates the encoded section of strain Glasgow E2. Mutant E2 coding sequences constructed in the pGEM1 vector were introduced into the pSFV1 expression vector using the *BglII* sites flanking the ORF. The resulting constructs were named as follows: pSFV/E1_{his}E2 Δ ₃₈₄₋₄₁₁, pSFV/E1_{his}E2 Δ ₄₁₅₋₄₅₄, pSFV/E1_{his}E2 Δ ₄₅₉₋₅₂₀, pSFV/E1_{his}E2 Δ ₅₂₈₋₆₀₉, pSFV/E1_{his}E2 Δ ₆₁₀₋₆₉₉, and pSFV/E1_{his}E2_{384-702Gla}. A schematic of the regions deleted in the E2 proteins is shown in Fig 6.4. Collectively, these plasmids were termed the pSFV/E1_{his}E2 Δ series.

For control purposes, a plasmid expressing E2/p7 and the N-terminal portion of NS2 polypeptides was constructed. This was achieved by excising a *BamHI*-*BglII* fragment from pGEM/E1_{his}E2 plasmid, and inserting it into the pSFV1 expression vector. The *BamHI* site in E1 is located at codon 340 and the resultant fragment incorporated the E1 transmembrane domain and E2 signal sequence. The *BglII* site is situated at the 3'-

terminal end of the ORF (extending to codon 851) in pGEM/E1_{his}E2. An ATG codon located at codon 347 was predicted to be utilised as a start codon. This construct was termed pSFV/E2_{340-851Gla}.

6.1.3. Properties of the E2 Deletion Mutants

Before investigating the consequences of mutagenesis of E2 on heteromeric complex formation with E1, proteins expressed by the constructs described above were initially characterised in terms of their molecular weights, glycosylation pattern and behaviour under non-reducing polyacrylamide gel electrophoresis conditions. Cells were electroporated with RNA transcribed from the constructs encoding strain Glasgow wild type (pSFV/E1_{his}E2) and mutant E2 (pSFV/E1_{his}E2 Δ) plasmids. For control purposes RNA from pSFV/E2_{340-851Gla} and pSFV1 were also used to electroporate cells. Following incubation for 12h at 37°C, extracts from cells were prepared in immunoprecipitation buffer. A portion of each extract was subjected to treatment with endo H_f to remove N-linked oligosaccharides and separated by polyacrylamide gel electrophoresis. Equal amounts of untreated crude cell extracts were separated by electrophoresis in parallel and analysed by Western blot using the E2-specific MAb, ALP98, (Fig 6.5, panels a and b). The same nitrocellulose blots were subsequently stripped of immunoglobulins and probed again with the anti-E2 antiserum, R141, (Fig 6.5, panels c and d). This second antibody was included because the linear epitope recognised by ALP98 MAb was found to lie within residues 610 and 699 of E2, a region which is deleted in mutant E2 Δ ₆₁₀₋₆₉₉ (panel a, lane 6).

Comparing the predicted molecular weights of wild type and E2 mutants (Table 6.1) and Western blot data of non-endo H_f treated proteins (panels a and c), all of the proteins migrated according to their estimated sizes. It was assumed that all eight potential glycosylation sites in wild type E2 were occupied by high mannose oligosaccharides. Three mutants had sequences removed that contained glycosylation sites. The E2 proteins in these cases lost 4 (pSFV/E1_{his}E2 Δ ₄₁₅₋₄₅₄), 3 (pSFV/E1_{his}E2 Δ ₅₂₈₋₆₀₉) and 2 (pSFV/E1_{his}E2 Δ ₆₁₀₋₆₉₉) glycosylation sites which account for their greater mobilities (panel a, lanes 3 and 5, panel c, lane 6; Table 6.1). This was further confirmed by data obtained from endo H_f-treated samples in which deglycosylated proteins migrated in accordance with the number of amino acids deleted from each mutant (Fig 6.5, panels b and d).

In addition, Western blot analysis of the endo H_F-treated samples identified more than one species of E2 for each construct. The proteins with molecular weights greater than the major species may represent incomplete deglycosylated forms. Since strain Glasgow p7 appears to be inefficiently cleaved from E2 (see Section 4.6), the minor lower molecular weight forms probably correspond to E2 proteins alone. The predicted molecular weights of E2 in the absence of p7 are shown in Table 6.1 (column 4) and support this hypothesis.

With reference to data presented in Chapter 4 (Section 4.2), some non-covalently linked complexes between strain Glasgow E1E2 are detected, but at a significantly lower level than the amount detected for strain H77 E1E2. Prior to assessing the effect of deletion mutations in E2 on non-covalent complex formation via immunoprecipitation assays, the ability of these mutants to form non-aggregated species was analysed by non-reducing SDS-PAGE. The crude extracts derived from cells electroporated with the constructs described above were separated under non-reducing conditions and examined by Western blot analysis for E2 again using ALP98. Non-aggregated E2 protein was detected in relatively equivalent amounts for the wild type form and proteins E2_{Δ384-411}, E2_{Δ415-454}, E2_{Δ459-520}, and E2_{Δ528-609} (Fig 6.6, lanes 1-5). However, the amount of non-aggregated material detected for E2₃₈₄₋₇₀₂ and E2 expressed in the absence of E1, was reduced (compare lane 1 with lanes 7 and 8). In the case of the E2₃₈₄₋₇₀₂Gla mutant protein, a proportion of this protein may have been secreted out of the cells since the C-terminal hydrophobic region, which contains the transmembrane domain, had been removed. However, using the available immunological reagents against E2, significant levels were never detected in cell supernatants. E2_{Δ610-699} mutant protein (lane 6) was not detected by ALP98 since the epitope for the antibody is located in the deleted region.

6.1.4. Complex Formation between Wild-type E1 and E2 Deletion Mutants

Having established that the E2 deletion mutants were able to form non-aggregated molecules (Western blot analysis), the involvement of these species in non-covalently linked complexes was subsequently examined. Cells were electroporated with RNA from pSFV/E1_{his}E2 and pSFV/E1_{his}E2_Δ constructs. For control purposes RNA from pSFV/E2₃₄₀₋₈₅₁Gla and pSFV1 were also electroporated. Following a 12 h incubation period in the presence of ³⁵S-methionine, cells were lysed in immunoprecipitation buffer.

Precipitation was performed using the E2-specific MAb ALP98 for crude extracts derived from all the constructs, with the exception of pSFV/E1_{his}E2 Δ 610-699, where the R141 anti-E2 antiserum was employed (Fig 6.7, panels a and b, lane 6). Examination of the precipitated material under non-reducing conditions (panel a) revealed that non-covalently linked E2 was not clearly detected. This was probably due to the diffuse nature of the oxidised forms of E2 and the presence of non-specifically precipitated cellular proteins of similar molecular weights. A diffuse band of approximately 35 kDa (panel a, lane 1) may correspond to the oxidised form of co-precipitated E1. This 35 kDa band was absent in the samples co-precipitated with the truncated E2 protein (panel a, lane 7), E2 expressed in the absence of E1 (panel a, lane 8) and in the control sample derived from cells electroporated with RNA from pSFV1 (panel a, lane 9). Reducing gel data suggested that E1 was clearly co-precipitated with all the mutant E2 proteins (panel b, lanes 2-6). For mutant E2_{384-702G1a} protein, a reduced amount of the glycoprotein is detected which probably accounts for the coincident reduction in the abundance of E1 protein (panel b, lane 7).

Examination of duplicate crude extracts with the E1-specific AP497 MAb revealed a similar immunoprecipitation pattern. Under non-reducing conditions neither oxidised E1 nor co-precipitated wild type and mutant E2 proteins were discernible (Fig 6.8, panel a). By contrast, the reducing gel data indicated the co-precipitation of wild type E2 and all of the E2 deletion mutants with E1 (panel b, lanes 1-7). Proportionally, there appeared to be more E1 than wild type or mutant E2, suggesting that not all E1 proteins detected by AP497 were associated with E2. Among the E2 mutant proteins, the only variant which co-precipitated with E1 to a comparatively lesser extent was E2_{384-702G1a} protein which lacks the E2 transmembrane domain and p7 sequence.

Although wild-type and mutant forms of E2 could be detected under non-reducing electrophoresis conditions by Western blot analysis, the data from immunoprecipitations would indicate that the majority of E2 and associated E1 are in the form of an aggregated complex. These covalent interactions between E1 and E2 were not abolished by removal of any specific region of E2. However, there was a relative reduction in the amount of complexes detected between truncated E2_{384-702G1a} and E1. This suggests that the transmembrane domain removed from mutant E2_{384-702G1a} (and perhaps p7) may have role in the formation of complexes with E1. Other studies at the time of obtaining this data had focused on the importance of the E2 transmembrane domain in terms of ER retention (Cocquerel *et al.*, 1998; Duvet *et al.*, 1998). In light of the attention paid to this region of

E2, an identical mutant was constructed in strain H77, to reinforce the observations made for the strain Glasgow truncated E2 mutant. Moreover, E1 and E2 from strain H77 produce a greater amount of non-covalently linked complexes and thus studies could be performed on the effects of the E2 transmembrane domain on native complex formation.

6.2.1. Construction of Plasmids Expressing Strain H77 Truncated E2 and E2 Alone

The C-terminal end of strain H77 E2 and the p7/NS2 sequences were replaced with an oligonucleotide, using pGEM/E1E2_{H77} as the parent plasmid. This plasmid was linearised at a *ScaI* site at codon position 701 in E2 and the *EcoRI* site situated 3' to the ORF in the pGEM1 vector. The termini generated by these enzymes were bridged with an oligonucleotide, which introduced a stop codon, followed by a *BglIII* site and regenerated the restriction enzyme sites (Fig 6.9). This construct was called pGEM/E1E2_{384-702H77}. The *BglIII* sites flanking the ORF were employed to generate a fragment that was inserted into the pSFV1 expression vector. This procedure yielded a construct called pSFV/E1E2_{384-702H77}.

In addition, the sequences encoding strain H77 E2 and p7 proteins as well as the 5' terminal part of NS2 were introduced into the pSFV1 vector using a strategy similar to that described for the construction of pSFV/E2_{340-851Gla}. Digestion with *BamHI*, (located at codon 340 in pGEM/E1E2_{H77}) and *BglIII* (located in the pGEM1 vector) generated a fragment that encoded the E2 coding sequence, in addition to the E2 signal sequence and p7/NS2 sequences (Appendix 3). This fragment was introduced directly into the pSFV1 vector and the resulting construct was named pSFV/E2_{340-836H77}.

6.2.2. Characterisation of the Strain Glasgow and H77 E2 Truncated Mutants

Reports by Selby *et al.* (1994) and Michalak *et al.* (1997) had suggested that E2 proteins truncated at residue 715 resulted in their secretion, although mutants truncated at residue 730 were not secreted. The truncated E2 proteins engineered in this study, which ended at residue 702, were assayed for secretion into the cell supernatant by Western blot analysis

and immunoprecipitation, as well as for cell surface expression by indirect immunofluorescence.

Cells co-expressing wild type E1 and truncated E2 from constructs pSFV/E1E2₃₈₄₋₇₀₂Gla and pSFV/E1E2₃₈₄₋₇₀₂H77 were fixed with methanol and analysed with anti-E1 antiserum R528 and ALP98, the E2 MAb. Localisation of the glycoproteins was examined by confocal microscopy (Fig 6.10). The E2₃₈₄₋₇₀₂ truncated mutants for both strains Glasgow and H77 were found to co-localise with E1 (panels c and f). Using paraformaldehyde only as a fixative, expression of truncated E2 on the cell surface was not detected (data not shown). Furthermore analysis of cell supernatants for these mutants also yielded negative results, implying that reagents against E2 may not efficiently recognise secreted forms of the protein or that significant amounts of E2 were not secreted. Therefore, it can be inferred from the data obtained in this experiment that truncated forms of E2 continue to be retained within the ER. From the immunofluorescence data it was noted that the level of strain Glasgow truncated E2 detected appeared to be reduced in amount compared to strain H77 truncated E2.

Endo H_f sensitivity has been used as an assay for ER retained proteins in earlier sections (3.9 and 5.5.2), and provided confirmation of the ER retention of strain Glasgow E2₃₈₄₋₇₀₂ protein (Fig 6.5). To examine strain H77 truncated E2₃₈₄₋₇₀₂ in a similar manner, crude extracts derived from cells electroporated with RNA from the pSFV/E1E2_{H77} and pSFV/E1E2₃₈₄₋₇₀₂H77 constructs were prepared and treated with endo H_f. The treated samples were analysed in parallel with untreated crude extracts by Western blot and probed for E2 using ALP98 (Fig 6.11). Data revealed a shift in the mobilities of endo H_f-treated strain H77 wild type E2 as well as truncated E2₃₈₄₋₇₀₂ (lanes 2 and 4), when compared to untreated samples (lanes 1 and 3). Therefore, data for both strains Glasgow and H77 indicated that deglycosylation of truncated mutant E2 proteins was possible with endo H_f, and in turn that these proteins were not retrieved from the secretory pathway.

The effects of truncation of strain H77 E2 at residue 702 on complex formation was assessed by immunoprecipitation using the E2 MAb, ALP98, and AP21.010 MAb (raised against E1), in addition to the anti-E1 antiserum R528. Cells were electroporated with RNA from construct pSFV/E1E2₃₈₄₋₇₀₂H77 and for control purposes with RNAs from pSFV/E1E2_{H77} and pSFV/E2₃₄₀₋₈₃₆H77. These cells were metabolically labelled with ³⁵S-methionine and their crude extracts were prepared in immunoprecipitation buffer. The

extracts were subjected to precipitation using ALP98, and eluted material was analysed under non-reducing and reducing conditions. Under non-reducing conditions, a non-covalently linked complex was only discernible for the wild type proteins (Fig 6.12, lane 1). For pSFV/E1E2₃₈₄₋₇₀₂H77 and pSFV/E2₃₄₀₋₈₃₆H77, there was little evidence of either monomeric E1 or E2 species (lanes 2 and 3). By contrast, examination of the reducing gel indicated that E2 was precipitated from all three extracts (lanes 4-6). However, compared to wild type proteins, little E1 was co-precipitated with truncated E2₃₈₄₋₇₀₂, thus suggesting that E1 was not even incorporated efficiently into the covalently associated aggregated material formed by this mutant (compare lanes 4 and 5). Data from reducing gels and previous Western blot analysis (Fig 6.11) hint that the level of truncated E2 detected for strain H77 does not appear to be significantly different from the amount detected for wild type protein, but there was a reduction in the amount of non-aggregated E2₃₈₄₋₇₀₂ (panel a, lane 2). In addition, a marked reduction in the amount of non-aggregated E2 expressed in the absence of E1 was also observed for strains Glasgow (Fig 6.7, panel a, lane 7) and H77 (Fig 6.12, lane 3). This correlates with data obtained in Chapter 5 (Section 5.3.3), where slight changes in E1, in terms of cysteine substitutions, affected the level of non-aggregated E2 detected. It appears therefore that the presence of wild type E1 is necessary for the formation increased production of non-aggregated E2.

To further investigate the behaviour of strain H77 E1 expressed along with truncated E2 with respect to complex formation, immunoprecipitations were performed using immune reagents directed against E1. Crude extracts were derived from cells electroporated with RNA from pSFV/E1E2₃₈₄₋₇₀₂H77, and from pSFV/E1E2_{H77} and pSFV/E1_{H77} for controls. Analysis of the precipitated material using the R528 antiserum and MAb AP21.010 revealed that a non-covalently linked complex between E1 and E2 was detected for the wild type proteins (Fig 6.13, panels a and b, lanes 2 and 5). Non-aggregated E1 was observed under the non-reducing polyacrylamide electrophoresis conditions for the other two constructs (Fig 6.13., panels a and b, lanes 1 and 3), but truncated E2 was not as clearly distinguished under either non-reducing or reducing conditions (panels a and b, lanes 3 and 6). Instead, in addition to monomeric E1, higher molecular weight species were precipitated from the sample derived from cells expressing pSFV/E1E2₃₈₄₋₇₀₂ with the anti-E1 antibodies (panels a and b, lane 6). Interestingly these higher molecular weight species co-migrated with oligomeric forms of E1 (panels a and b, compare lanes 4 and 6), which predominantly occur when E1 is synthesised in the absence of E2 (see Fig 5.4). It is likely that any truncated E2 which co-precipitated with E1 would co-migrate with the presumed

dimeric (x2) form of E1 (panels a and b, lane 6). These data, from two different immunological reagents therefore suggest that although E1 is able to oligomerise in the absence of E2, it behaves similarly when made in the presence of a truncated form of E2 lacking the E2 transmembrane domain. Furthermore, in the absence of E2, E1 tends to form increased amounts of differentially glycosylated species and does not attain an oxidised state. In the presence of truncated E2₃₈₄₋₇₀₂, the differentially glycosylated forms of E1 are apparent but occur to a lesser extent than with E1 expressed on its own. Also, the diffuse material representing monomeric E1 may include some oxidised E1 that corresponds to a population of E1 molecules which can interact with truncated E2. However, the substantial amount of E1 which oligomerises in the presence of truncated E2 suggests that the transmembrane domain of E2 has a role in the folding of E1.

6.3.1. Use of Foreign Glycoprotein Sequences to Study the Properties of E2

Due to the aberrant behaviour of E1, in terms of formation of oligomers and the loss of non-covalently associated complexes with truncated E2 lacking the transmembrane domain, the influence of the E2 transmembrane domain (TM) on the properties of E1 was examined further. The first strategy tested the affect of substituting foreign glycoprotein transmembrane domains for the E2 TM. In parallel studies sequences containing the E2 TM were fused to the ectodomain of a foreign glycoprotein to determine whether E1 could associate with the resultant chimeric protein.

6.3.2.1. Construction Of Chimeric Proteins - Cloning Of Full Length HSV-1 Glycoprotein D

In order to examine whether the E2 TM had a specific role in determining the behaviour of E1 described in section 6.1.6, this region was substituted with two different foreign transmembrane domains. The first of these was derived from HSV-1 gD. Published data had shown that gD is efficiently transported to the cell surface in the absence of any other HSV-1 proteins (Long *et al.*, 1992). The second transmembrane sequence employed was derived from the rubella virus E1 glycoprotein. By contrast to gD, this glycoprotein is retained in the ER and sequences within its transmembrane domain have been implicated in conferring this property (Hobman *et al.*, 1997).

As an initial cloning step, a DNA fragment containing the coding sequences for full length gD protein was amplified and cloned into pGEM1. A plasmid called pAT153_{BamHIJ}, that contained a region of the HSV-1 genome which spans nucleotides 136290 and 142746 was used as the template for PCR amplification (nomenclature as in McGeoch *et al.*, 1988a). The primers (1 and 2; Fig 6.14) used were designed to introduce *HindIII* and *BglIII* sites at the 5'-terminus and *EcoRI* and *BglIII* sites at the 3'-terminus in the resulting PCR product. Digestion of this product with *HindIII* and *EcoRI* allowed the gD coding sequences to be cloned into the pGEM1 vector, which was subsequently termed pGEM/gD.

6.3.2.2. Construction of E2 Chimeric Proteins

To create the first of the E2 chimeric proteins, the predicted E2 TM coding sequence was substituted by inserting a PCR fragment encoding the gD TM and cytoplasmic tail (Fig 6.15, panel a). A PCR product was generated from pGEM/gD template plasmid using a forward primer (spanning nucleotides 139436-139454) that incorporated an in-frame *ScaI* site and codon 702 of E2 at the 5'-terminus; Primer 2 (shown in Fig 6.14) was used as the reverse primer, thus maintaining the stop codon, *BglIII* and *EcoRI* sites. The PCR fragment, after treatment with *ScaI* and *EcoRI* restriction enzymes, was introduced into the pGEM/E1E2_{H77} plasmid linearised with the same enzymes. The *ScaI* site in this plasmid was positioned at codon 701 in the polyprotein and the *EcoRI* site was located in the pGEM1 vector downstream of the HCV coding sequences. The resulting plasmid was named pGEM/E1E2_{702gDtm}, which signified the last codon of E2 (aa 702) directly followed by residues representing the gD TM and cytoplasmic tail. The equivalent SFV expression plasmid was termed pSFV/ E1E2_{702gDtm}.

Oligonucleotides were employed to substitute the E2 transmembrane domain with that of rubella virus E1 glycoprotein (RVE1) (Fig 6.15, panel b). The RVE1 transmembrane domain is comparatively well defined and is composed of a stretch of 35 amino acid residues (Hobman *et al.*, 1997). Therefore, the oligonucleotides encoded the sequence for the RVE1 transmembrane domain followed by an artificial stop codon. Suitable restriction enzyme sites were also included in the design of the oligonucleotides to allow their introduction at the *ScaI* site at residue 701 and the *EcoRI* site in pGEM/E1E2_{H77}. Following insertion of the oligonucleotides the construct was called pGEM/E1E2_{701RVE1tm}.

The *BglIII* restriction sites flanking the coding sequences were used once again to introduce the ORF into pSFV1, which was subsequently called pSFV/E1E2_{701RVE1tm}; this plasmid encoded E2 from aa 384 to 701 followed by the rubella virus E1 transmembrane domain.

6.3.2.3. Construction of gD Chimeric Proteins

The strategy employed placed the HSV-1 gD coding sequence downstream of and in-frame with the HCV E1 coding sequences. This enabled E1 and gD to be expressed as a single polyprotein, a normal feature of E1 and E2 protein synthesis. In order to prevent disturbing the functions of the E1 C-terminal sequences, which act both as the E1 transmembrane domain and E2 signal sequence, the coding sequence for this region in E1 was maintained. Therefore, the segment encoding the gD signal sequence and the N-terminus (aa -24 to +10, see Appendix 4) was omitted from the E1-gD chimera. A *HindIII-SalI* (at codon position 381 in E1) fragment derived from pGEM/E1E2_{H77} and an *EcoRI-PvuII* (at residue 29 in gD) fragment obtained from the pGEM/gD were ligated into *HindIII-EcoRI* linearised pGEM1 vector using an oligonucleotide to bridge between the *SalI* and *PvuII* sites (Fig 6.16, panel a). The oligonucleotide encoded residues in E1 and E2 that spanned the proteolytic cleavage site between E1 and E2 and a portion of the N-terminal region of gD, from residues 11 to 29. These residues contain the epitope for a gD-specific monoclonal antibody, 4846, which had been fully characterised (A. Cross and H. S. Marsden, unpublished results). The resulting construct was termed pGEM/E1₃₈₇gD₁₁₋₃₇₀.

The transmembrane and cytoplasmic domain of gD, in the construct pGEM/E1₃₈₇gD₁₁₋₃₇₀, was substituted with that of a region of the HCV polyprotein encoding the E2 TM coding sequence in order to study its properties in the absence of the E2 ectodomain (Fig 6.16, panel b). Due to the length of the E2 TM, the sequence encoding this region was amplified by PCR using primers that introduced an *EaeI* at the 5' terminus of the amplified product. In addition, *BglIII* and *EcoRI* sites were incorporated at the 3' terminus by the second primer, along with an artificial stop codon. The PCR product encoded residues 703 to 758 of the HCV polyprotein and thus extended a short distance (13 residues) beyond the cleavage site between E2 and p7. It was presumed that cleavage by host cell signalases would occur at this site, although this need not necessarily be the case. Therefore, in the following section, this segment will be referred to as the E2TM with the caveat that

chimeric proteins may contain also this short segment of p7. The PCR fragment was inserted at the *EaeI* site at codon 310 in gD, and *EcoRI* site within the pGEM1 vector of plasmid pGEM/E1₃₈₇gD₁₁₋₃₇₀. This construct was called pGEM/E1₃₈₇gD₁₁₋₃₁₂E2_{tm}.

For control purposes the gD TM was substituted with the rubella virus E1 TM, and details for the design of this construct are shown in Fig 6.16 (panel c). An oligonucleotide encoding the RVE1 TM sequence was introduced at the *EaeI* and *EcoRI* sites in the plasmid pGEM/E1₃₈₇gD₁₁₋₃₇₀, resulting in pGEM/E1₃₈₇gD₁₁₋₃₁₂RVE1_{tm}. For all the constructs, *BglII* fragments were transferred into pSFV1, and the resultant plasmids were called pSFV/E1₃₈₇gD₁₁₋₃₇₀, pSFV/E1₃₈₇gD₁₁₋₃₁₂E2_{tm} and pSFV/E1₃₈₇gD₁₁₋₃₁₂RVE1_{tm}.

The constructs made in section 6.3.2.2 and 6.3.2.3 are referred to as the E2/gD chimera plasmids, and are shown diagrammatically in Fig 6.17.

6.3.3.1. Cellular Distribution of Chimeric Proteins - Indirect Immunofluorescence

To examine the effect of transmembrane domain substitutions on E2 and gD glycoproteins, the localisation of these proteins, along with E1, was investigated using indirect immunofluorescence combined with confocal microscopy. Cells electroporated with RNA from pSFV/E1E2_{H77}, pSFV/E1E2₇₀₂gD_{tm} and pSFV/E1E2₇₀₁RVE1_{tm} were analysed by immunofluorescence using R528 anti-E1 antiserum and ALP98, the E2-specific MAb. Confocal images of these cells revealed the localisation of E2₇₀₁RVE1_{tm} was identical to that for E1 (Fig 6.18, compare panels g-i); an identical result was obtained for the wild type E1 and E2 proteins (panel a-c). For E2₇₀₂gD_{tm}, the bulk of the detected staining coincided with that for E1, although some staining at the periphery of the cell was also found which did not overlap with E1 (compare panels d and e). This additional staining for E2₇₀₂gD_{tm} is apparent in the merged image of panels d and e (panel f).

A similar assay was performed for the gD chimeric constructs, where cells were electroporated with RNA from plasmids pSFV/E1₃₈₇gD₁₁₋₃₇₀, pSFV/E1₃₈₇gD₁₁₋₃₁₂E2_{tm} and pSFV/E1₃₈₇gD₁₁₋₃₁₂RVE1_{tm}. Immunofluorescence analysis was performed with the same anti-E1 antiserum (R528), but gD was detected using 4846 MAb. The confocal images obtained indicated that gD₁₁₋₃₇₀ expressed downstream of E1 in the polyprotein (Fig 6.19,

panel b), had a distinctive staining pattern at the cell periphery which was not apparent for co-expressed E1 (panels a and c). Such peripheral localisation of gD was observed on replacing the gD transmembrane and cytoplasmic domains with the transmembrane domain from RVE1, although the amount of staining at the cell periphery was comparatively reduced (compare panels b and h). However, fusion of the HCV E2 TM to gD gave a staining pattern for gD which was identical to that for HCV E1 (panels d-f) and there was little evidence of localisation at the cell periphery (compare panels b and e).

The above evidence suggested that the peripheral staining observed for E2 and gD chimeric proteins may represent protein localisation on the cell surface. This was examined further by comparing immunofluorescence patterns in permeabilised and non-permeabilised cells. Non-permeabilised cells were prepared by fixing with paraformaldehyde in the absence of any detergent (PFA). For analysis of permeabilised cells, fixation was carried out using a combination of paraformaldehyde and Triton X-100 (PFAT). Cells electroporated with RNA from construct pSFV/E1E2_{702gDtm}, in addition to pSFV/E1E2_{701RVE1tm} and pSFV/E1E2_{H77} constructs for control purposes, were fixed using the above described fixatives and processed for immunofluorescence. The confocal images derived from these cells revealed that cells expressing all of the above constructs, when fixed with PFAT, stained for both E1 and E2 protein that localised intracellularly (Fig 6.20., panels b, d and f). However, in the case of PFA-treated cells (panels a, c and e) intracellular staining was absent for E1 indicating that these cells were not permeabilised. Thus, any fluorescence that was observed for cells fixed in this manner would represent proteins that were present on the cell surface. No cell surface expression of wild type E2 was observed in cells expressing proteins from construct pSFV/E1E2_{H77} (panel a). Weak fluorescence was observed for cells expressing E2_{702gDtm} (panel c), suggesting that the MAb ALP98 was able to detect some but barely detectable cell surface localisation of this chimeric E2 protein. There did appear also to be some cell surface expression of E2_{701RVE1tm}, where E2 was fused to the rubella virus E1 TM (panel e). This data suggest that the HSV-1 gD and RVE1 transmembrane domains may not completely retain E2 in the ER.

For similar studies on cell surface expression of gD expressed in the context of E1, cells were electroporated with RNA transcripts from constructs pSFV/E1_{387gD}₁₁₋₃₇₀ and pSFV/E1_{387gD}_{11-312E2tm} and fixed with either PFA or PFAT. In this instance, gD was visualised using MAb 4846. The PFA-treated cells expressing gD₁₁₋₃₇₀ readily exhibited

the presence of this protein on the cell surface (Fig 6.21, panel a) and in the parallel culture fixed with PFAT, there was additional intracellular localisation of the protein (panel b). Investigation of the cells expressing construct pSFV/E1₃₈₇gD₁₁₋₃₁₂E2_{tm} fixed with PFAT revealed intracellular localisation for the chimeric protein, gD₁₁₋₃₁₂E2_{tm} (panel d). However, in PFA-fixed cells, only a small amount of gD fused to the E2 transmembrane domain was evident on the cell surface (panel c). The level of immunofluorescence was significantly less than that detected for gD₁₁₋₃₇₀, which contained its authentic transmembrane domain.

6.3.3.2. Endo H_f and PNGase F Sensitivity of Chimeric Proteins

Further analysis was conducted to investigate the cell surface expression seen for the gD/E2 chimeric proteins and gD₁₁₋₃₇₀, when expressed downstream of E1, with the use of two endoglycosidase enzyme digestion assays, endo H_f and PNGase F. The predicted molecular weights of the nascent and fully glycosylated chimeric proteins are stated in Table 6.2, where the typical molecular weight of complex type N-linked oligosaccharides present on mature VSV G protein is taken as ~3.3 kDa (Kornfeld & Kornfeld, 1985).

For endoglycosidase assays, cells were electroporated with RNA transcribed from pSFV/E1E2_{H77}, pSFV/E1E2_{702gDtm}, pSFV/E1 E2_{701RVE1tm}, pSFV/E1₃₈₇gD₁₁₋₃₇₀ and pSFV/E1₃₈₇gD₁₁₋₃₁₂E2_{tm}. Portions of the crude extracts derived from these cells were treated with endo H_f and PNGase F, and separated by polyacrylamide gel electrophoresis alongside corresponding untreated samples. Western blot analysis using ALP98 was performed on samples obtained from cells expressing wild type E2, E2_{702gDtm} and E2_{701RVE1tm}. The data revealed that the predominant species for wild-type E2 was sensitive to endo H_f digestion (Fig. 6.22, panel a, lanes 1-3). A higher molecular weight species (indicated as ◆) also was sensitive to the enzyme and was assumed to be a glycosylated form of uncleaved E1E2 (see Fig 3.19). For E2_{702gDtm}, the predominant species was digested with endo H_f, however a proportion of a higher molecular weight band remained resistant to endo H_f but not PNGase F (indicated as ■, panel a, lanes 4-6). This material in lane 4 may represent both glycosylated, uncleaved E1E2_{702gDtm} and some E2_{702gDtm} which is modified by transit beyond the ER. Such further modified protein would be resistant to endo H_f and it was noted from the immunofluorescence analysis that some of the E2_{702gDtm} protein was detected on the cell surface (Fig. 6.20, panel c). For E2_{701RVE1tm}, endo H_f treatment gave a smear of material (panel a, lanes 7-9) which was digested by the enzyme.

By contrast, PNGase F digestion gave a single band. The smeared material could represent a complex series of molecules which have varying extents of endo H_f -sensitivity. Partially-resistant forms could occur through modification of glycans as a result of cycling between the ER and Golgi apparatus and again, it was noted that very small amounts of E2_{701RVE1tm} were present on the cell surface (Fig. 6.20, panel e). A higher molecular weight species (indicated as \odot , panel a, lane 7) was sensitive to endo H_f and is therefore likely to be glycosylated, uncleaved E1E2_{701RVE1tm}. The slight increase in mobilities of PNGase F treated proteins is likely to result from the presence of single saccharide residues at each of the glycosylation sites in endo H_f -treated samples (panel a compare lanes 2, 5 and 8 with lanes 3, 6 and 9); these saccharides are removed in PNGase F digests.

Western blot analysis of samples derived from cells expressing E1_{387gD}₁₁₋₃₇₀ and E1_{387gD}_{11-312E2tm}, was performed using 4846 MAb. Results revealed the presence of two gD₁₁₋₃₇₀-specific bands (Fig 6.22, panel b, lane 13). The lower molecular weight band probably corresponds to intracellular gD₁₁₋₃₇₀ protein. The higher molecular weight species on the other hand is likely to represent the cell surface gD₁₁₋₃₇₀ species, which contains complex type oligosaccharides that increase the molecular weight of the proteins. This observation is consistent with observations made by Long et al. (1992) in their studies characterising HSV-1 gD. After endo H_f treatment, proteins represented by the upper band and some of the lower band remained resistant to deglycosylation with this enzyme (panel b, lane 14); this confirmed the notion that these proteins contain oligosaccharides modified in compartments other than the ER. The lower molecular weight proteins resistant to endo H_f may correspond to Golgi apparatus-localised species containing hybrid-type oligosaccharides. However, the production of a novel lower molecular weight band following endo H_f treatment revealed that a proportion of gD₁₁₋₃₇₀ proteins did localise in the ER. Treatment with PNGase F (panel b, lane 15) gave a single prominent band. The same blot shows the result for gD_{11-312E2tm}, containing the E2 transmembrane domain. Upon treatment with endo H_f , the major species detected in the untreated sample was reduced in size to the same apparent molecular weight as in the PNGase F-treated sample (panel b, lanes 11 and 12), suggesting that the bulk of this protein remained in the ER. There were two minor bands (panel b, lane 10, indicated by *), which migrate faster after endo H_f treatment, but which almost disappear after PNGase F treatment. These bands may represent forms of gD_{11-312E2tm}, which have been further processed such that they contain partially-resistant oligosaccharide moieties. Although Western blot analyses data cannot be quantitatively interpreted, it was apparent that any cell surface species as

identified by resistance to endo H_f for E2_{702gDtm}, E2_{701RVE1tm} and gD_{11-312E2tm}, were significantly less prominent compared to the equivalent gD₁₁₋₃₇₀ species (compare lanes 4, 7, and 10 with lane 13).

The behaviour of E1 expressed from the chimera constructs used above was also analysed by endo H_f treatment. Endo H_f -treated samples were analysed again alongside corresponding untreated crude extracts using AP21.010 MAb to probe for E1 (Fig 6.23). As observed previously, E1 was sensitive to endo H_f digestion. Furthermore, a distinct band representing a protein of approximately 65 kDa (marked by *) corresponded to a species detected by the gD-specific MAb 4846 (Fig 6.22, panel b, lane 10), which was sensitive to endo H_f . Given that both E1 and gD-specific antibodies detected this species, it is likely to represent the uncleaved precursor polyprotein of E1₃₈₇gD_{11-312E2tm}, which has been glycosylated.

Collectively, the data thus far have indicated that in contrast to wild type E2, a portion of chimeric E2 proteins carrying the foreign transmembrane segments is transported beyond the ER and these forms are resistant to endo H_f -deglycosylation but remain sensitive to PNGase F. Additionally, gD containing its own TM is localised in intracellular organelles as well as at the cell surface. However, the bulk of chimeric gD fused to the E2 TM is rendered sensitive to endo H_f , and is thus an indication that this protein is retained in the ER by virtue of the HCV sequences.

6.3.4. Effect of Foreign Ecto- and Transmembrane Domains on E1E2 Complex Formation

The principle aim in generating these chimeric proteins was to study the influence of sequences containing the E2 transmembrane domain on E1E2 complex formation and their influence on the behaviour of E1. Therefore, immunoprecipitation analysis was performed using various immunological reagents available. Cells were electroporated with RNA from constructs expressing wild type and chimeric proteins and pSFV1. Following radiolabelling with ³⁵S-methionine, crude extracts of these cells were prepared in immunoprecipitation buffer. Immunoprecipitations were performed initially using the E2-specific MAb, ALP98, for the E2-containing extracts, and the gD-specific MAb, 4846, for the gD-containing extracts. The precipitated material was examined under both non-

reducing and reducing conditions to enable non-covalently linked complexes to be distinguished from aggregated material. From the non-reducing gel, native E1E2 complexes were identified for wild-type forms of the glycoproteins, in agreement with previous data (Fig 6.24, panel a, lane 1). However, on replacing the E2 TM with either gD or RVE1 TM segments, E1 was not present and chimeric E2 proteins were either absent or present in small amounts under such electrophoretic conditions (panel a, lanes 2 and 3). Examination of these samples under reducing conditions indicated that complex formation did occur between E1 and the chimeric E2 proteins (Fig 6.24, panel b, lanes 1-3). However, these complexes are likely to represent covalently-linked aggregates and contained reduced amounts of E1. Therefore, substitution of the E2 transmembrane domain with equivalent domains from foreign glycoproteins does not prevent complex formation but non-covalently linked E1 and E2 is not readily detected. In particular, substitution with the RVE1 TM containing an ER retention signal was not sufficient to compensate for the authentic E2 transmembrane domain which has similar retention properties.

Immunoprecipitation with 4846 MAb did not reveal any significant co-precipitation of E1 with gD₁₁₋₃₇₀ protein (Fig 6.24, panel a, lane 5). A small amount of E1 was observed to precipitate with gD₁₁₋₃₇₀, however this was only detected under reducing conditions, and thus may represent aggregated material (panel b, lane 5). E1 did co-precipitate with gD containing the E2 TM and this was observed under reducing and non-reducing conditions (panels a and b, lane 4). This suggests that E1 is non-covalently bound to the gD/E2 chimeric protein. It was observed that increased levels of E1 precipitated with the gD/E2 chimeric protein as compared to wild type E2 protein however this may be a consequence of the different avidities of ALP98 and 4846 antibodies for E2 and gD respectively. A higher molecular weight band of ~65 kDa represented the E1₃₈₇gD₁₁₋₃₁₂E2_{tm} precursor polyprotein (panel a and b, lanes 4, indicated by *), and a species of ~55 kDa present in pSFV/E1₃₈₇gD₁₁₋₃₇₀ (panel b, lane 5) was presumably gD protein that had been further processed beyond the ER (Fig 6.22, lane 13).

In parallel, identical ALP98 and 4846 immunoprecipitation samples from non-radiolabelled cell extracts were examined by Western blot analysis using R528 to probe for E1 (Fig 6.25). This analysis, performed on precipitated samples subjected to non-reducing gel electrophoresis, yielded a similar result to that obtained for radiolabelled precipitates. Hence, non-covalently-linked E1 was only precipitated by wild type E2 (lane 1) and the

gD₁₁₋₃₁₂E2_{tm} chimera with the gD ectodomain fused to the E2 TM (lane 4). This result confirms that the E2 transmembrane domain, when fused to gD allows the formation of a non-covalently linked complex with E1.

Complex formation was investigated also using reagents against E1, for which crude extracts were derived from cells expressing strain H77 E1 alone, wild type E1 and E2 proteins, and the chimeric E2/gD proteins. Analysis with anti-E1 antiserum, R528 (Fig 6.26) revealed that, under non-reducing conditions, the differentially glycosylated forms of E1 were detected preferentially when E1 was expressed in the absence of E2 (panel a, lane 1); this is in agreement with data presented in Chapter 5. Non-covalently-linked complexes, which occur between the wild type E1 and E2 proteins (panel a, lane 2), were observed also under non-reducing conditions. In this case, only one major glycosylated E1 species was detected which corresponded to the oxidised form of the protein according to its migration relative to E1 expressed in the absence of E2. Barely detectable levels of monomeric E1 were found when it was expressed in the context of E2_{702gDtm}, E2_{701RVE1tm} and gD₁₁₋₃₇₀ chimeric proteins (panel a, lanes 4, 5 and 6, respectively). Also co-precipitation of non-aggregated E2_{702gDtm}, E2_{701RVE1tm} and gD₁₁₋₃₇₀ was not observed. However, a prominent band representing E1 expressed along with gD₁₁₋₃₁₂E2_{tm} fusion protein (panel a, lane 3) was detected in approximately similar amounts to that for E1 expressed along with wild type E2 (panel a, compare lanes 2 and 3). Comparing the migratory characteristics of this band with E1 expressed in the presence and absence of E2, it would appear to be an oxidised form of E1. Two other novel proteins co-precipitated with E1 expressed in the context of gD₁₁₋₃₁₂E2_{tm} had apparent molecular weights of approximately 45 and 65 kDa. Western blot analysis using 4846 antibody of duplicate immunoprecipitations showed that the 45 kDa protein was gD containing the E2 TM and the 65 kDa protein, corresponded in size to uncleaved but glycosylated E1-gD₁₁₋₃₁₂E2_{tm} precursor protein (Fig 6.27, lane 5).

In an earlier section (6.1.6), oligomers of E1 were identified under reducing electrophoretic conditions on expression with E2 truncated at residue 702. To test whether oligomers could be detected with the E2 and gD chimeric proteins, immunoprecipitations with two anti-E1 reagents, R528 and AP497, were analysed under reducing conditions. Results revealed that both the monoclonal antibody and the rabbit antiserum were able to precipitate oligomeric forms of E1 expressed along with E2_{702gDtm}, E2_{701RVE1tm} and gD₁₁₋₃₇₀ (Fig 6.26, panel b, lanes 4-6; Fig 6.28, lanes 4-6), which migrated with the E1 oligomers

produced in absence of E2 (Fig 6.26, panel b, lane 1 and Fig 6.28, lane 1). Additional protein bands (marked by ■ and ⊙) are predicted to correspond to E2_{702gDtm} and E2_{701RVE1tm} proteins, respectively, that are precipitated in the aggregated material (Fig 6.26, panel b, lanes 4 and 5; Fig 6.28, lanes 4 and 5). Significantly, oligomeric E1 was not precipitated from extracts containing wild type E2 and gD_{11-312E2tm} (Fig 6.26, panel b, lanes 2 and 3; Fig 6.28, lanes 2 and 3). Additional species observed in lanes 2 and 3 with apparent molecular weights of ~97 kDa and ~65 kDa (indicated by ◆ and *) represent glycosylated forms of the precursor E1E2_{H77} and E1_{387gD}_{11-312E2tm} polyproteins, respectively (see Section 6.3.3.2). These results provide substantial evidence for a critical role for sequences containing the E2 TM in complex formation with E1 and the ability of E1 to form oligomers.

Further conclusive evidence of the specific influence sequences containing of the E2 TM on chimeric gD was provided from the examination of gD chimeric protein fused with the RVE1 TM. Cells were electroporated with RNA from constructs pSFV/E1_{387gD}_{11-312E2tm}, pSFV/E1_{387gD}_{11-312RVE1tm} and pSFV/E1_{H77}. ³⁵S-methionine-labelled crude extracts were prepared from these cells and subjected to immunoprecipitation using R528, the anti-E1 antiserum. Analysis of the isolated material under non-reducing conditions showed that although E1 could be precipitated from all three crude extracts, gD_{11-312E2tm} protein but not gD_{11-312RVE1tm} protein co-precipitation was detected (Fig 6.29, panel a, lanes 1 and 2). E1 precipitated by R528 from the pSFV/E1_{387gD}_{RVE1tm} extract had the same pattern of multiple species as E1 expressed alone whereas E1 isolated from pSFV/E1_{387gD}_{11-312E2tm} extract migrated as a smear of oxidised species (Fig 6.29, panel a, lanes 1-3). Therefore, oxidation of E1 is dependent on sequences containing the E2 TM and cannot be substituted by other TM sequences with retention properties. Consistent with this conclusion, oligomeric forms of E1 were detected on expression with gD fused to the rubella virus E1 TM (panel b, lanes 1-3) under reducing conditions. An additional observation was that the deglycosylated forms of E1 were less evident on co-precipitation of gD with either the E2 or rubella virus E1 TM sequences (panel b, compare lanes 1 and 2 with lane 3). This may indicate that the multiple monomeric forms of E1 are a feature of E1 expression in the absence of any other glycoprotein.

6.4. Discussion

An ER retention signal has been mapped to the C-terminal 29 amino acid residues of E2 which constitutes the hydrophobic transmembrane domain (Cocquerel *et al.*, 1998). E2 proteins from strain H truncated at residues 661, 669, 688, 704, 710 and 715, expressed by recombinant vaccinia viruses, were detected in HepG2 cellular extracts as well as in the supernatant of infected cells (Selby *et al.*, 1994; Michalak *et al.*, 1997). However, significantly reduced levels of truncation mutants E2₃₇₀₋₆₈₈, E2₃₇₀₋₇₀₄, and E2₃₇₀₋₇₁₅ were detected in the supernatant when expressed by recombinant Sindbis virus (Michalak *et al.*, 1997). Similarly in this study, expression of strains Glasgow and H77 E2 truncated at residue 702 using SFV, another alphavirus, led to intracellular retention of this mutant rather than secretion. It is unlikely that retention is solely a consequence of misfolding of E2 since most forms of E2 secreted by the vaccinia virus system are apparently aggregated (Michalak *et al.*, 1997). Therefore, it is possible that viral expression systems can have different effects on the ability of cells to secrete proteins. Sindbis virus and SFV expression systems might inhibit secretion of some proteins, or alternatively, the vaccinia virus system facilitates secretion. In this study, SFV-expressed gD, carrying its authentic transmembrane domain was efficiently transported through the secretory pathway to the cell surface. This provides evidence that SFV does not prevent movement of proteins from the ER. The level of protein expression achieved by vaccinia virus appears to be significantly higher than that achieved by recombinant Sindbis virus or SFV (Michalak *et al.*, 1997; from this study data not presented). Therefore, the over-expression of glycoproteins by recombinant vaccinia virus perhaps imparts a greater strain on the ER quality control mechanisms that result in the escape of misfolded proteins from the ER.

Evidence for the existence of an ER-retention signal at the C-terminus of E2 was provided by expression of chimeric proteins in which E2 ecto- and transmembrane domains were exchanged with corresponding domains of a protein normally transported to the plasma membrane. E2₃₇₁₋₇₁₇ fused to the CD4 anchor and cytoplasmic domain led to the expression of this E2 ectodomain on the cell surface (Cocquerel *et al.*, 1998). In a separate study, E2 ectodomain (aa 371-742) fused to the transmembrane domain and cytoplasmic tail of VSV G was also transported to the cell surface (Lagging *et al.*, 1998). Here, the

fusion of HSV-1 gD and rubella virus E1 transmembrane and cytoplasmic domains to E2₃₈₄₋₇₀₂ led to the transport of only a small amount of chimeric E2 to the plasma membrane. The bulk of this chimeric protein was retained in the ER in a manner similar to truncated E2 (at aa 702) as shown by confocal microscopy and endo H_f sensitivity. No previous studies have analysed the ER-retention properties of the HSV-1 gD transmembrane and cytoplasmic domain. In the case of the rubella virus E1 transmembrane domain, it efficiently retains the rubella virus E1 protein in the ER (Hobman *et al.*, 1997). However, when associated to ectodomains of foreign glycoproteins this signal was less effective and fusion proteins did leave the ER (Hobman *et al.*, 1997). Further studies revealed that sequences upstream of the rubella virus E1 transmembrane domain influenced the efficiency of ER retention by this domain (Hobman *et al.*, 1997). The region containing the E2 transmembrane domain did efficiently sequester gD in the ER, indicative of an ER-retention signal. This agrees with an earlier study in which an ER-retention signal was mapped to the C-terminal 29 amino acids of E2 that constitutes the hydrophobic transmembrane domain (Cocquerel *et al.*, 1998). However, there may be some contribution from ectodomain sequences in E2 since small amounts of gD fused to the segment containing the E2 transmembrane domain are present on the cell surface.

Data with strain Glasgow E2 deletion mutants failed to identify any regions in the protein whose removal prevented aggregated complex formation with E1. As mentioned previously for this strain, the abundance of native complex is barely detectable with the available antibodies. The overall conclusion therefore is that removing certain regions of E2 which may be involved in aggregate formation does not enhance production of native complex. In fact, it may be the case that any alteration to E2 promotes aggregation. Interestingly, analysis of secreted E2 truncated mutants revealed that apart from E2₃₄₀₋₆₆₁ the remaining mutants did not migrate into non-reducing gels, which suggested that they formed aggregates (Michalak *et al.*, 1997). This was confirmed by immunoprecipitations studies using a conformation-sensitive MAbs H2, which failed to recognise secreted E2₃₇₁₋₇₁₅ but efficiently precipitated secreted E2₃₇₁₋₆₆₁. These findings indicate that the nature of truncated E2 in functional studies (e.g. receptor binding; Pileri *et al.*, 1998) has to be evaluated carefully and reconfirmed using E1E2 complexes.

Fusion of a transmembrane domain from foreign glycoproteins (gD and rubella virus E1) to the ectodomain of E2 did not assist the formation of a native complex for either strain H77 or strain Glasgow. This was despite the presence of an ER retention signal in the

rubella virus E1 transmembrane domain to anchor chimeric E2 to the ER. This suggests that retention of E2 in the ER *per se* was not sufficient to permit native complex formation. Pseudotype virus particles have been produced using chimeric HCV structural proteins (Lagging *et al.*, 1998), in which the transmembrane domains of E1 and E2 have been substituted by the equivalent region of VSV G, a viral glycoprotein which readily forms virus-like particles in the absence of other viral proteins (Rolls *et al.*, 1994). From observations in this study using E2 fused to the transmembrane domain of foreign glycoproteins, caution should perhaps be applied to the use of the such pseudotype virus particles for infectivity of mammalian cells and neutralisation studies (Lagging *et al.*, 1998) since any E1E2 complexes are unlikely to be in a native form.

The amount of E1 associated with E2 truncated at residue 702 was reduced, indicative of a role for the C-terminal region of E2 in complex formation; this agrees with previous studies (Selby *et al.*, 1994; Cocquerel *et al.*, 1998). Examination of the behaviour of E1 in the presence of truncated E2 revealed that E1 oligomerised in a manner similar to that in the absence of E2. E1 oligomerisation was observed also when E1 was made along with chimeric E2 molecules, consisting of the E2 ectodomain fused to either the gD or rubella virus E1 transmembrane and cytoplasmic domains, and HSV-1 gD. Oligomerisation was one of the criteria established in Chapter 5 to enable discrimination of E1 folding in the presence and absence of E2. In addition, oxidised forms of E1 were not precipitated in appreciable amounts with these chimeras. A significant finding in this study was that the oligomerisation of E1 was not detected when a foreign glycoprotein (gD) ectodomain was fused to sequences containing the E2 transmembrane domain (aa 703-758). An E1gD_{E2tm} complex was immunoprecipitated using both anti-gD and anti-E1 antibodies and the complexed proteins migrated into non-reducing gels. In addition, E1 protein was in an oxidised state and there was little evidence of E1 oligomerisation products. This suggests that chimeric gD, containing the E2 TM, was able to form a native-like complex with E1. Therefore, in relation to the necessity for E2 for proper folding of E1 and native complex formation, a region spanning the C-terminal end of E2 and p7, which includes the E2 transmembrane domain appears critical for these two events. Previous studies have suggested that E2 performs a chaperone function to aid E1 folding (Section 1.3.11.1). The data presented here imply that a region containing the E2 transmembrane domain is important for this process while the ectodomain can be substituted by a foreign sequence. There may be additional important interactions between E1 and the E2 ectodomain, however, these are not sufficient to prevent E1 misfolding. Further studies are required to

determine the limits of the region in E2 (and perhaps p7) that allow oxidation of E1. Assuming that the E2 transmembrane domain is important in E1 oxidation, it would be predicted that interaction would occur with a hydrophobic region in E1. There are two hydrophobic segments in E1 (Fig 3.8) and either of them may be involved in E2 association.

Strain Glasgow proteins	No. of aa deleted	No. of potential N-glycan sites lost (aa position)	Predicted MW of nascent E2 (kDa)	Predicted MW of nascent E2 + p7 (kDa)	Predicted MW of modified E2 (~kDa)*	Predicted MW of modified E2 + p7 (~kDa)*
E2 ₃₈₄₋₇₄₅	-	-	40.0	47.1	58.0	65.1
E2 _{Δ384-411}	26	0	37.3	44.4	55.3	62.4
E2 _{Δ415-454}	38	4 (417,423, 430,448)	35.7	42.8	45.7	52.8
E2 _{Δ459-520}	60	0	33.5	40.6	51.5	58.6
E2 _{Δ528-609}	80	3 (540,556, 576)	31.2	38.3	43.2	50.3
E2 _{Δ610-699}	88	2 (623, 645)	29.6	36.7	43.6	50.7
E2 ₃₈₄₋₇₀₂	43	0	35.0	-	53.0	-

Table 6.1. Predicted molecular weights (MW) of nascent and glycosylated forms of E2 deletion mutants. * It is assumed that addition of a high mannose oligosaccharide increased the molecular weight of nascent protein by ~2.0 kDa (Kornfeld & Kornfeld, 1985).

Protein	Predicted MW of nascent protein (kDa)	Predicted MW with ER retained protein (~kDa)*	Predicted MW with transported protein (~kDa)[#]
E2 ₃₈₄₋₈₁₀ (H77 wt E2-p7)	47.2	69.2	83.5
E2 _{702gDtm}	41.3	63.3	77.6
E2 _{701RVE1tm}	38.8	60.8	75.1
gD ₁₁₋₃₇₀	39.9	45.9	49.8
gD _{11-312E2tm}	39.6	45.6	49.5
gD _{11-312RVE1tm}	37.7	43.7	47.6

Table 6.2. Predicted molecular weights (MW) for nascent and glycosylated forms of HCV E2, HSV-1 gD and chimeric proteins. * It is assumed that addition of a high mannose oligosaccharide increased the molecular weight of nascent protein by ~2.0 kDa (Kornfeld & Kornfeld, 1985). [#] It is assumed that addition of complex oligosaccharides increased the molecular weight of nascent protein by 3.3 kDa (Kornfeld & Kornfeld, 1985).

CTG *GTC* *GAC* *GTG* *CAG* *TAC* *TTG* *TAC* *GGG* *GTG* *ACC* *TGA* *TCT* *AGA* *TCT* *AAG* *CTT*
 L V D V Q T L Y G V T *
 607 610 699 851

Fig 6.2. Construction of strain Glasgow E2 internal deletion Δ 610-699. An *EcoRI-SalI*_{c609} fragment derived from pGEM/E1_{his}E2 and pGEM1 plasmid linearised using *EcoRI* and *HindIII* restriction enzymes were ligated along with a PCR fragment that was prepared by digestion with *SalI* and *HindIII* restriction enzymes. The PCR fragment contained the coding sequence for the C-terminus of E2, p7 and the N-terminal part of NS2 from residues 699-851. The PCR fragment included an in frame 5' *SalI* site (at codon 608) and a stop codon (*) following residue 851. This construct was called pGEM/E1_{his}E2 Δ 610-699. Restriction sites are shown in italics, E2 amino acid residues are shown in green while the residue in black represents NS2.

GTG *CAG* *TAC* *TTG* *TAG* *AGA* *TCT* *AAG* *CTT* GG
 V Q Y L *
 699 702

Fig 6.3. Construction of a truncated form of strain Glasgow E2. An *EcoRI-Scal*_{c701} fragment derived from pGEM/E1_{his}E2 and *HindIII* and *EcoRI* linearised pGEM1 plasmid were ligated along with an oligonucleotide (over-lined in red), which encoded sequences adjacent to the *Scal* site (nucleotide position 2443; Appendix 1) and introduced a stop codon (*) after codon 702 in E2. This construct was called pGEM/E1E2_{384-702Gla}. The restriction sites are shown in italics and E2 amino acid residues are shown in green.

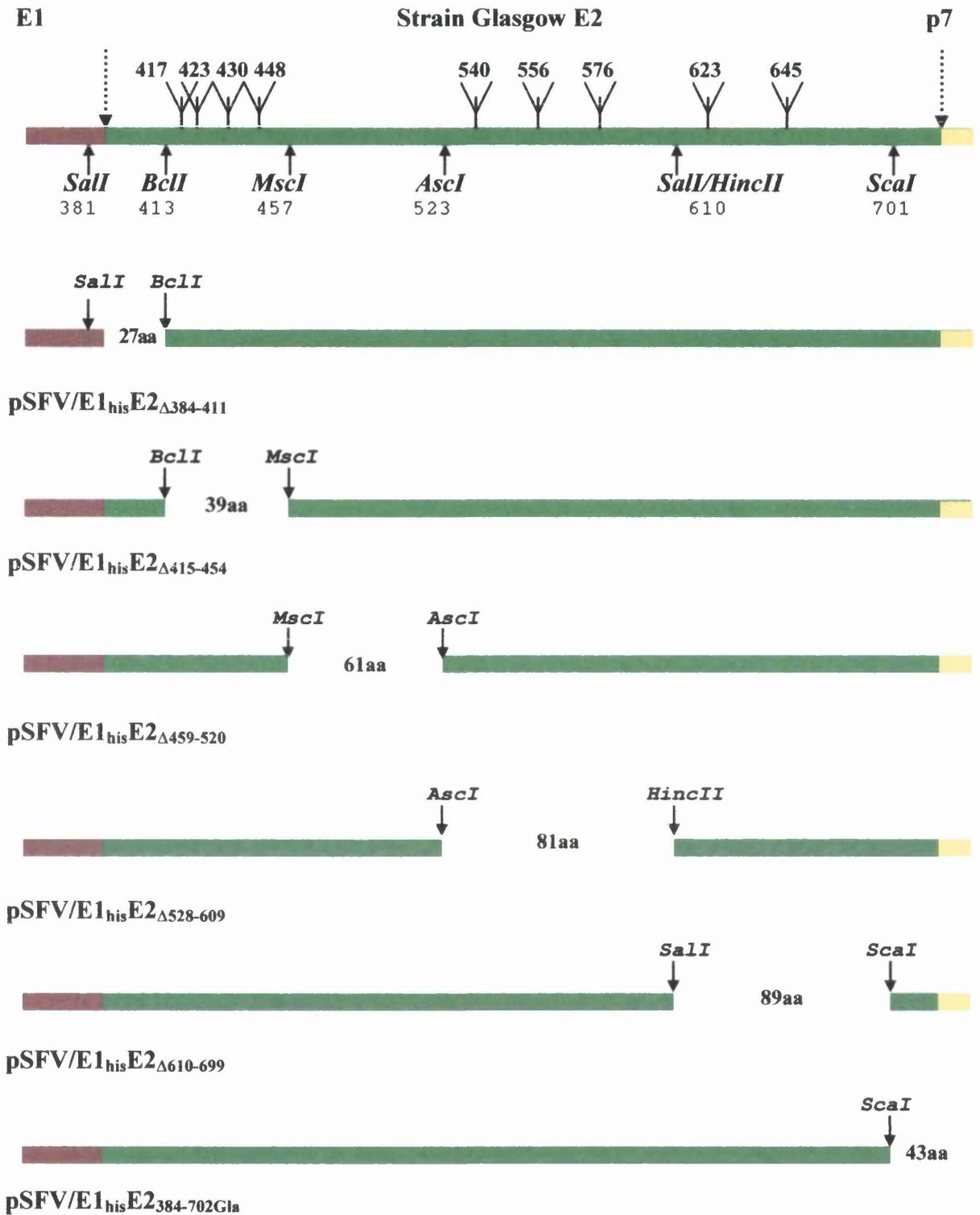


Fig 6.4. A schematic representation of the amino acid deletions created in strain Glasgow E2 glycoprotein. Restriction enzyme sites used in the construction of each mutant are shown, as are the numbers of residues removed in each case. Location of predicted glycosylation sites are indicated. Downward arrows indicate the predicted cleavage sites between E1-E2 and E2-p7. The brown block represents E1, green represents E2 and yellow indicates p7.

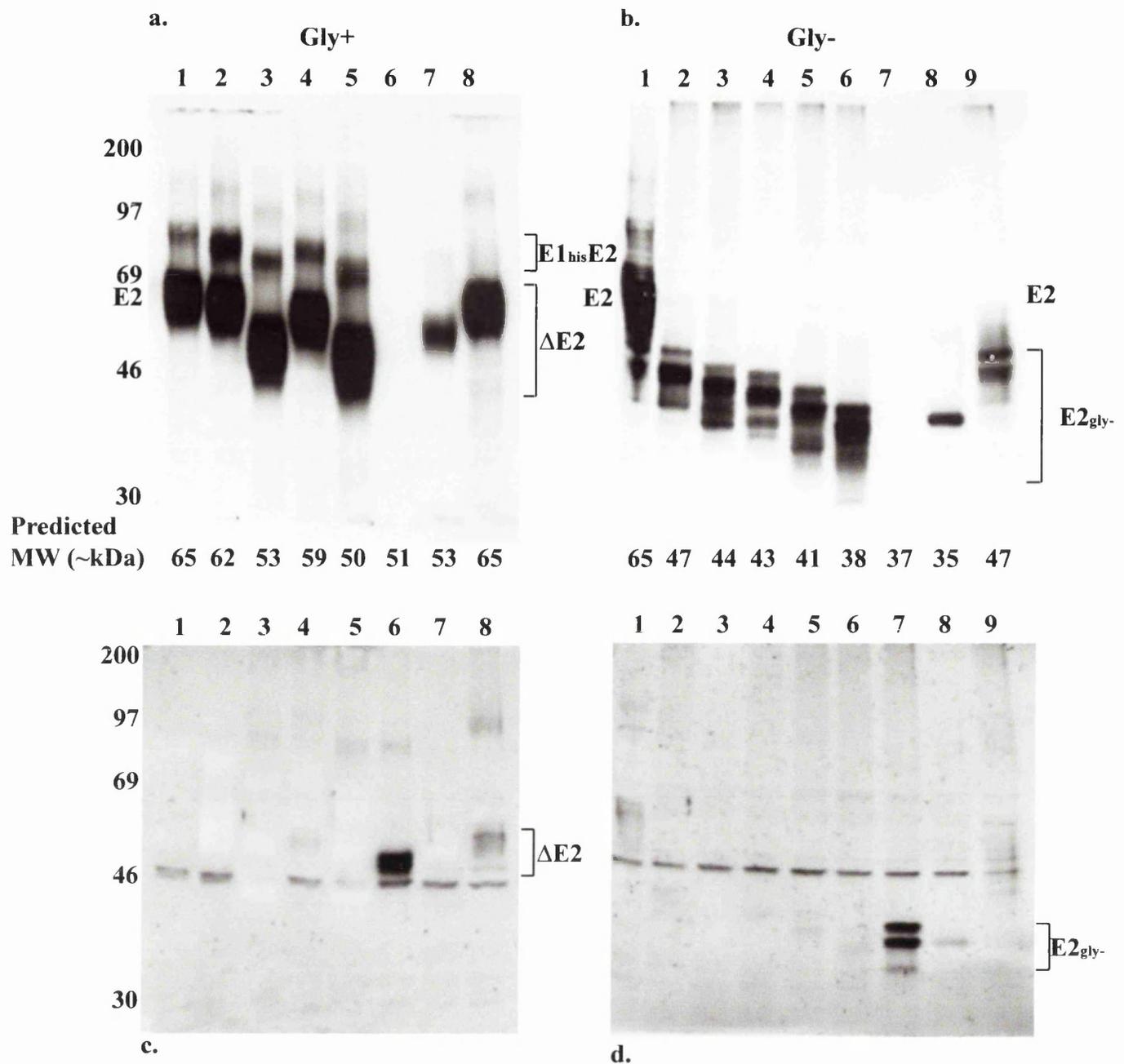


Fig 6.5. Sensitivity of strain Glasgow E2 deletion mutants to endo H_f glycosidase. Cells were electroporated with RNA from pSFV strain Glasgow and incubated at 37°C for 12 h. These cells were lysed and a portion of crude cell extract was endo H_f-treated (panels b and d). An equal amount of untreated crude extract was analysed in parallel (panels a and c) by electrophoresis on 10% reducing, polyacrylamide gels. After transfer of proteins to membranes, Western blot analysis was conducted initially using E2-specific MAb, ALP98, (panels a and b). Subsequently, blots were stripped and reprobed with anti-E2 antiserum, R141, (panels c and d). Samples were pSFV/E1_{his}E2 (panel a, lane 1; panel b, lanes 1 and 2), pSFV/E1_{his}E2_{Δ384-411} (panel a, lane 2; panel b, lane 3), pSFV/E1_{his}E2_{Δ415-454} (panel a, lane 3; panel b, lane 4), pSFV/E1_{his}E2_{Δ459-520} (panel a, lane 4; panel b, lane 5), pSFV/E1_{his}E2_{Δ528-609} (panel a, lane 5; panel b, lane 6), pSFV/E1_{his}E2_{Δ610-699} (panel c, lane 6; panel d, lane 7), pSFV/E1_{his}E2₃₈₄₋₇₀₂G1a (panel a, lane 7; panel b, lane 8), and pSFV/E2₃₄₀₋₈₅₁G1a (panel a, lane 8; panel b, lane 9). Lane 1 in panels b and d show the untreated full-length E2 control samples. The predicted molecular weights (MW) of the wild type and mutant proteins are indicated below each corresponding lane; these molecular weights assume that p7 is inefficiently cleaved from E2. The positions of precursor polyprotein E1_{his}E2, E2, mutant E2 (ΔE2) and deglycosylated (gly-) proteins are indicated.

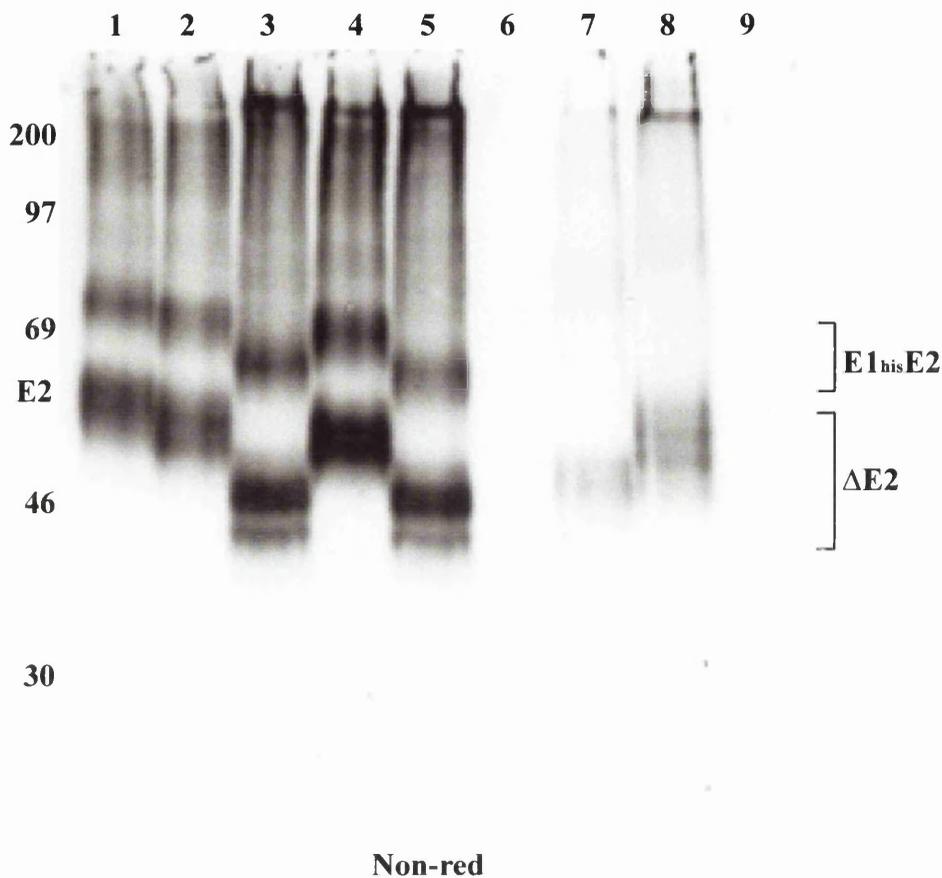


Fig 6.6. Analysis of strain Glasgow E2 deletion mutants under non-reducing electrophoresis conditions. Experimental conditions were as described in Fig 6.5. except the crude extracts were analysed by non-reducing electrophoresis on a 10% polyacrylamide gel. Following transfer of proteins to membrane, Western blot analysis was performed using E2-specific MAb ALP98. Crude extracts were derived from cells electroporated with RNA from pSFV/E1_{his}E2 (lane 1), pSFV/E1_{his}E2 $_{\Delta 384-411}$ (lane 2), pSFV/E1_{his}E2 $_{\Delta 415-454}$ (lane 3), pSFV/E1_{his}E2 $_{\Delta 459-520}$ (lane 4), pSFV/E1_{his}E2 $_{\Delta 528-609}$ (lane 5), pSFV/E1_{his}E2 $_{\Delta 610-699}$ (lane 6), pSFV/E1_{his}E2 $_{384-702Gla}$ (lane 7), pSFV/E2 $_{340-851Gla}$ (lane 8), pSFV1 (lane 9). The positions of E1_{his}E2 precursor polyproteins, E2 and E2 mutants (Δ E2) proteins are shown.

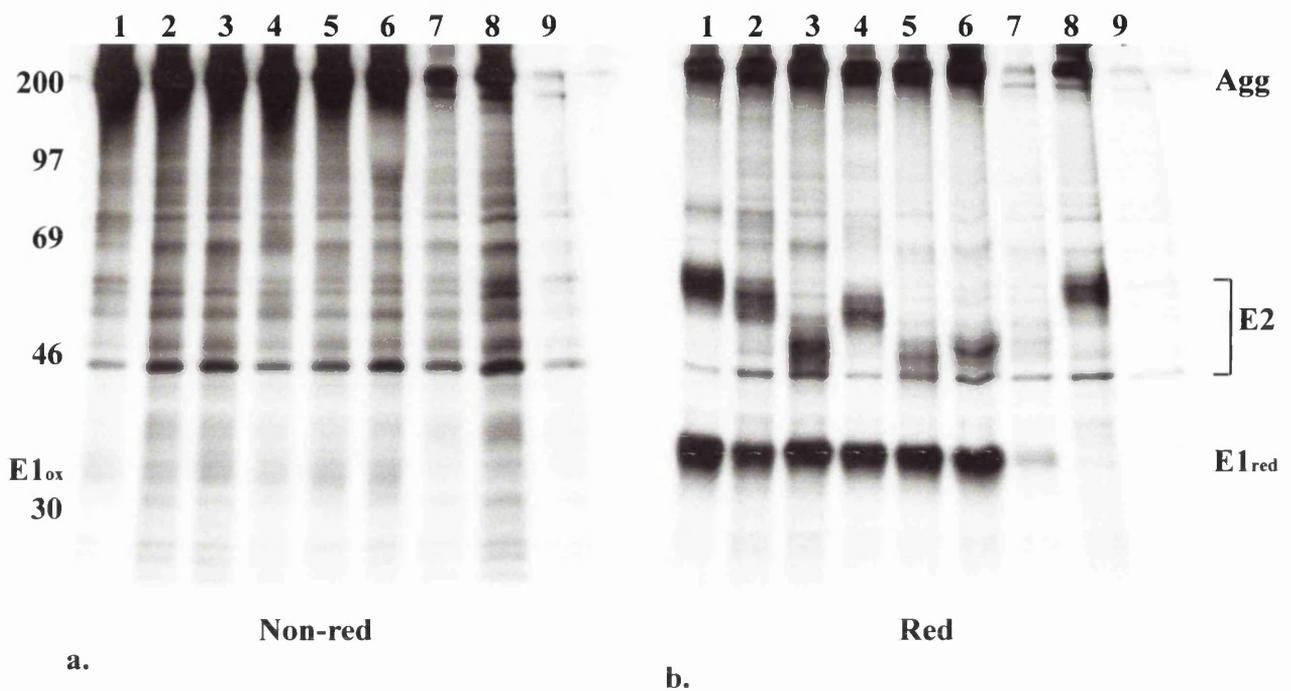


Fig 6.7. Complex formation between E2 deletion mutants and E1 protein. Cells were electroporated with the RNA from pSFV strain Glasgow constructs and labelled from 4-12 h after electroporation in the presence of ^{35}S -methionine. Cells were lysed and all crude extracts were subjected to immunoprecipitation using E2-specific antibody ALP98, except anti-E2 antiserum R141 was used for extract derived from pSFV/E1_{his}E2 Δ ₆₁₀₋₆₉₉-electroporated cells (lanes 6). Precipitated material was analysed under non-reducing (panel a) and reducing conditions (panel b) on 10% polyacrylamide gels. Samples were pSFV/E1_{his}E2 (lanes 1), pSFV/E1_{his}E2 Δ ₃₈₄₋₄₁₁ (lanes 2), pSFV/E1_{his}E2 Δ ₄₁₅₋₄₅₄ (lanes 3), pSFV/E1_{his}E2 Δ ₄₅₉₋₅₂₀ (lanes 4), pSFV/E1_{his}E2 Δ ₅₂₈₋₆₀₉ (lanes 5), pSFV/E1_{his}E2 Δ ₆₁₀₋₆₉₉ (lanes 6), pSFV/E1_{his}E2 Δ ₃₈₄₋₇₀₂G1a (lanes 7), pSFV/E2 Δ ₃₄₀₋₈₅₁G1a (lanes 8), pSFV1 (lanes 9). The positions of the aggregated material (Agg), oxidised (ox) and reduced (red) E1 and E2 proteins are indicated.

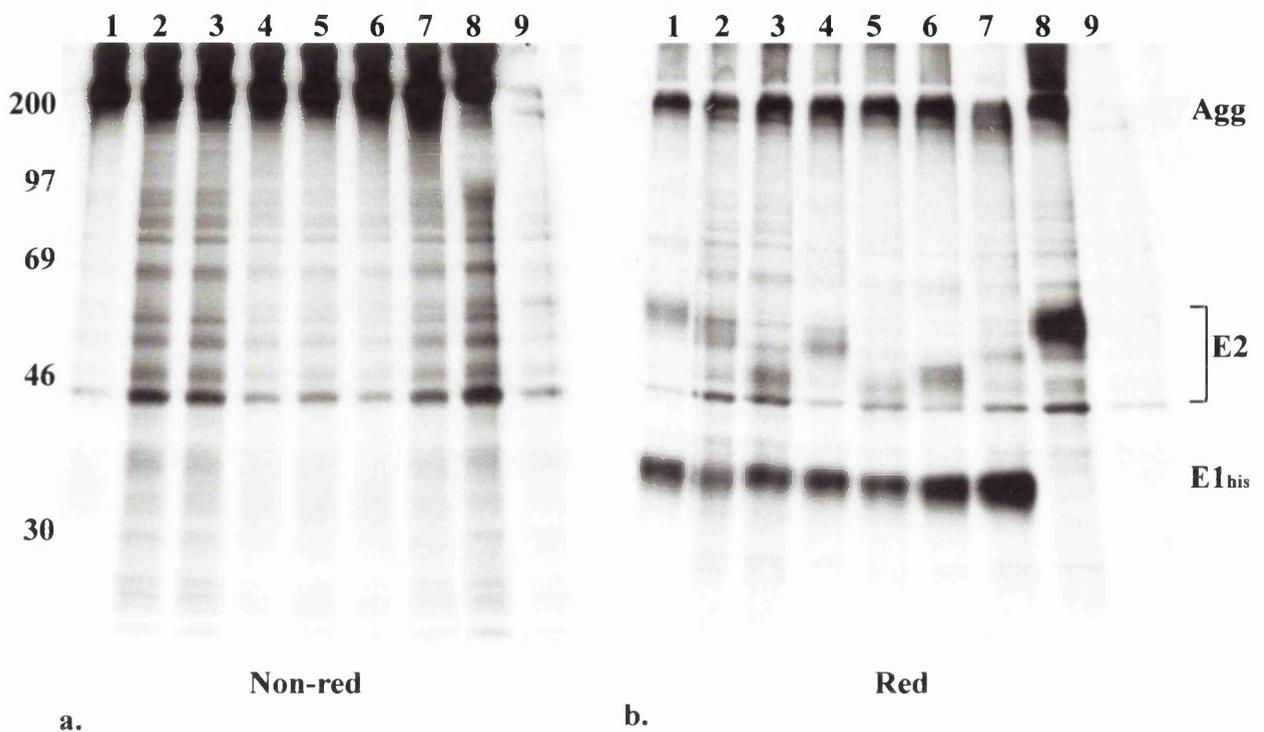


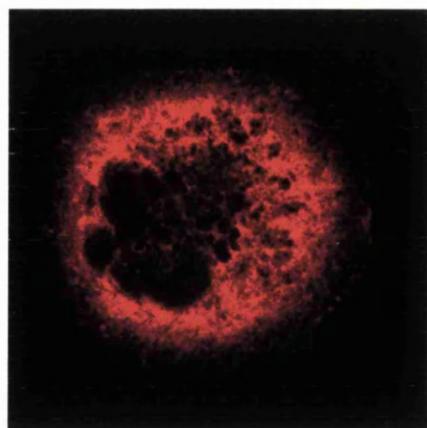
Fig 6.8. Complex formation between E2 deletion mutants and E1 protein. Experimental conditions were as for those described for Fig 6.7 except immunoprecipitations on crude extracts were performed using E1-specific MAb AP497. The precipitated samples were analysed under non-reducing (panel a) and reducing (panel b) conditions on 10% polyacrylamide gels. Samples were pSFV/E1_{his}E2 (lanes 1), pSFV/E1_{his}E2 Δ ₃₈₄₋₄₁₁ (lanes 2), pSFV/E1_{his}E2 Δ ₄₁₅₋₄₅₄ (lanes 3), pSFV/E1_{his}E2 Δ ₄₅₉₋₅₂₀ (lanes 4), pSFV/E1_{his}E2 Δ ₅₂₈₋₆₀₉ (lanes 5), pSFV/E1_{his}E2 Δ ₆₁₀₋₆₉₉ (lanes 6), pSFV/E1_{his}E2₃₈₄₋₇₀₂G1a (lanes 7), and pSFV1 (lanes 9). The pSFV/E2₃₄₀₋₈₅₁G1a (lanes 8) cell extract was precipitated with ALP98. The bands corresponding to the aggregated material (Agg), E1_{his} and E2 proteins are labelled.

—————
 GTG CAG TAC TTG TAG AGA TCT GAC AGT GAA TTC CGG
 V Q Y L *
 699 702

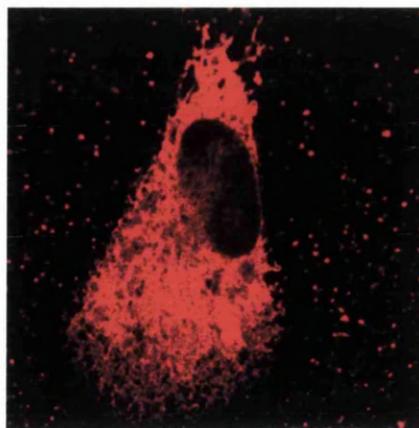
Fig 6.9. Schematic illustrating the construction of the truncated form of strain H77 E2. pGEM/E1E2_{H77} (see Fig 4.11) was linearised at *ScaI* (at codon 701) and *EcoRI* (positioned 3' of the ORF within the pGEM1 plasmid). This fragment which encoded E1 and the ectodomain of E2 was ligated using an oligonucleotide (over-lined in red) which bridged between the *ScaI* and *EcoRI* sites. The oligonucleotide encoded sequences adjacent to the *ScaI* site (nucleotide position 2442; Appendix 3) and a stop codon (*) following residue 702. This construct was called pGEM/E1E2_{384-702H77}.

Glasgow E1E2₇₀₂

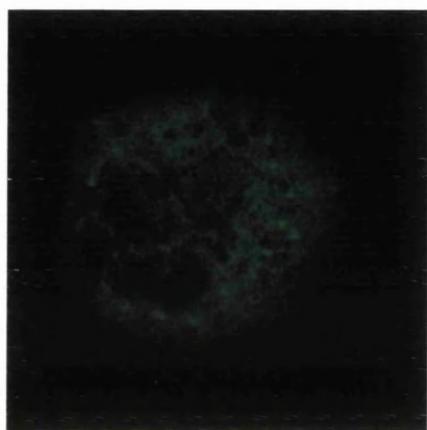
H77 E1E2₇₀₂



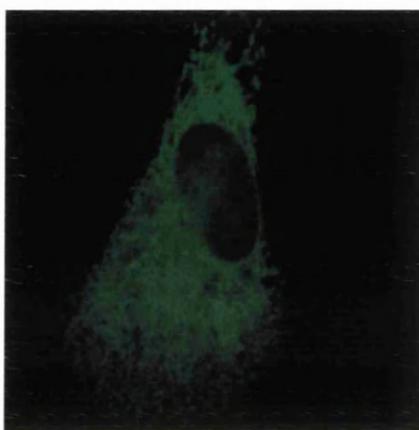
a.



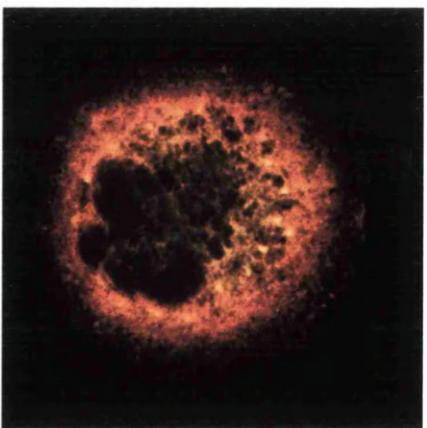
d.



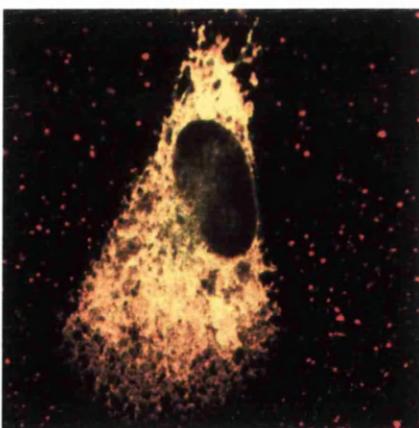
b.



e.



c.



f.

Fig 6.10. Localisation of E1 and truncated E2 glycoproteins of strains Glasgow and H77. Cells were electroporated with RNA from pSFV/E1_{his}E2₃₈₄₋₇₀₂Gla (panels a-c) and pSFV/E1E2₃₈₄₋₇₀₂H77 (panels d-f) and incubated for approximately 12 h at 37°C. Methanol-fixed cells were analysed using anti-E1 antiserum, R528, (panels a and d) and E2-specific MAb, ALP98, (panels b and e). Panels c and f represent the merged images of panels a and b, and d and e, respectively.

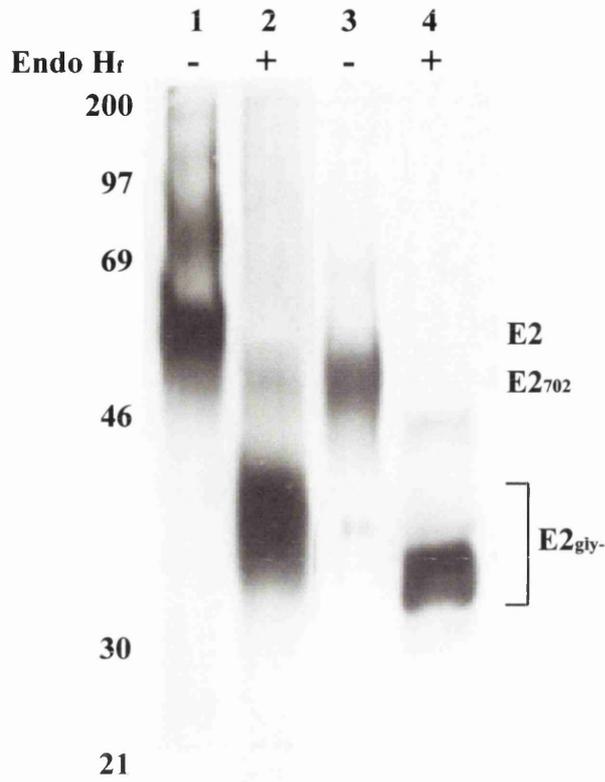


Fig 6.11. Endo H_f sensitivity of strain H77 truncated and full-length E2 proteins. Cells were electroporated with RNA from pSFV strain H77 constructs. A portion of crude extracts was treated with endo H_f (lanes 2 and 4) and analysed alongside untreated crude extracts (lanes 1 and 3) on a 10% polyacrylamide gel. Proteins were visualised by Western blot analysis using E2-specific MAb ALP98. Samples were pSFV/E1E2_{H77} (lanes 1 and 2) and pSFV/E1E2_{384-702H77} (lanes 3 and 4). Positions of E2, truncated E2 (E2₇₀₂) and deglycosylated E2 (E2_{gly-}) proteins are shown.

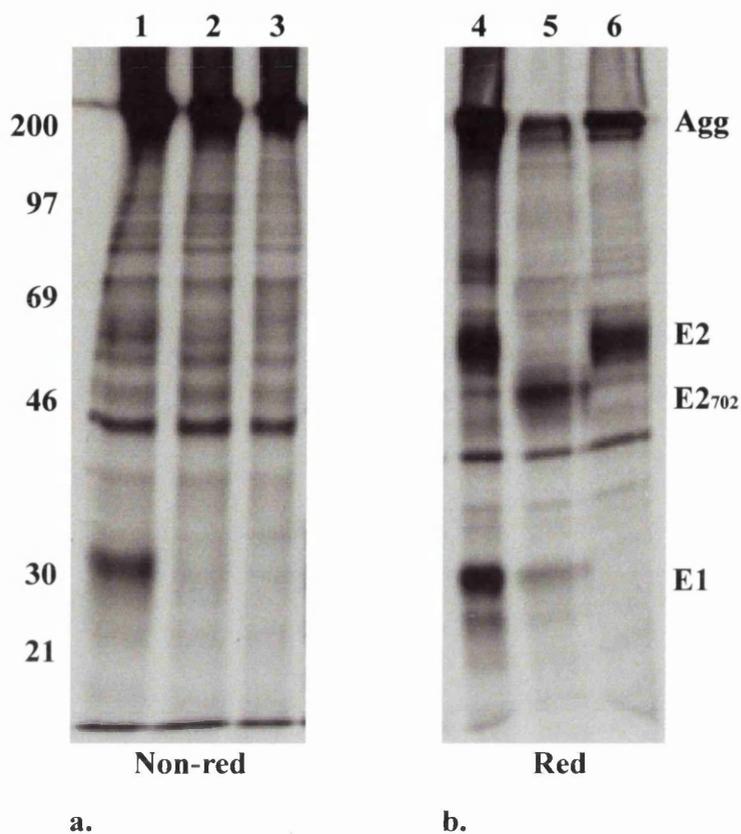


Fig 6.12. Examination of complex formation between strain H77 truncated E2 and E1 proteins. Cells were electroporated with RNA and ³⁵S-labelled between 4-12 h after electroporation. Crude extracts derived from these cells were subjected to immunoprecipitation with the E2-specific MAb, ALP98. Precipitated material was analysed on 10% polyacrylamide gels under non-reducing (panel a) and reducing (panel b) conditions. Samples were pSFV/E1E2_{H77} (lanes 1 and 4), pSFV/E1E2_{384-702H77} (lanes 2 and 5), and pSFV/E2_{340-836H77} (lanes 3 and 6). The bands corresponding to E2 and truncated E2 (E2₇₀₂) are labelled, as are aggregated material (Agg) and E1.

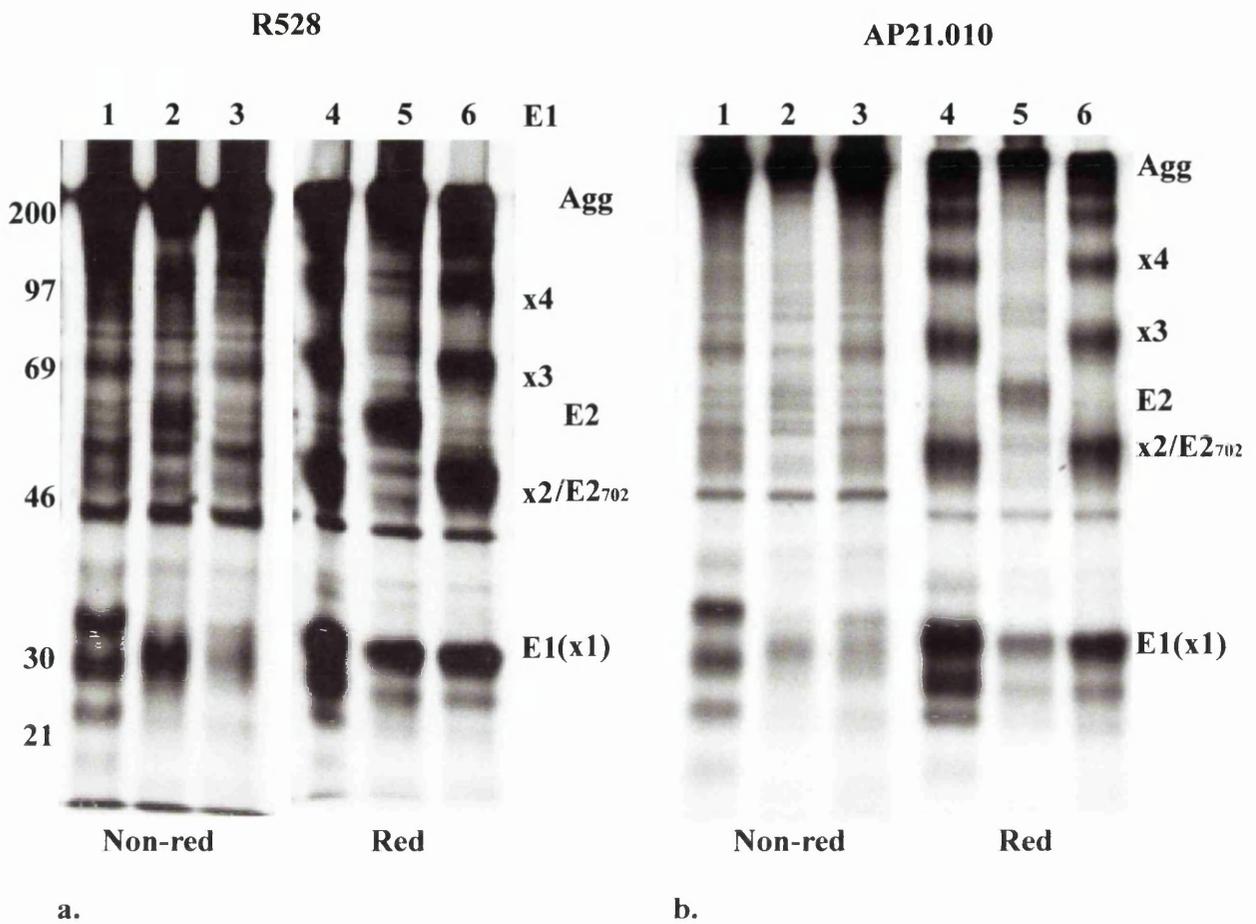


Fig 6.13. Examination of the behaviour of E1 expressed along with truncated E2 using E1-specific antiserum R528 and E1-specific MAb AP21.010. Experimental conditions were as described in Fig 6.12. except immunoprecipitations were performed using E1-specific antiserum R528 (panel a) and MAb AP21.010 (panel b). The precipitated material was analysed under non-reducing (Non-red) and reducing (Red) electrophoresis conditions on 10% polyacrylamide gels. Samples were pSFV/E1_{H77} (lanes 1 and 4), pSFV/E1E2_{H77} (lanes 2 and 5), and pSFV/E1E2_{384-702H77} (lanes 3 and 6). The bands corresponding to E2, truncated E2 (E2₇₀₂) and aggregated material (Agg) are labelled, as are the oligomeric forms of E1.

a.

E2 gD

GTG	CAG	TAC	TTG	ATG	GGC	CTG	ATC	GCC	...	CCC	TTG	TTT	TAC	TAG	AGA	TCT	GAA
V	Q	Y	L	M	G	L	I	A		P	L	F	Y	*			
699			702	316												370	

b.

E2 RVE1

GTG	CAG	TAC	TTG	TGG	AAT	CTG	ACA	TTA	GGG	GCC	ATT	TGT	GCT	CTG	CCC	CTA	GTG	GGT	TTG
V	Q	Y	I	W	N	L	T	L	G	A	I	C	A	L	P	L	V	G	L
699		701	1																17

TTA	GCA	TGC	TGC	GCC	AAG	TGT	TTG	TAT	TAC	CTT	CGC	GGA	GCG	ATT	GCC	CCT	AGG	TAG	AGA
L	A	C	C	A	K	C	L	Y	Y	L	R	G	A	I	A	P	R	*	
18																		35	

TCT GAA TTC CGG

Fig 6.15. Fusion of the HSV-1 gD and rubella virus E1 transmembrane domains to the ectodomain of E2.

a. pGEM/E1E2_{H77} was linearised at *ScaI* (at codon 701, nucleotide position 2442; Appendix 3) and *EcoRI* (located 3' of the ORF within pGEM1 plasmid) to generate a fragment coding for E1 and the ectodomain of E2. This fragment was ligated along with a PCR fragment (over-lined in red) which bridged the *ScaI* and *EcoRI* sites. The PCR product was generated using pGEM/gD plasmid as a template and coded for the gD transmembrane domain (residues 316-370; Appendix 4) as well as the authentic gD stop codon. This construct was called pGEM/E1E2_{702gDtm}.

b. The *ScaI-EcoRI* fragment from pGEM/E1E2_{H77} described in part a was ligated along with an oligonucleotide (over-lined in red) which bridged the *ScaI* and *EcoRI* sites. The oligonucleotide encoded the rubella virus E1 transmembrane domain and a stop codon. This construct was called pGEM/E1E2_{701RVE1tm}.

Restriction sites are shown in italics while amino acid sequences in green, lime and light blue represent HCV E2, HSV-1 gD and rubella virus E1 respectively.

Fig 6.16. Construction of HSV-1 gD chimeric proteins.

a. A *HindIII-Sall*_{c381} fragment derived from pGEM/E1E2_{H77} (see Fig 4.11) and *PvuII*_{c29}-*EcoRI* fragment from pGEM/gD (see Fig 6.14 and Appendix 4) were ligated along with an oligonucleotide (over-lined in red) and pGEM1 linearised with *HindIII* and *EcoRI*. The oligonucleotide bridged the *Sall* (nucleotide residue 1482; Appendix 3) and *PvuII* (nucleotide position 138574; Appendix 4) sites and encoded residues which span the E1-E2 proteolytic cleavage site and residues 11-28 of gD. This construct was called pGEM/E1₃₈₇gD₁₁₋₃₇₀.

b. pGEM/E1₃₈₇gD₁₁₋₃₇₀ plasmid was partially digested with *EaeI* restriction enzyme (a site at codon 310 in gD; nucleotide position 139419; Appendix 4) and *EcoRI* (which lies downstream of the ORF within the pGEM1 vector) to generate a linear fragment containing sequences for E1 and the ectodomain of gD. This fragment was ligated with a PCR fragment (over-lined in red) that was produced using pGEM/E1E2_{H77} plasmid as a template. The PCR fragment bridged the *EaeI* and *EcoRI* sites and coded for the C-terminus of E2 and N-terminus of p7 (residues 703-758) followed by a stop codon. This construct was called pGEM/E1₃₈₇gD₁₁₋₃₁₂E2_{tm}.

c. The *EaeI-EcoRI* fragment generated from pGEM/E1₃₈₇gD₁₁₋₃₇₀ in part b was ligated with an oligonucleotide (over-lined in red), which bridged the *EaeI* and *EcoRI* sites. The oligonucleotide encoded the rubella virus E1 transmembrane domain followed by a stop codon. This construct was called pGEM/E1₃₈₇gD₁₁₋₃₁₂RVE1_{tm}.

Restriction sites are in italics, while amino acid sequences in brown, green, yellow, lime and light blue represent E1, E2, p7, gD and rubella virus E1 respectively.

Fig 6.16. a.

E2 gD

GCC GGC **GTC GAC** GCG GAA ACC CAC **GTC** AAG ATG GCC GAC CCC AAT CGC TTT CGC GGC AAA
 A G V D A E T H V K M A D P N R F R G K
 379 387 11 21

GAC CTT CCG **GTC CTG GAC CAG CTG** ACC GAC
 D L P V L D Q L T D
 23 28 31

b.

gD E2 E7

CCC CCG **GCC ACC** TAC GGG GTA... **CTC** AAT GCA GCA TCC TGA AGA TCT GAA TTC
 P P A T Y G V L N A A S *
 309 312 703 758

c.

gD RVE1

CCC CCG **GCC ACC** TGG TGG AAT CTG ACA TTA GGG GCC AIT TGT GCT CTG CCC CTA GTG GGT
 P P A T W W N L T L G A I C A L P L V G
 309 312 1 16

TTG TTA GCA TGC TGC GCC AAG TGT TTG TAT TAC CTT CGC GGA GCG ATT GCC CCT AGG TAG
 L L A C C A K C L Y Y L R G A I A P R *
 17 35

AGA TCT GAA TTC CGG

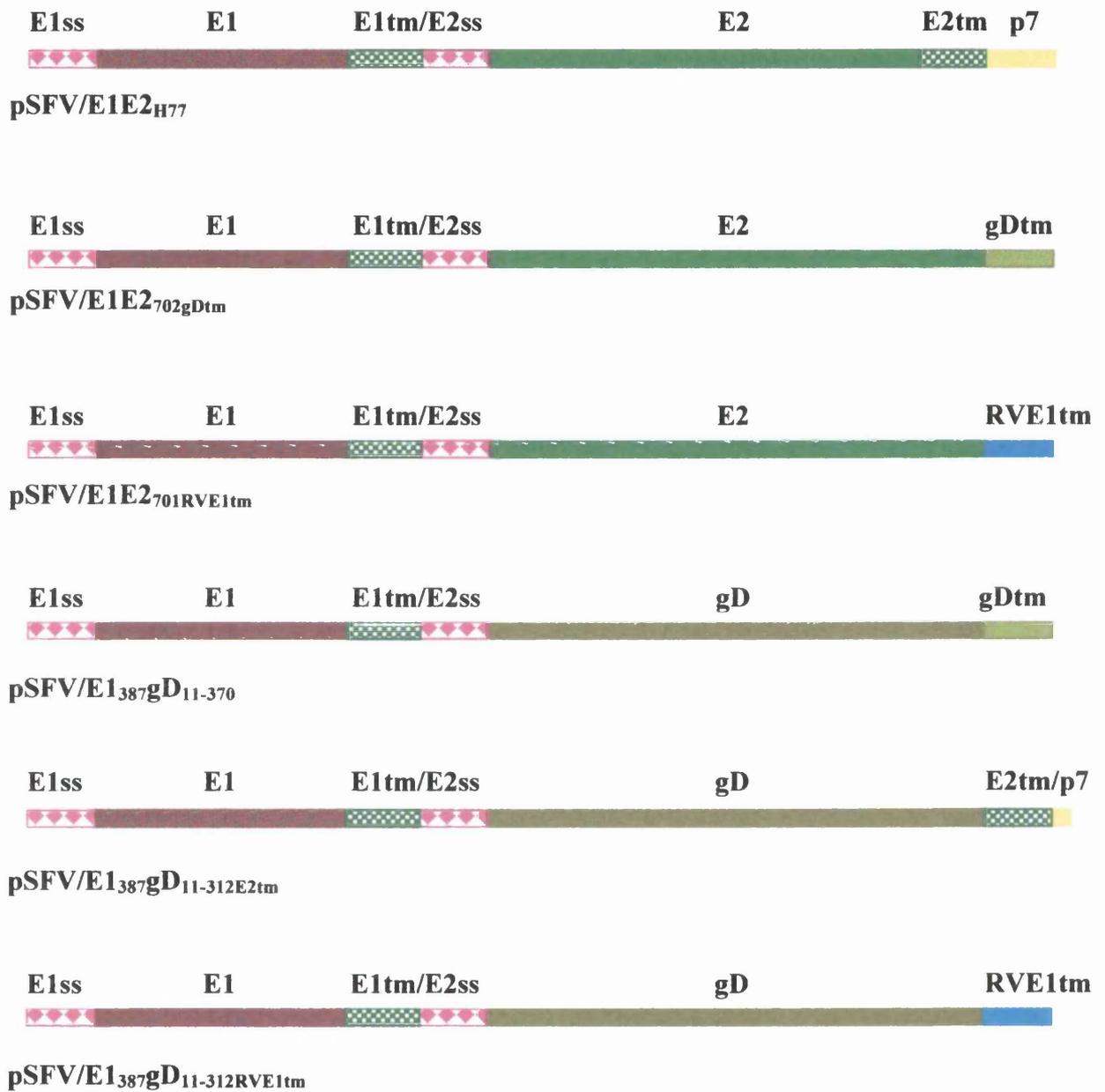


Fig 6.17. A schematic representation of the chimeric HCV E2 and HSV-1 gD proteins. ss and tm indicate signal sequence and transmembrane segments respectively.

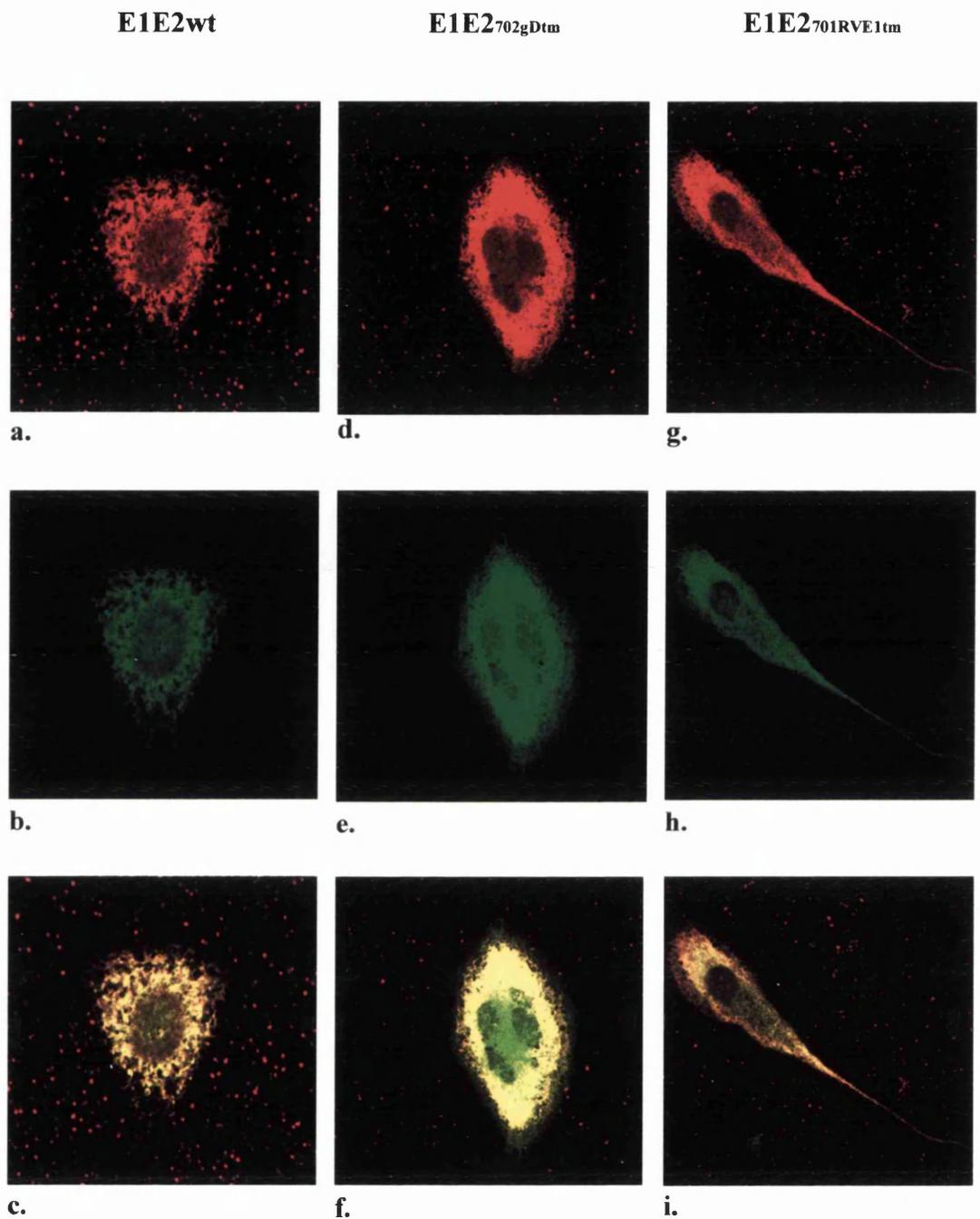


Fig 6.18. Localisation of E2 proteins fused to transmembrane domains of HSV-1 gD and rubella virus E1. BHK cells electroporated with RNA from pSFV/E1E2_{H77} (panels a-c), pSFV/E1E2_{702gDtm} (panels d-f) and pSFV/E1E2_{701RVE1tm} (panels g-i) were incubated at 37°C for 12 h, fixed in methanol and processed for indirect immunofluorescence using anti-E1 antiserum R528 (panels a, d and g) and E2-specific MA b ALP98 (panels b, e and h). Panels c, f and i represent merged images of panels a and b, d and e, and g and h, respectively.

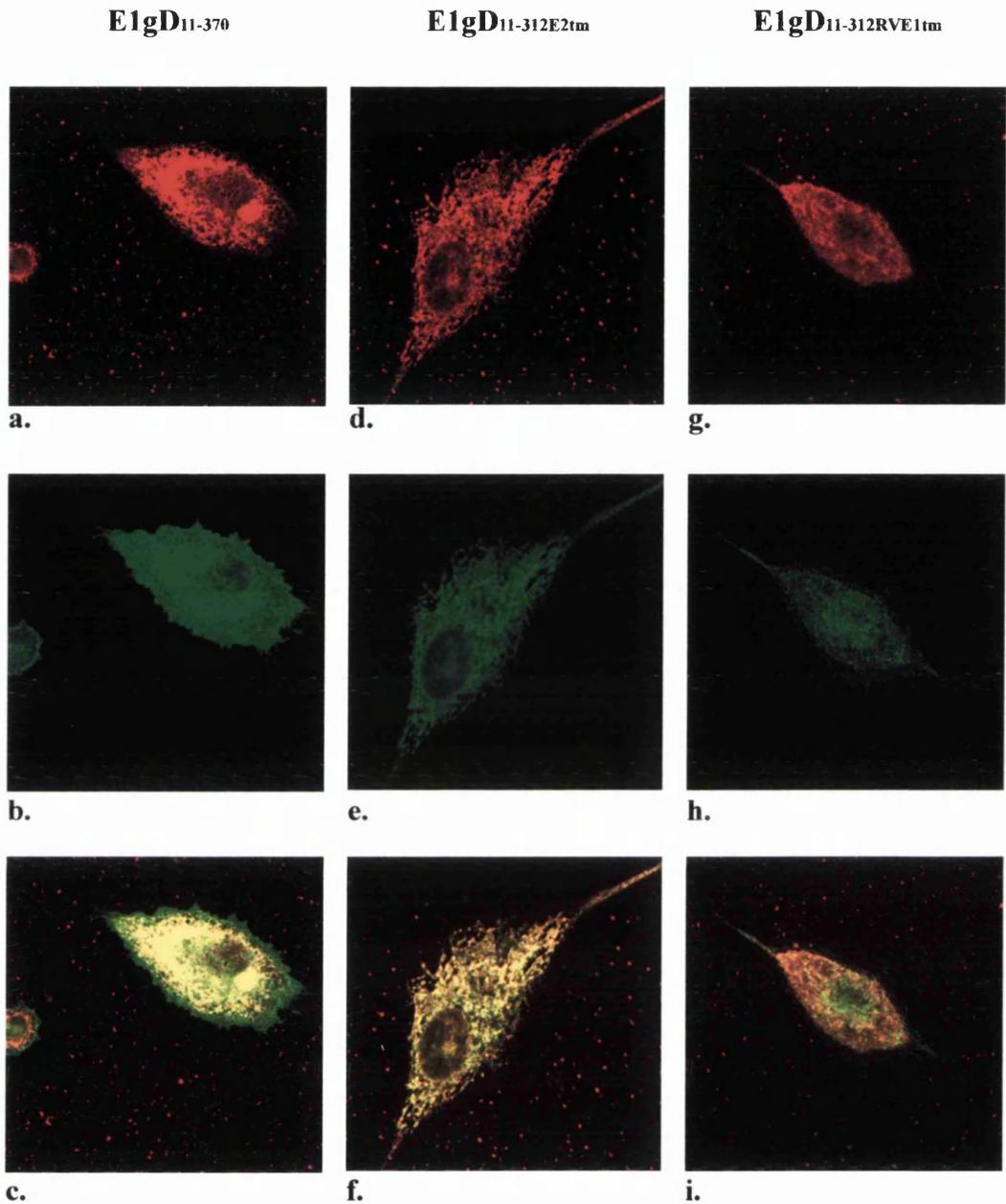


Fig 6.19. Localisation of HSV-1 gD fused to transmembrane domains of HCV E2 and rubella virus E1. BHK cells were electroporated with RNA from pSFV/E1₃₈₇gD₁₁₋₃₇₀ (panels a-c), pSFV/E1₃₈₇gD_{11-312E2tm} (panels d-f), and pSFV/E1₃₈₇gD_{11-312RVE1tm} (panels g-i) and incubated for 12 h. Subsequently, cells were fixed in methanol before being processed for indirect immunofluorescence using anti-E1 antiserum, R528, (panels a, d, and g) and gD-specific MAb 4846 (panels b, e, and h). Panels c, f and i show the merged images of panels a and b, d and e, and g and h, respectively.

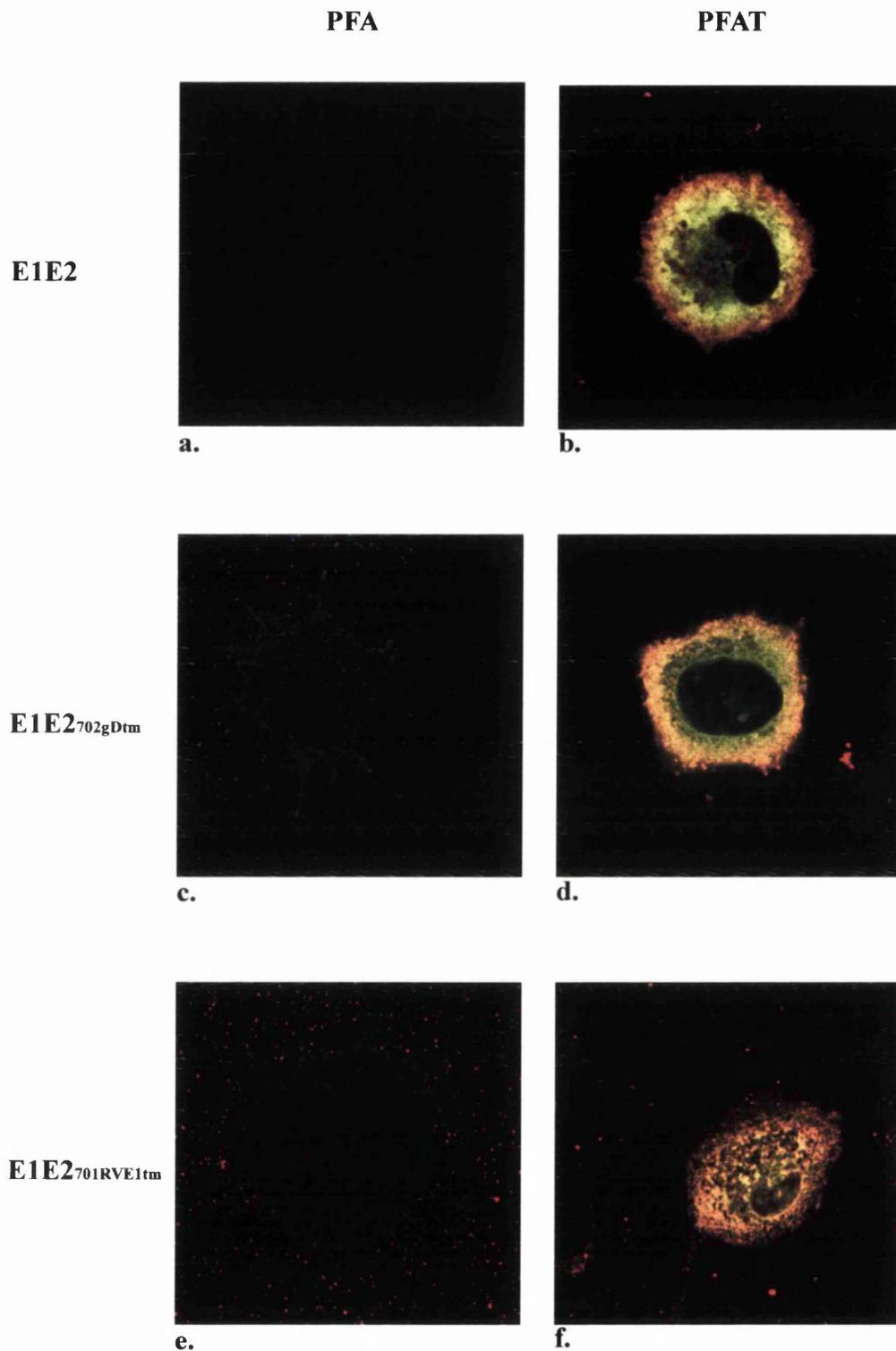


Fig 6.20. Cell surface localisation of E2 fused to the HSV-1 gD and rubella virus E1 transmembrane domains. BHK cells electroporated with RNA transcripts were incubated at 37°C for 12 h and fixed with paraformaldehyde (PFA). Parallel cultures were fixed and permeabilised with PFA containing 0.1% Triton X-100 (PFAT). These cells were processed for indirect immunofluorescence using anti-E1 antiserum R528 and E2-specific MAbs ALP98 and images were captured by confocal microscopy. Panels a, c and e are images derived from cells fixed with PFA alone, whereas panels b, d and f are images obtained from PFAT treated cells. Constructs were pSFV/E1E2_{H77} (panels a and b), pSFV/E1E2_{702gDtm} (panels c and d), and pSFV/E1E2_{701RVE1tm} (panels e and f).

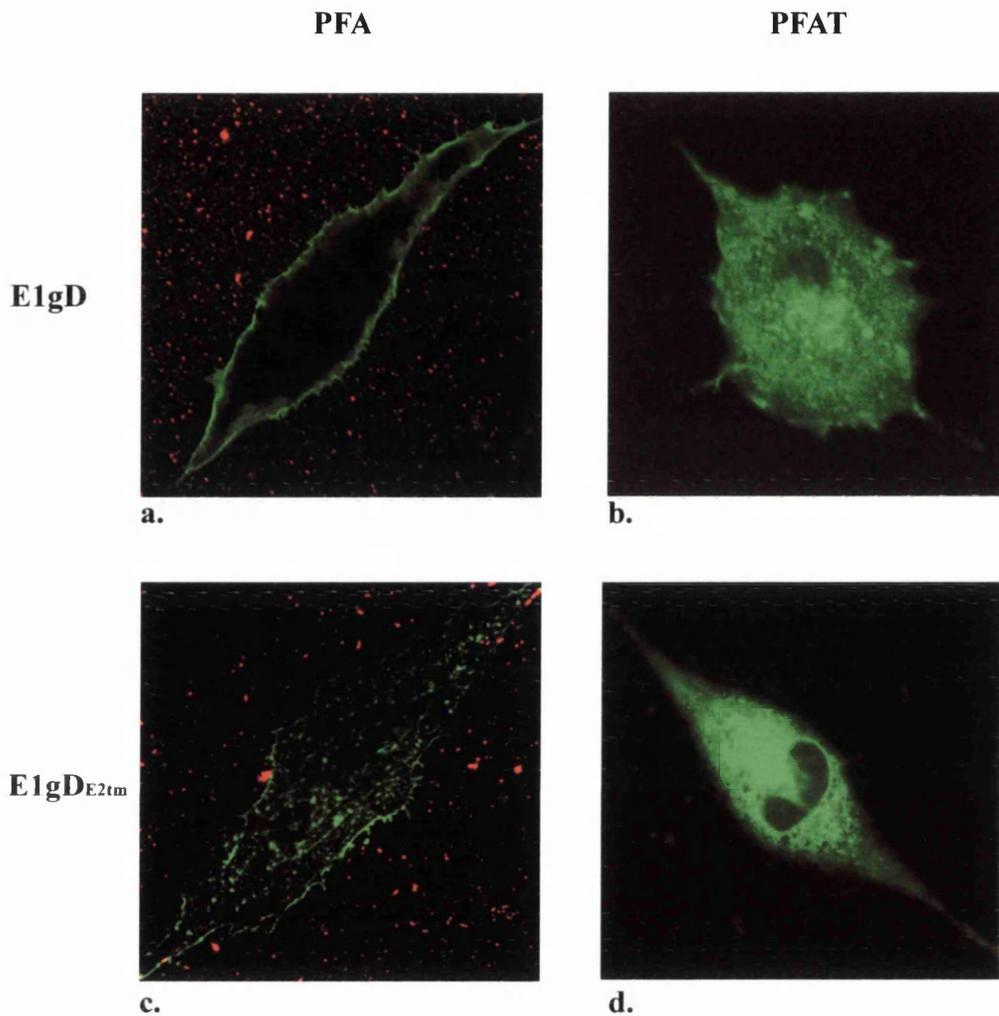


Fig 6.21. Cell surface expression of HSV-1 gD fused to the HCV E2 transmembrane domain. Experimental conditions were as described in Fig 6.20, except fixed cells were processed for indirect immunofluorescence using gD-specific MAb 4846. Panels a and c are images obtained from cells treated with PFA and panels b and d are those derived from PFAT treated cells. Constructs were pSFV/E1₃₈₇gD₁₁₋₃₇₀ (panels a and b) and pSFV/E1₃₈₇gD₁₁₋₃₁₂E_{2tm} (panels c and d).

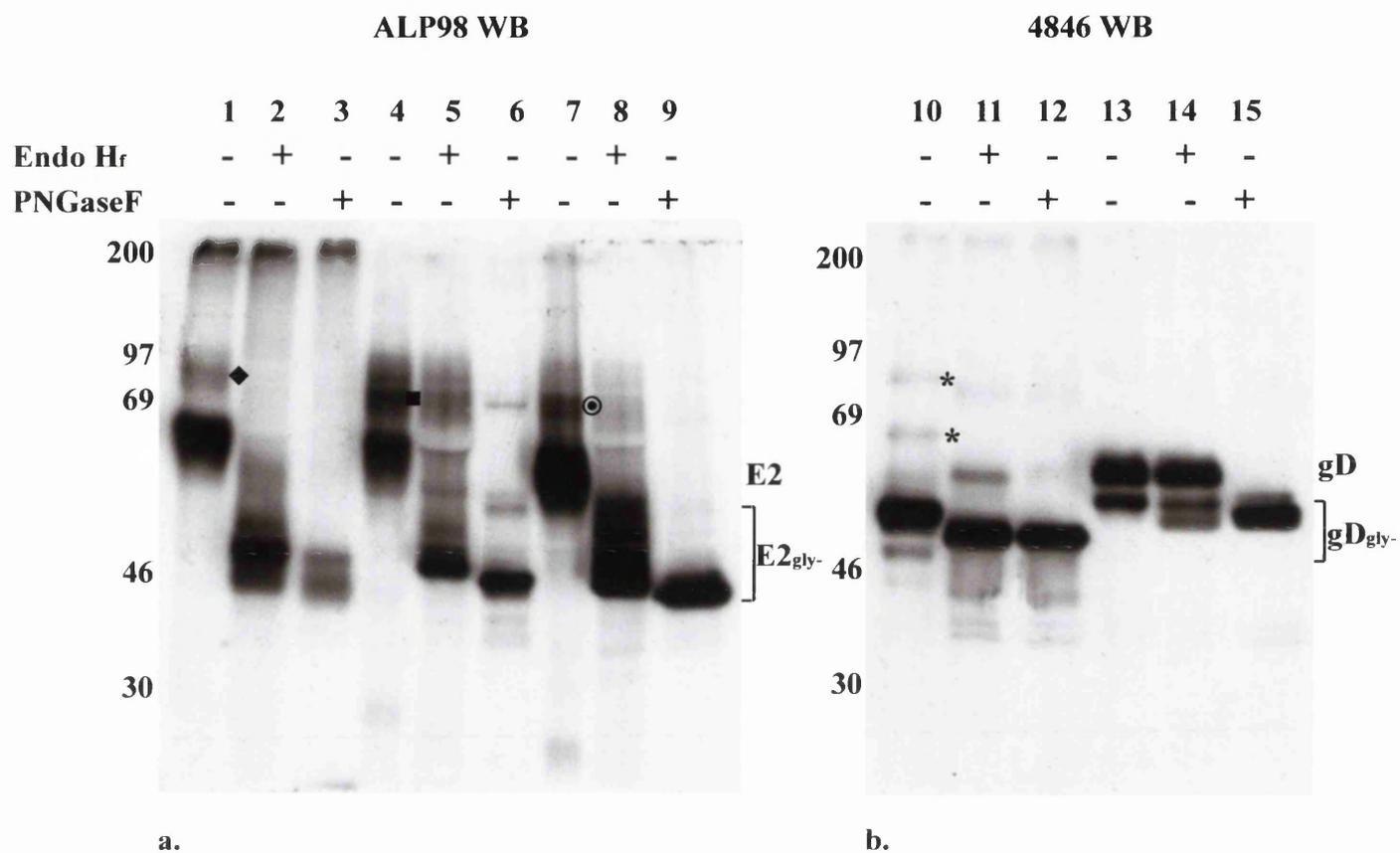


Fig 6.22. Endo H_f sensitivity of HCV E2 and HSV-1 gD chimeric proteins. Cells, electroporated with RNA transcripts, were incubated for 12 h at 37°C and lysed. Portions of the crude extracts were treated with endo H_f and PNGase F and analysed in parallel with untreated extracts by electrophoresis on 10% polyacrylamide gels. After transferring proteins to membranes, Western blot analysis was performed using E2-specific MAb ALP98 (panel a) and gD-specific MAb 4846 (panel b). Lanes 2, 5, 8, 11 and 14 show samples subjected to endo H_f treatment and those in lanes 3, 6, 9, 12 and 15 were PNGase F-treated. Remaining lanes represent untreated proteins. Samples were pSFV/E1E2_{H177} (lanes 1-3), pSFV/E1E2_{702gDtm} (lanes 4-6), pSFV/E1E2_{701RVE1tm} (lanes 7-9), pSFV/E1_{387gD}_{11-312E2tm} (lanes 10-12), pSFV/E1_{387gD}₁₁₋₃₇₀ (lanes 13-15). Positions of E2, deglycosylated E2 (gly-), gD and deglycosylated gD are shown. Bands indicated by symbols are described in the text.

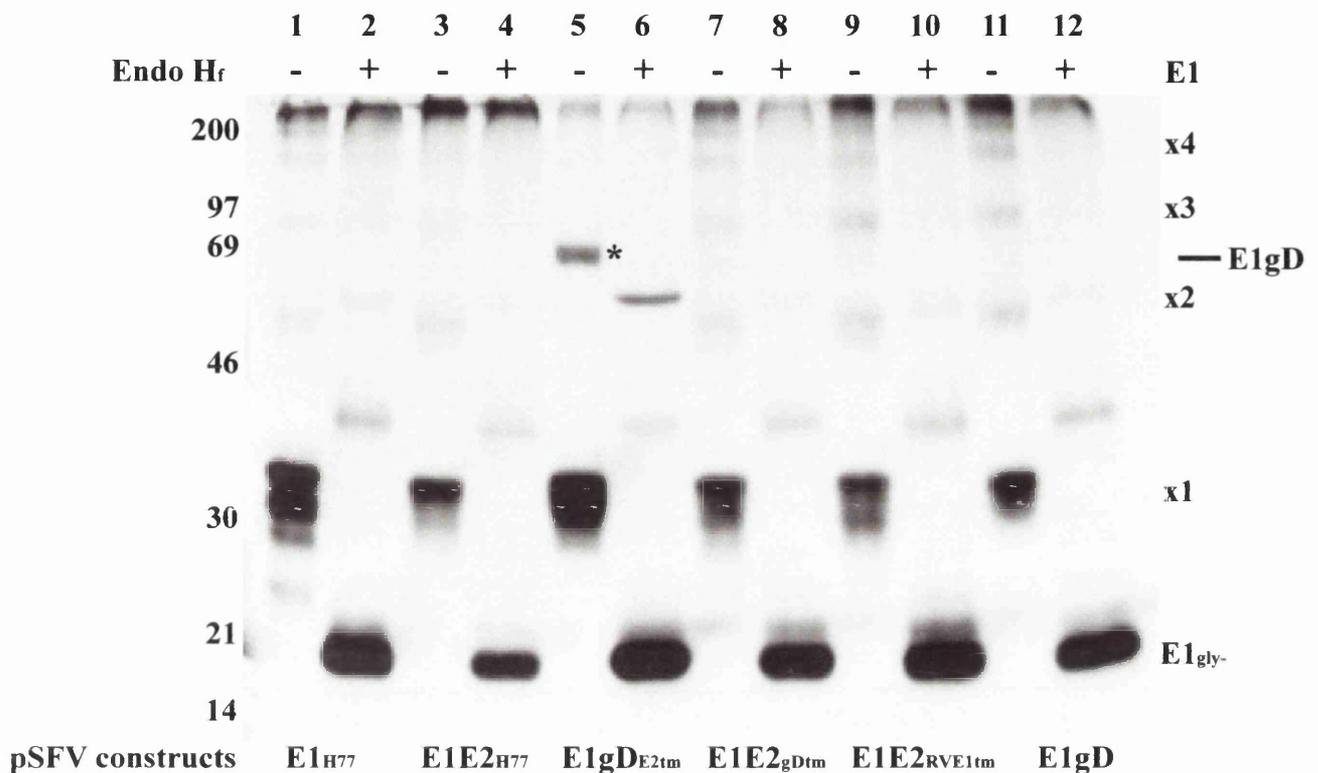


Fig 6.23. Sensitivity to endo H_r of E1 expressed in the context of chimeric proteins. Experimental conditions were as described in Fig 6.22 except crude extracts were treated only with endo H_r. Following electrophoresis on a 12% polyacrylamide gel, samples were analysed alongside untreated material by Western blot analysis using E1-specific AP21.010 MAbs. Lanes 2, 4, 6, 8, 10 and 12 show the endo H_r treated samples, while remaining lanes represent untreated proteins. Samples were pSFV/E1_{H77} (lanes 1 and 2), pSFV/E1E2_{H77} (lanes 3 and 4), pSFV/E1_{387gD}_{11-312E2tm} (lanes 5 and 6), pSFV/E1E2_{702gDtm} (lanes 7 and 8), pSFV/E1E2_{701RVE1tm} (lanes 9 and 10), and pSFV/E1_{387gD}₁₁₋₃₇₀ (lanes 11 and 12). Oligomers of E1 and deglycosylated forms of E1 (-gly) are shown along with a cleaved form of E1gD polyprotein (*).

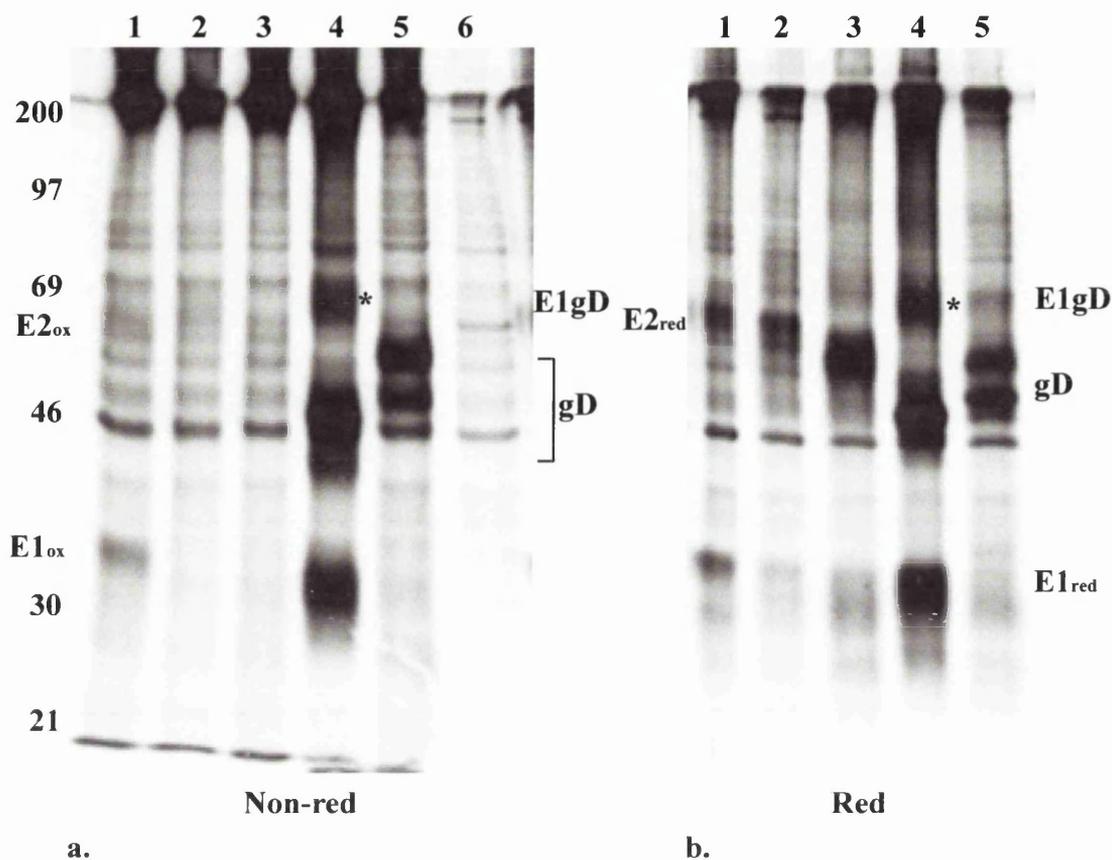


Fig 6.24. Examination of complex formation by chimeric E2 and gD proteins with E1 using E2-specific MAb ALP98 and gD-specific MAb 4846. Cells electroporated with RNA were ^{35}S -labelled from 4-12 h after electroporation. Following cell lysis, crude extracts were subjected to immunoprecipitation using E2-specific MAb ALP98 (lanes 1-3) and gD-specific MAb 4846 (lanes 4-6). Precipitated material was examined under non-reducing (panel a) and reducing (panel b) conditions on 10% polyacrylamide gels. Samples were derived from pSFV/E1E2_{H77} (lanes 1), pSFV/E1E2_{702gDtm} (lanes 2), pSFV/E1E2_{701RVE1tm} (lanes 3), pSFV/E1_{387gD}_{11-312E2tm} (lanes 4), pSFV/E1_{387gD}₁₁₋₃₇₀ (lanes 5), and pSFV1 (lane 6). Positions of oxidised (ox) and reduced (red) E1 and E2 are indicated, as are those of gD and uncleaved E1gD precursor protein. Bands indicated by * are discussed in the text.

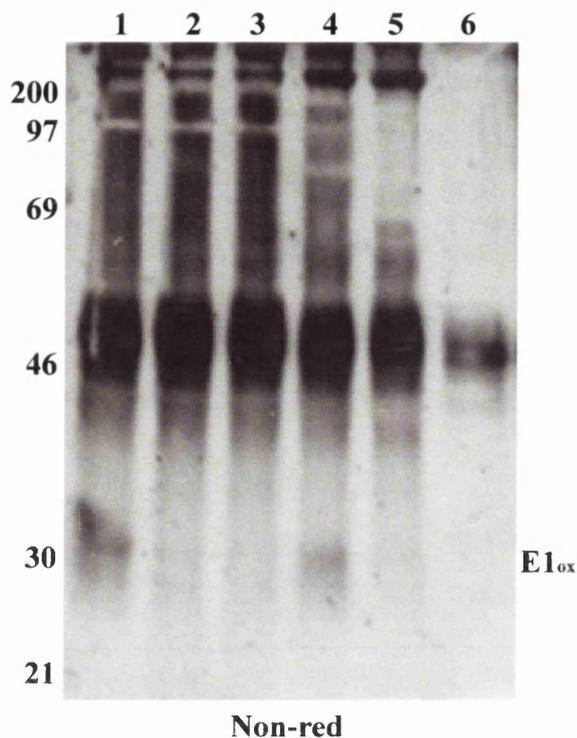


Fig 6.25. Co-precipitation of E1 with wild type E2 and chimeric gD proteins. Experimental conditions were as described for Fig 6.24 except cells were not ^{35}S -labelled. Following immunoprecipitation using E2-specific ALP98 and gD-specific 4846, precipitates were separated by electrophoresis on a non-reducing 10% polyacrylamide gel and transferred to membrane. Western blot analysis was performed using anti-E1 antiserum R528. Samples were pSFV/E1E2_{H77} (lane 1), pSFV/E1E2_{702gDtm} (lane 2), pSFV/E1E2_{701RVE1tm} (lane 3), pSFV/E1_{387gD11-312E2tm} (lane 4), pSFV/E1_{387gD11-370} (lane 5), and pSFV1 (lane 6). Bands representing oxidised E1 (ox) are indicated. The additional bands represent the heavy and light chains of the immunoglobulins used for immunoprecipitation.

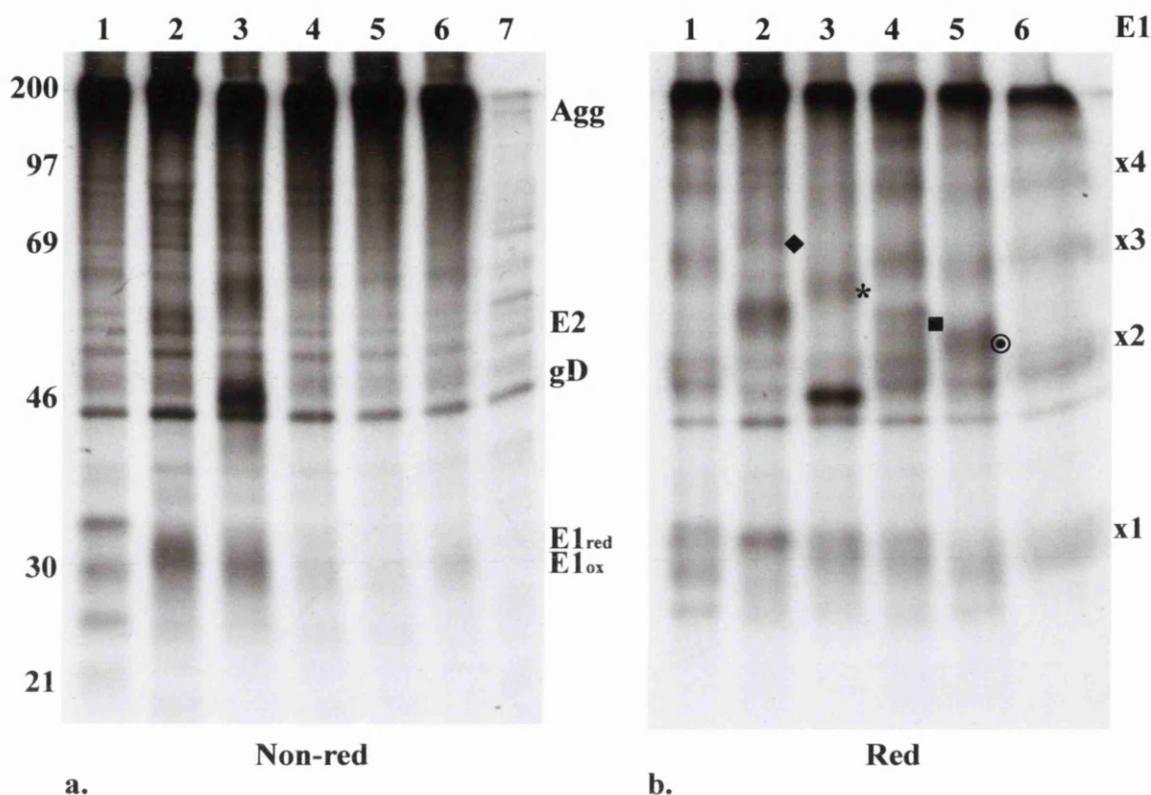


Fig 6.26. Examination of complex formation between E1 and chimeric E2 and gD proteins. Experimental conditions were as described for Fig 6.24 except immunoprecipitations were performed using anti-E1 antiserum R528. The precipitated samples were analysed under non-reducing (panel a) and reducing (panel b) electrophoresis conditions on 10% polyacrylamide gels. Samples were pSFV/E1_{H77} (lanes 1), pSFV/E1E2_{H77} (lanes 2), pSFV/E1₃₈₇gD₁₁₋₃₁₂E2_{tm} (lanes 3), pSFV/E1E2₇₀₂gD_{tm} (lanes 4), pSFV/E1E2_{701R}VE1_{tm} (lanes 5), pSFV/E1₃₈₇gD₁₁₋₃₇₀ (lanes 6), and pSFV1 (lane 7). Bands representing aggregated material (Agg), gD and E2 are indicated. Oxidised (ox), reduced (red) and oligomeric E1 are also shown. Bands indicated by symbols are discussed in the text.

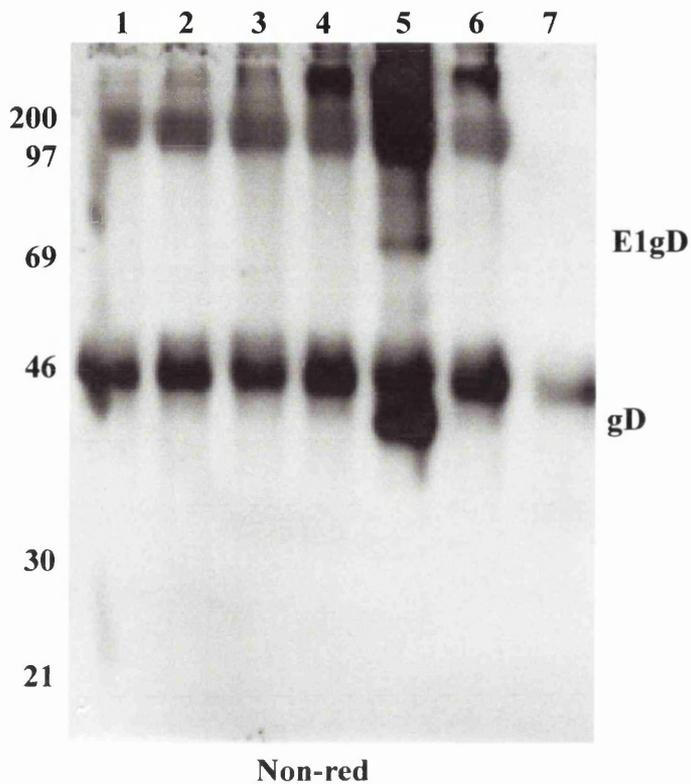


Fig 6.27. Examination of E1-specific R528 immunoprecipitated material for the co-precipitation of gD. Experimental conditions were as described for Fig 6.26 except proteins were not radiolabelled. Samples were separated under non-reducing conditions on a 10% polyacrylamide gel and transferred to membrane for Western blot analysis using gD-specific MAb 4846. Samples analysed were pSFV/E1_{H77} (lane 1), pSFV/E1E2_{702gDtm} (lane 2), pSFV/E1E2_{701RVE1tm} (lane 3), pSFV/E1E2_{H77} (lane 4), pSFV/E1_{387gD11-312E2tm} (lane 5), pSFV/E1_{387gD11-370} (lane 6), and pSFV1 (lane 7). Bands representing gD and precursor E1gD polyprotein are indicated. The remaining bands represent heavy and light chains of immunoglobulins used for immunoprecipitation.

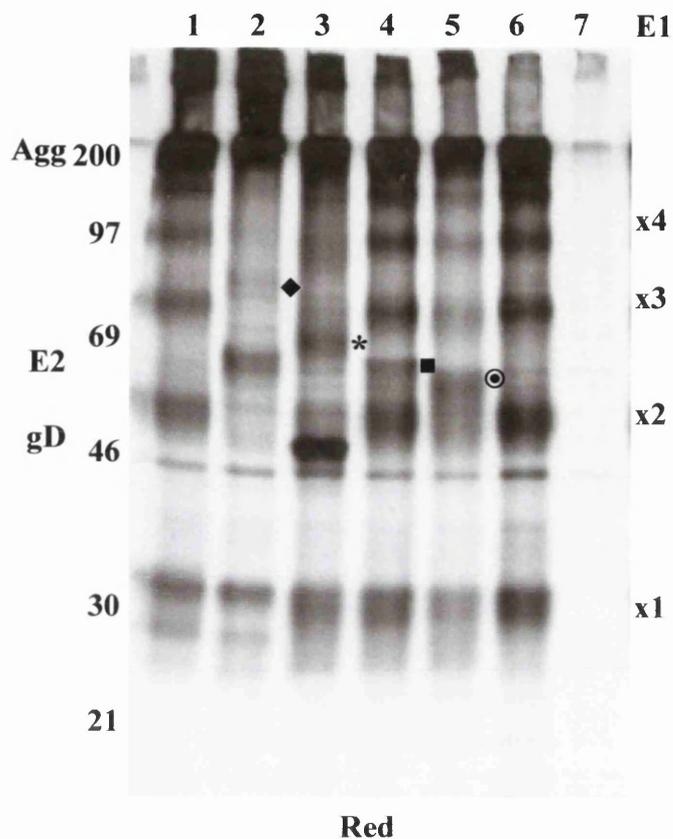


Fig 6.28. Analysis of complex formation of chimeric E2 and gD proteins with E1. Experimental conditions were as described for Fig 6.26 except immunoprecipitations were performed using E1-specific MAAb AP497 and precipitated material was examined under reducing electrophoretic conditions on a 10% polyacrylamide gel. Samples were pSFV/E1_{H77} (lane 1), pSFV/E1E2_{H77} (lane 2), pSFV/E1₃₈₇gD₁₁₋₃₁₂E2_{tm} (lane 3), pSFV/E1E2₇₀₂gD_{tm} (lane 4), pSFV/E1E2₇₀₁RVE1_{tm} (lane 5), pSFV/E1₃₈₇gD₁₁₋₃₇₀ (lane 6), and pSFV1 (lane 7). Positions of E2, gD, aggregated material (Agg), and oligomerised forms of E1 are indicated. Bands indicated by symbols are discussed in the text.

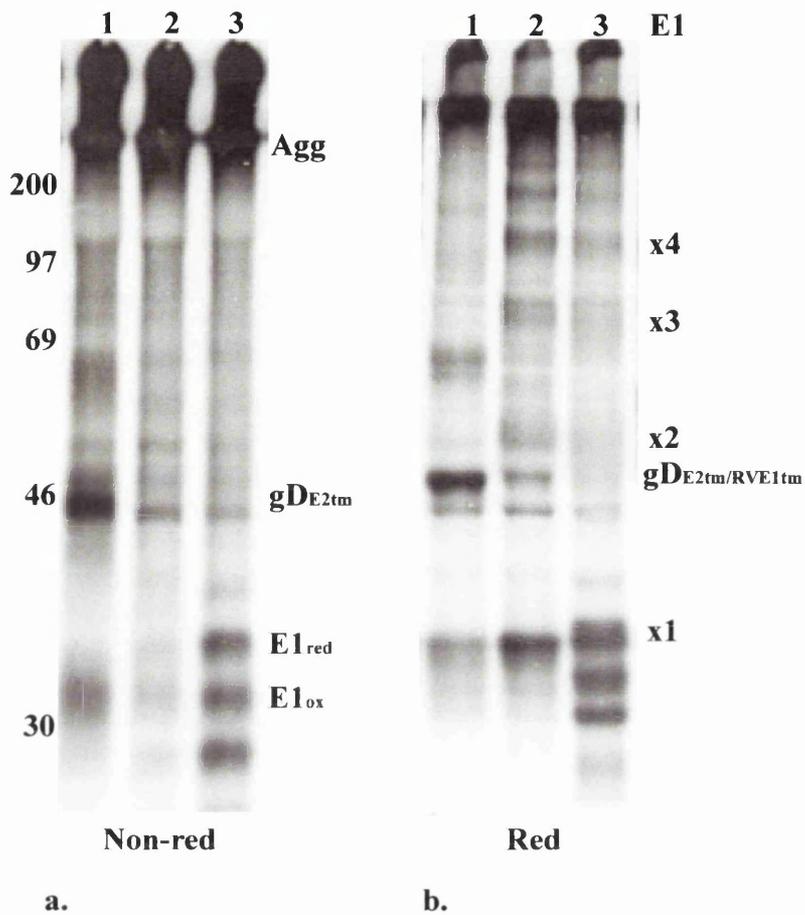


Fig 6.29. Examination of the behaviour of E1 expressed along with gD fused to rubella virus E1 transmembrane domain. RNA-electroporated cells were incubated at 37°C and labelled between 4-12 h after electroporation with ³⁵S-methionine. Crude extracts derived from these cells were subjected to immunoprecipitation using anti-E1 antiserum R528. Precipitates were examined under non-reducing (panel a) and reducing (panel b) electrophoresis conditions on a 10% polyacrylamide gel. Samples were pSFV/E1₃₈₇gD₁₁₋₃₁₂E2_{tm} (lanes 1), pSFV/E1₃₈₇gD₁₁₋₃₁₂RVE1_{tm} (lanes 2), and pSFV/E1_{H77} (lanes 3). Positions of aggregated material (Agg), gD, oxidised (ox), reduced (red) and oligomerised E1 proteins are indicated.

Chapter 7 – General Discussion

7.1. The Arrangement of E1 and E2 on the HCV Polyprotein

The presence of E2 is necessary for E1 to attain its stable oxidised state, suggesting that E2 plays a critical role in E1 folding by interacting directly with the protein. It is possible that correct protein folding and hetero-oligomerisation in some enveloped viruses require certain structural constraints which are provided by protein-protein associations between monomers within the assembled oligomer. Since expression of E2 *in vivo* occurs from the same polyprotein as E1 (in *cis*), this could ensure that the subunits for a complex are generated by proteolytic cleavage of a common translation product. This local recruitment of subunits may increase the efficiency of oligomerisation and folding. Structural proteins from several viruses are expressed in a similar fashion (e.g., flavi-, rubella-, alpha-, and bunya-viruses). For TBE virus, a flavivirus, co-expression of PrM and E membrane-bound proteins is necessary for proper folding (Section 1.5.2; Heinz *et al.*, 1994). PrM functions in the PrM-E heterodimer to prevent E from undergoing an acid-catalysed conformational change during transit of the immature virions through an acidic intracellular compartment. In the case of rubella virus, despite its Golgi localisation signal, E2 remains associated with E1 in the ER during which shuffling of the intramolecular disulphide bonds in E1 takes place (Hobman *et al.*, 1993). In addition, complete folding of Sindbis virus glycoprotein E1 protein is not achieved in the absence of oligomerisation with PE2 (Carleton *et al.*, 1997). Therefore, it may be that stable protein-protein interactions between HCV E1 and E2 enable E1 to reach an oxidised form rather than E2 performing a specific chaperone-like function where interactions would be transient. Indeed, results in this study have shown that the substitution of the E2 ectodomain with that of a foreign glycoprotein was sufficient to prevent the misfolding of E1 (see Section 6.3.4).

The nature of HCV genomic organisation is such that the E1 polypeptide sequence is translated first and translocated into the ER before the E2 sequences. Dimerisation of SFV structural proteins E1 and preE2 takes place rapidly and efficiently in the ER, with both proteins originating from the same translation product (termed heterodimerisation in *cis*; Barth *et al.*, 1995). In this case, preE2 is made before E1 during 26S mRNA translation. It is proposed that the preE2 polypeptide chain is retained at its translocation site until the E1 chain is synthesised and translocated into the same site. Since the production of HCV E2

follows the synthesis of E1, the mechanism by which E2 affects E1 folding is not consistent with this model.

As an alternative to heterodimerisation in *cis* during translocation into the ER, E1 protein may encounter pre-folded E2 subunits made from different polyprotein molecules, which block E1 misfolding and form heteromeric E1E2 complexes. Heterodimerisation in *trans* takes place for the envelope heterodimeric complex (G1G2) of the bunyavirus Uukuniemi virus (Persson & Pettersson, 1991). G1 and G2 polypeptides are separated following cleavage from a common translation product soon after synthesis. The two proteins fold independently of each other and then form a complex. However, G1 and G2 in the complex need not have arisen from the same precursor protein. A similar mechanism may also operate for HCV E1 and E2 heterodimerisation, in which slowly-folding E1 may interact with E2 molecules that are already present in the ER. Speed with which aggregate formation takes place (as established in this study) and the reliance of E1 on E2 to prevent E1 oligomerisation suggests that the presence of E2 in the lumen of the ER may be necessary during the synthesis of E1 polypeptide. Since E1 is synthesised first, in the event of insufficient pre-formed E2 molecules, E1 polypeptides would remain in an unstable reduced form that may result in association with similarly reduced neighbouring E1 molecules and lead to aggregation. Although these may reflect a phenomenon of the method of expression *in vitro*, a similar process could occur *in vivo*.

7.2. Formation of E1E2 Aggregates – An Intrinsic Property of the Glycoproteins or an *In Vitro* Artefact?

Two types of complex have been identified between HCV glycoproteins E1 and E2. In a native complex, E1 and E2 are held together by non-covalent interactions while aggregated complex is stabilised by both non-covalent and covalent associations. For the structural proteins from the related flavi- and pestiviruses, only one form of complex is found. Examination of pestivirus CSFV structural proteins either expressed by recombinant vaccinia virus or derived from CSFV virions has revealed that structural proteins E^{rns} and E2 form disulphide-linked homodimers, while E1 and E2 form disulphide-linked heterodimers (Thiel *et al.*, 1991). By contrast, heterodimers formed by structural proteins PrM and E of three flaviviruses (West Nile virus, YFV and TBE virus) contain non-covalent interactions (Nowak & Wengler, 1987). Apart from the pestivirus glycoproteins,

several other viral envelope proteins require intermolecular disulphide bonds to attain an appropriately folded state. As examples, pseudorabies virus gM and gN, HCMV gH and gL, and Moloney murine leukemia virus gp70 and Pr15E form heteromeric disulphide-linked complexes (Jons *et al.*, 1998; Opstelten *et al.*, 1998; Huber & Compton, 1999). In common with the flaviviruses, the structural proteins of rubella virus (E1 and E2), SFV (preE2 and E1), and Sindbis virus (pE2 and E1) form non-covalently linked heteromeric complexes (Hobman *et al.*, 1993; Barth *et al.*, 1995; Carleton *et al.*, 1997).

The failure to detect intermediate oligomers of E1E2 heterodimers has led to the conclusion that the high molecular weight covalently-linked complexes represent misfolded protein aggregates (Dubuisson *et al.*, 1994; Dubuisson & Rice, 1996). These observations were made directly from tissue culture experiments where recombinant viral systems may account for the incorrect folding of highly expressed proteins. Aggregation may result from over-expression, an incompatible environment for protein folding or exhaustion of normal chaperone mechanisms. However, aggregated complexes have been detected also in cell lines where expression is tightly regulated (Duvet *et al.*, 1998; Moradpour *et al.*, 1998). Moreover, various sera derived from HCV-infected patients have been able to detect native as well as aggregated complexes (Cardoso *et al.*, 1998; Habersetzer *et al.*, 1998). This suggests that proteins in the aggregate do adopt at least a partially-folded structure that is recognised by conformation-sensitive antibodies which would not be present on denatured proteins (Lee *et al.*, 1997). Examination of aggregates (or inclusion bodies) formed as a result of expression of tailspike endorhamnosidase in *E. coli* showed that aggregation occurs by the association of partially folded monomeric intermediates with significant secondary structure rather than the fully native or completely unfolded species (Speed *et al.*, 1995 and 1996).

Aggregation is observed for several other virus structural proteins but occurs mainly as a result of mutations in the proteins (Einfeld & Hunter, 1997) or non-typical methods of expression (Andersson *et al.*, 1997). The SFV E1 protein, expressed independently from upstream E2 sequences, forms disulphide-linked aggregates (Andersson *et al.*, 1997) which indicates that E1 protein synthesised in a manner dissimilar to the situation *in vivo*, results in abnormally folded products. In addition, a glycosylated precursor protein (Env), encoded by retroviral Rous sarcoma virus, is cleaved during intracellular maturation to a transmembrane subunit (TMS) and a surface subunit (SUS). Mutations to the C-terminus of TMS, which is a type 1 glycoprotein, result in non-specific aggregation of mutant TMS

(Einfeld & Hunter, 1997). This could be applicable to E1 which appears to be particularly sensitive to alteration. Studies have shown that while E2 behaviour remains comparable when synthesised in the presence or absence of E1, this is not the case for E1. E1 expressed in the absence of the E2 transmembrane domain forms oligomers, which suggests that E1 is more prone to aggregation than E2. Moreover, some E1 deletion mutants are not rescued from the oligomerisation-aggregation pathway in the presence of E2 (Section 5.3.4). It is perhaps possible that the ability of E1 to aggregate may be an intrinsic property of the protein. An example of homo-oligomerisation is provided by the HCV NS3 proteinase, where the presence of the co-factor NS4A is necessary for efficient proteolytic activity (Love *et al.*, 1996). The crystal structure of NS3, revealed that, in the absence of NS4A, hydrophobic patches are exposed on NS3 molecules which interact with adjacent molecules to form NS3-NS3 complexes (Love *et al.*, 1996). However, these hydrophobic patches are buried under conditions where NS4A is present (Kim *et al.*, 1996). The presence of pre-folded E2 may prevent exposure of hydrophobic regions of E1 to other neighbouring E1 molecules.

Covalent oligomerisation, however, may not necessarily be a non-functional event, since aggregation of a baculovirus structural protein is essential for infection (Markovic *et al.*, 1998). Baculovirus gp64 monomers are not transported to the surface of the infected cell, but instead remain in the ER and are degraded. Phosphorylated, acylated and N-glycosylated trimers of gp64 are successfully transported to the surface of infected cells, where they are stabilised by intermolecular disulphide bonds formed post-translationally. These disulphide-linked aggregates are thought to mediate fusion of host cell membranes. However, in the case of HCV envelope protein aggregation, it is not possible to comment on the extent to which aggregation is a functional feature of the infectious process.

Protein aggregates have been implicated as aetiologic agents for human diseases such as Alzheimer's and Creutzfeldt-Jacob (Kocisko *et al.*, 1995). Alzheimer's disease is clinically characterised by the abnormal accumulation of amyloid aggregates. The aggregate plaques are composed of variously-sized amyloid peptides (e.g., A β ₄₂ and A β ₄₀) derived through proteolytic processing of the β -amyloid precursor protein (β APP). Mutations within β APP that alter its metabolism may lead to Alzheimer's disease. From tissue culture experiments, A β ₄₂ is made and retained within the ER in an insoluble state while A β ₄₀ is generated exclusively within the TGN (Greenfield *et al.*, 1999). The A β ₄₀

species assume soluble and detergent-insoluble complexes; the latter form remains within the TGN. HCV may use E1E2 aggregation to control the production of virus particles. Since these complexes are recognised by sera from HCV-infected patients, they may exist *in vivo*. Stable aggregates which are not subject to degradation also may limit the processing of proteins to peptide for presentation on the cell surface. In this way, HCV could avoid detection by the immune system (Aridor & Balch, 1999) and such a mechanism may contribute to persistence. CSFV also establishes a persistent infection, with the majority of assembled infectious virions remaining cell-associated (Moormann & Hulst, 1988). Accumulation of HCV E1E2 aggregates also may play a role in pathogenesis. For example, ER accumulation of VSV G reduced levels of COPII vesicle formation in cultured cells (Aridor *et al.*, 1999). This has a generalised negative impact on the normal transport pathway from the ER. In HCV-infected hepatocytes, accumulation of aggregates of E1 and E2 could also impair trafficking of cellular proteins from the ER. Indeed, a variant of α 1-antitrypsin called Pi2 accumulates in the ER of hepatocytes and is associated with liver injury in certain individuals (Sifers, 1995). Thus, aggregation of E1 and E2 could contribute to development of liver disease in HCV-infected individuals.

The continued sensitivity to DTT by both the aggregated and native complexes of HCV E1 and E2 may suggest that neither of the complexes has acquired a stable state. Taking Sindbis virus E1-PE2 as an example, these proteins are sensitive to DTT reduction at early stages after synthesis (Carleton & Brown, 1996). However, resistance to DTT is achieved following the adoption of fully oxidised conformations in the ER. Sindbis virus E1 folds through a series of disulphide intermediates that are associated with the ER-resident molecular chaperone BiP. Once E1 assumes a particular late intermediate conformation it forms a heterodimer with PE2 resulting in the simultaneous completion of E1 folding and dissociation from BiP. Heterodimerisation bestows increased resistance to DTT to the later folding intermediates, which is indicative of inaccessibility by DTT and stabilisation of the complex (Carleton & Brown, 1996). A similar series of events occurs during maturation of influenza virus HA and VSV G proteins (Tatu *et al.*, 1993; Tatu & Helenius, 1995). However, though strain H77 E1 and E2 in the native complex are believed to be in their fully oxidised states, no DTT-resistant species of E1 were observed after a 12 h incubation period (see Fig 4.15). DTT-resistant forms of E1 would be expected to migrate faster than the DTT-sensitive forms due to their more compact structure. Therefore, the native E1E2 complexes may not represent an authentic complex found on HCV particles. Clearly, it is also possible that a DTT-resistant state is never achieved even by authentic

complexes. Such questions can only be addressed in a system allowing efficient propagation of the virus.

7.3. Retention of E1 and E2 Glycoproteins within the ER

Endogenous ER-resident membrane proteins are thought to be retained in the ER rather than retrieved from the *cis* Golgi, as is the case for several KDEL, K(X)KXX, and RR containing luminal or membrane ER proteins (see Section 1.4.8). The distinction between retrieval and retention is made by the absence of Golgi modifications to N-linked oligosaccharides attached to the endogenous proteins. HCV E1 and E2 are resistant to deglycosylation with endo D, which suggests that these proteins remain in the ER without cycling between this organelle and the *cis* Golgi (Duvet *et al.*, 1998). Luminal chaperones including BiP, calnexin and calreticulin interact with newly synthesised E1 and E2 in the ER, and this event is presumed to assist their folding (Choukhi *et al.*, 1998). However, following dissociation of some E1E2 complexes from these chaperones, they remain in the ER. Chimeric proteins constructed utilising the E1 and E2 ectodomains have shown that the transmembrane domains from both proteins are necessary and sufficient for ER retention (Cocquerel *et al.*, 1998; Cocquerel *et al.*, 1999; Flint & McKeating, 1999). The experiments presented here confirm this role for the E2 transmembrane domain.

Examination of endogenous membrane-spanning ER proteins has led to the proposal of two models for retention, one is based on retention through transmembrane domain oligomerisation and the second relies on retention through differences in membrane thickness along the exocytic pathway (Nilsson & Warren, 1994). Although it is not certain which of these mechanisms is correct or indeed if both operate, examples of Golgi-localised proteins that use each of these models are known. The replacement of the membrane-spanning domain of α 2,6-sialyltransferase (SialylT; TGN resident protein) with a stretch of polyleucines did not prevent localisation to the Golgi (Munro, 1991). This and the observation that the membrane-spanning domains of Golgi resident proteins are shorter than those of plasma membrane proteins led to the suggestion that the thickness of the membrane determined the point at which resident Golgi enzymes are retained (Bretscher & Munro, 1993). As an example for the second model, the first membrane-spanning domain (stretch of 22 amino acids) of the M glycoprotein of an avian coronavirus, infectious bronchitis virus (IBV), mediates oligomerisation in the Golgi when this segment was fused

to VSV G protein (Weisz *et al.*, 1993). Extensive mutational analysis of the IBV E1 membrane-spanning domain showed that polar residues, lining one face of a predicted α -helix, are important both for the retention and oligomerisation of this molecule (Machamer *et al.*, 1993). There are 4 polar amino acid residues in the predicted hydrophobic/membrane-spanning region of HCV E1, while 9 polar residues are identified in the equivalent region in HCV E2 (Appendix 1 and 3). However, the importance of these residues in both proteins to either retention or oligomerisation is yet to be examined.

The transmembrane domains are important in heterodimer complex formation between structural proteins E1 and E2 of rubella virus (Hobman *et al.*, 1997). For this virus, wild type E1 is efficiently transported to the Golgi apparatus due to its association with E2 (Hobman *et al.*, 1993). This is despite the ER retention signal identified in the E1 transmembrane and cytoplasmic tail, suggesting that association with E2 results in the appropriate folding of E1. This process may override the ER retention signal, thus allowing transport of E1 away from the ER. Such a phenomenon is also demonstrated by pseudorabies virus glycoproteins gN and gM (Jons *et al.*, 1998) and bunyavirus G1 and G2 glycoproteins (Persson & Pettersson, 1991), where gN and G1 are retained in the ER until heteromeric complex formation permits their relocation. By contrast to the above examples, HCV E1 and E2 never leave the ER in detectable amounts in heterologous expression systems. In addition, ER-retained rubella virus E1E2 heterodimers do not appear to form very large complexes, whereas HCV E1 and E2 retained proteins do accumulate as aggregates.

In terms of the influence of length of transmembrane domains on ER retention, an example has been provided by the analysis of ER-associated UBC6, an ER resident protein with no known retention/retrieval signal. Functionally the protein has been implicated in targeting the ER translocator, Sec 61, for degradation by catalysing ubiquitin-conjugation. It is a member of the C-terminal anchored type IV transmembrane proteins and is characterised by the N-terminal of the protein facing the cytosolic side of the ER membrane. Examination of chimeric proteins in which the cytosolic domain of UBC6 was linked to a heterologous transmembrane domain, or in which the UBC6 transmembrane domain was appended to an unrelated soluble protein, led to the conclusion that the transmembrane domain of UBC6 played a dominant role in its ER localisation (Yang *et al.*, 1997).

Lengthening the wild type UBC6 hydrophobic segment from 17 to 21 amino acids resulted in re-targeting to the Golgi complex from the ER. Furthermore, an increase in length to 26

amino acids allowed the modified protein to traverse the secretory pathway and locate to the plasma membrane. This finding suggested that in the absence of specific targeting determinants, proteins could be sorted within the secretory pathway based on interactions between their transmembrane domains and the surrounding lipid bilayer. The minimum length of the C-terminal portion of rubella virus E1 required for its retention in the ER is 35 amino acids (Hobman *et al.*, 1997). For rotavirus VP7, another ER-retained protein, 31 residues (with three essential residues Ile-9, Thr-10, and Gly -11) from the N-terminus are sufficient to retain a reporter protein (Maass & Atkinson, 1994). Hydrophobicity profiles predict a stretch of 20 and 32 C-terminal hydrophobic amino acid residues for HCV E1 and E2 respectively (Appendices 1 and 3). While these residues may represent the transmembrane domain sequences, it is not clear whether any of them correspond to the cytoplasmic tails for E1 and E2. Therefore, it is difficult to state how the length of E1 and E2 transmembrane domains influence their ER retention. Furthermore, if there are C-terminal E1- and E2-specific residues on the cytoplasmic side of the ER they may also play a role in retention by interaction with cytoplasmic components. This is observed for rubella virus E1 retention (Hobman *et al.*, 1997). Since cytoplasmic tails are undefined for HCV E1 and E2, their role in ER retention cannot be excluded.

7.4. Role of Covalent Modifications on E1 Folding

As discussed earlier in this chapter, a consequence of the arrangement of HCV glycoprotein coding sequences is that E1 is translocated into the ER before E2. If aggregation is an intrinsic property of E1, then the virus may have evolved mechanisms to counteract this process. One mechanism could involve the N-linked glycosylation sites of E1 which are located close to the N-terminus. Three of the four glycosylation sites are located within 43 residues of the N-terminus of E1. Thus, these sites would be recognised and glycosylated during the earlier stages of E1 synthesis. Glycosylation is associated with enhancing solubility of proteins in the ER and interactions with chaperones such as calreticulin and calnexin (Section 1.4.5). Therefore, glycosylation in the N-terminal region of E1 may aid retention of the protein in a non-aggregated state prior to the availability of E2. Interaction of E2 with E1 would then facilitate correct folding of E1.

E2 is apparently essential for folding of E1 to an oxidised form. As described previously (Section 1.4.3), disulphide bond formation is catalysed by PDI, however this enzyme is unlikely to play a specific role in protein folding. The data in this study would indicate

that the C-terminal region of E2 including the transmembrane domain plays a key role in intramolecular disulphide bond formation in E1. The ability to replace most of E2 with a foreign glycoprotein suggests that these ectodomain sequences may act indirectly to block E1 misfolding, perhaps by acting as a buffer between adjacent ectodomains of E1 molecules. Additional studies are required to examine the oxidised state of E1 in the presence of E2 and foreign ectodomain sequences to reveal whether intramolecular disulphide bonds occur between the same cysteine residues in both situations. As well as E2, the primary sequence of E1 is likely to contribute also to its folding and further analysis is required to identify these residues. Possible interactions with other cellular proteins, such as ERp57 which may play a more active role in folding and oxidation (Section 1.4.4), may also shed further light on E1 folding.

7.5. Model for the Role of E1 and E2 in HCV Particle Assembly

The failure to detect HCV E1 and E2 on the cell surface is comparable with pestivirus structural proteins, which also are absent from this location (Gray & Nettleton, 1987). Electron microscopic examination of BVDV and border disease virus suggests that pestiviruses mature in intracellular vesicles and may be released by exocytosis (Gray & Nettleton, 1987). The majority of CSFV infectious particles also remain cell associated (Moormann & Hulst, 1988). Electron microscopic studies of YFV-infected cells observed morphologically mature virions in the ER, which then accumulate within disordered membrane-bound vesicles (Ishak *et al.*, 1988). Thus, for pesti- and flaviviruses, vesicular transport through the host secretory pathway is believed to be involved in transporting nascent virions from the ER to the plasma membrane where they are released by exocytosis. Evidence for this is provided by the observed modification of West Nile virus E and PrM glycans, to forms which are characteristic of carbohydrates present on host plasma membrane glycoproteins (Nowak *et al.*, 1989). Despite budding of flaviviruses into the ER, there have been no reports to date of ER retention signals identified in the structural proteins E and PrM. The failure to detect budding intermediates and cytoplasmic nucleocapsids has suggested that the assembly process of flaviviruses occurs rapidly. This may abolish the need for prolonged retention of the structural proteins in the ER.

HCV particle assembly may also occur in the ER, however reports on virion morphogenesis suffer from a lack of reproducibility. Virus-like particles of 50-60nm have been identified in cytoplasmic vesicles following transfection of Daudi cells with full-

length genomic RNA (Shimizu *et al.*, 1996b). Moreover, particles have also been produced in insect cells from recombinant baculovirus expressing core, E1 and E2 (Baumert *et al.*, 1998). However, these observations have not always been successfully reproduced (Wang *et al.*, 1997) and it is also not clear whether these particles are capable of attachment to cells.

In vivo, in infected hepatocytes, budding of capsids into the ER may drive virion morphogenesis. Data presented earlier (see Section 3.9) revealed that E1 and E2 are not distributed throughout the entire ER but are likely to be present in sub-compartments. Within these sub-compartments, E1 and E2 could exist either as native or aggregated complexes. Therefore, budding would presumably have to occur at such sub-compartments. The ectodomains of the HCV glycoproteins need not have a direct influence on budding, however their status, whether aggregated or native, may be relayed via their transmembrane/cytoplasmic domains to allow incorporation of appropriate folded proteins into the virus envelope. Sequences corresponding to a cytoplasmic tail are not defined for E1 and E2, although comparison with equivalent flavivirus glycoproteins suggests that these segments may be small. Despite the size of cytoplasmic domains, the transmembrane anchors of flaviviruses are highly conserved and are believed to be involved in envelope-nucleocapsid interaction (Rice, 1996). Once budding into the ER occurs, the E1 and E2 transmembrane domains would not remain continuous with the ER membrane and therefore would no longer operate as ER retention signals. Apart from the segments of E1 and E2 interacting with the nucleocapsid, other viral transmembrane proteins may also influence budding and force the formation of premature particles. For example, efficiently processed p7 may perform this function. Subsequent maturation could occur during exit from the cell. However, these events can only be fully understood in an appropriate virus propagation system either *in vitro* or *in vivo*.

Chapter 8 - References

- Alberini, C. M., Bet, P., Milstein, C. & Sitia, R. (1990). Secretion of immunoglobulin-M assembly intermediates in the presence of reducing agents. *Nature* **347**, 485-487.
- Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J. D. (1994). The endoplasmic reticulum. *Molecular Biology of the Cell*, pp. 577-598. Garland Publishing, Inc, New York.
- Aleta, J. M., Cimato, T. R. & Ettinger, M. J. (1998). Protein methylation: a signal event in post-translational modification. *Trends in Biochemical Sciences* **23**, 89-91.
- Allan, R. B. & Balch, W. E. (1999). Cell biology - protein sorting by directed maturation of Golgi compartments. *Science* **285**, 63-66.
- Allison, S. L., Schalich, J., Stiasny, K., Mandl, C. W., Kunz, C. & Heinz, F. X. (1995). Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. *Journal of Virology* **69**, 695-700.
- Allison, S. L., Stiasny, K., Stadler, K., Mandl, C. W. & Heinz, F. X. (1999). Mapping of functional elements in the stem-anchor region of tick-borne encephalitis virus envelope protein E. *Journal of Virology* **73**, 5605-5612.
- Alter, H. J., Sanchezpescador, R., Urdea, M. S., Wilber, J. C., Lagier, R. J., DiBisceglie, A. M., Shih, J. W. & Neuwald, P. D. (1995). Evaluation of branched DNA signal amplification for the detection of hepatitis C virus RNA. *Journal of Viral Hepatitis* **2**, 121-132.
- Alter, M. J. (1993). The detection, transmission, and outcome of hepatitis C virus infection. *Infectious Agents and Disease-Reviews Issues and Commentary* **2**, 155-166.
- Alter, M. J. (1995). Epidemiology of hepatitis C in the West. *Seminars in Liver Disease* **15**, 5-14.
- Alter, M. J., Hadler, S. C., Judson, F. N., Mares, A., Alexander, W. J., Hu, P. Y., Miller, J. K., Moyer, L. A., Fields, H. A., Bradley, D. W. & Margolis, H. S. (1990). Risk factors for acute non-A, non-B hepatitis in the United States and association with hepatitis C virus infection. *Jama-Journal of the American Medical Association* **264**, 2231-2235.
- Andersson, H., Barth, B. U., Ekstrom, M. & Garoff, H. (1997). Oligomerization-dependent folding of the membrane fusion protein of Semliki Forest virus. *Journal of Virology* **71**, 9654-9663.

- Aridor, M., Bannykh, S. I., Rowe, T. & Balch, W.E. (1999). Cargo can modulate COPII vesicle formation from the endoplasmic reticulum. *Journal of Biological Chemistry* **274**, 4389-4399.
- Aridor, M. & Balch, W. E. (1999). Integration of endoplasmic reticulum signaling in health and disease. *Nature Medicine* **5**, 745-751.
- Asabe, S. I., Tanji, Y., Satoh, S., Kaneko, T., Kimura, K. & Shimotohno, K. (1997). The N-terminal region of hepatitis C virus-encoded NS5A is important for NS4A-dependent phosphorylation. *Journal of Virology* **71**, 790-796.
- Atkins, G. J., Sheahan, B. J. & Liljestrom, P. (1999). The molecular pathogenesis of Semliki Forest virus: a model virus made useful? *Journal of General Virology* **80**, 2287-2297.
- Barba, G., Harper, F., Harada, T., Kohara, M., Goulinet, S., Matsuura, Y., Eder, G., Schaff, Z., Chapman, M. J., Miyamura, T. & Brechot, C. (1997). Hepatitis C virus core protein shows a cytoplasmic localization and associates to cellular lipid storage droplets. *Proceedings of the National Academy of Sciences, USA* **94**, 1200-1205.
- Bartenschlager, R. (1999). The NS3/4A proteinase of the hepatitis C virus: unravelling structure and function of an unusual enzyme and a prime target for antiviral therapy. *Journal of Viral Hepatitis* **6**, 165-181.
- Bartenschlager, R., Ahlbornlaake, L., Mous, J. & Jacobsen, H. (1993). Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *Journal of Virology* **67**, 3835-3844.
- Bartenschlager, R., Ahlbornlaake, L., Mous, J. & Jacobsen, H. (1994). Kinetic and structural analyses of hepatitis C virus polyprotein processing. *Journal of Virology* **68**, 5045-5055.
- Bartenschlager, R., Lohmann, V., Wilkinson, T. & Koch, J. O. (1995). Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation. *Journal of Virology* **69**, 7519-7528.
- Barth, B. U., Wahlberg, J. M. & Garoff, H. (1995). The oligomerization reaction of the Semliki Forest virus membrane protein subunits. *Journal of Cell Biology* **128**, 283-291.
- Bartholomeus, A. & Thompson, P. (1999). Flaviviridae polymerase and RNA replication. *Journal of Viral Hepatitis* **6**, 261-270.
- Baumert, T. F., Ito, S., Wong, D. T. & Liang, T. J. (1998). Hepatitis C virus structural proteins assemble into viruslike particles in insect cells. *Journal of Virology* **72**, 3827-3836.

- Bause, E. (1983). Structural requirements of N-glycosylation of proteins. *Biochemical Journal* **209**, 331-336.
- Behrens, S. E., Tomei, L. & DeFrancesco, R. (1996). Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO Journal* **15**, 12-22.
- Behrens, S. E., Grassmann, C. W., Thiel, H. J., Meyers, G. & Tautz, N. (1998). Characterization of an autonomous subgenomic pestivirus RNA replicon. *Journal of Virology* **72**, 2364-2372.
- Berg, J. M. & Shi, Y. G. (1996). The galvanization of biology: a growing appreciation for the roles of zinc. *Science* **271**, 1081-1085.
- Berglund, P., Sjoberg, M., Garoff, H., Atkins, G. J., Sheahan, B. J. & Liljestrom, P. (1993). Semliki Forest virus expression system - production of conditionally infectious recombinant particles. *Bio-Technology* **11**, 916-920.
- Berglund, P., Smerdou, C., Fleeton, M. N., Tubulekas, I. & Liljestrom, P. (1998). Enhancing immune responses using suicidal DNA vaccines. *Nature Biotechnology* **16**, 562-565.
- Bishop, W. R., & Bell, R. M. (1988). Assembly of phospholipids into cellular membranes: biosynthesis, transmembrane movement, and intracellular translocation. *Annual Reviews in Cell Biology* **4**, 579-610.
- Blitvich, B. J., Mackenzie, J. S., Coelen, R. J., Howard, M. J. & Hall, R. A. (1995). A novel complex formed between the flavivirus-E and NS1 proteins - analysis of its structure and function. *Archives of Virology* **140**, 145-156.
- Blobel, G. & Dobberstein, B. (1975). Transfer of proteins across membranes. *Journal of Cell Biology* **67**, 835-851.
- Booth, J. C. L. (1998). Chronic hepatitis C: the virus, its discovery and the natural history of the disease. *Journal of Viral Hepatitis* **5**, 213-222.
- Borman, A., Howell, M. T., Patton, J. G. & Jackson, R. J. (1993). The involvement of a spliceosome component in internal initiation of human rhinovirus RNA translation. *Journal of General Virology* **74**, 1775-1788.
- Borowski, P., Heiland, M., Oehlmann, K., Becker, B., Kornetzky, L., Feucht, H. & Laufs, R. (1996). Non-structural protein 3 of hepatitis C virus inhibits phosphorylation mediated by cAMP-dependent protein kinase. *European Journal of Biochemistry* **237**, 611-618.
- Braakman, I., Helenius, J. & Helenius, A. (1992a). Manipulating disulfide bond formation and protein folding in the endoplasmic reticulum. *EMBO Journal* **11**, 1717-1722.

- Braakman, I., Helenius, J. & Helenius, A. (1992b). Role of ATP and disulfide bonds during protein folding in the endoplasmic reticulum. *Nature* **356**, 260-262.
- Brands, R., Snider, M. D., Hino, Y., Park, S. S., Gelboin, H. V. & Rothman, J. E. (1985). Retention of membrane proteins by the endoplasmic reticulum. *Journal of Cell Biology* **101**, 1724-1732.
- Bretscher, M. S. & Munro, S. (1993). Cholesterol and the Golgi apparatus. *Science* **261**, 1280-1281.
- Brillanti, S., Garson, J., Foli, M., Whitby, K., Deaville, R., Masci, C., Miglioli, M. & Barbara, L. (1994). Pilot study of combination therapy with ribavirin plus interferon alfa for interferon alfa-resistant chronic hepatitis C. *Gastroenterology* **107**, 812-817.
- Brinton, M. A. & Dispoto, J. H. (1988). Sequence and secondary structure - analysis of the 5'-terminal region of flavivirus genome RNA. *Virology* **162**, 290-299.
- Brock, K. V., Deng, R. & Riblet, S. M. (1992). Nucleotide sequencing of 5' and 3' termini of bovine viral diarrhea virus by RNA ligation and PCR. *Journal of Virological Methods* **38**, 39-46.
- Brodsky, J. L. & McCracken, A. A. (1997). ER-associated and proteasome-mediated protein degradation: how two topologically restricted events came together. *Trends in Cell Biology* **7**, 151-156.
- Brown, D. A. (1992). Interactions between GPI-anchored proteins and membrane lipids. *Trends in Cell Biology* **2**, 338-343.
- Brown, E. A., Zhang, H. C., Ping, L. H. & Lemon, S. M. (1992). Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Research* **20**, 5041-5045.
- Bruschke, C. J. M., Hulst, M. M., Moormann, R. J. M., vanRijn, P. A. & vanOirschot, J. T. (1997). Glycoprotein E-rns of pestiviruses induces apoptosis in lymphocytes of several species. *Journal of Virology* **71**, 6692-6696.
- Bulleid, N. J., Dalley, J. A. & Lees, J. F. (1997). The C-propeptide domain of procollagen can be replaced with a transmembrane domain without affecting trimer formation or collagen triple helix folding during biosynthesis. *EMBO Journal* **16**, 6694-6701.
- Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. (1994). Structure of influenza hemagglutinin at the pH of membrane fusion. *Nature* **371**, 37-43.
- Cardoso, M. D., Siemoneit, K., Strum, D., Krone, C., Moradpour, D. & Kubanek, B. (1998). Isolation and characterization of human monoclonal antibodies against hepatitis C virus envelope glycoproteins. *Journal of Medical Virology* **55**, 28-34.

- Carleton, M. & Brown, D. T. (1996). Disulfide bridge-mediated folding of Sindbis virus glycoproteins. *Journal of Virology* **70**, 5541-5547.
- Carleton, M., Lee, H., Mulvey, M. & Brown, D. T. (1997). Role of glycoprotein PE2 in formation and maturation of the Sindbis virus spike. *Journal of Virology* **71**, 1558-1566.
- Chamberlain, R. W., Adams, N., Saeed, A. A., Simmonds, P. & Elliott, R. M. (1997). Complete nucleotide sequence of a type 4 hepatitis C virus variant, the predominant genotype in the Middle East. *Journal of General Virology* **78**, 1341-1347.
- Chang, J. C., Seidel, C., Ofenloch, B., Jue, D. L., Fields, H. A. & Khudyakov, Y. E. (1999). Antigenic heterogeneity of the hepatitis C virus NS4 protein as modeled with synthetic peptides. *Virology* **257**, 177-190.
- Cheng, J. C., Chang, M. F. & Chang, S. C. (1999). Specific interaction between the hepatitis C virus NS5B RNA polymerase and the 3' end of the viral RNA. *Journal of Virology* **73**, 7044-7049.
- Choo, Q. L., Kuo, G., Ralston, R., Weiner, A., Chien, D., Vannest, G., Han, J., Berger, K., Thudium, K., Kuo, C., Kansopon, J., McFarland, J., Tabrizi, A., Ching, K., Moss, B., Cummins, L. B., Houghton, M. & Muchmore, E. (1994). Vaccination of chimpanzees against infection by the hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **91**, 1294-1298.
- Choo, Q. L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W. & Houghton, M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral-hepatitis genome. *Science* **244**, 359-362.
- Choo, Q. L., Richman, K. H., Han, J. H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medinaselby, A., Barr, P. J., Weiner, A. J., Bradley, D. W., Kuo, G. & Houghton, M. (1991). Genetic organization and diversity of the hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **88**, 2451-2455.
- Choukhi, A., Ung, S., Wychowski, C. & Dubuisson, J. (1998). Involvement of endoplasmic reticulum chaperones in the folding of hepatitis C virus glycoproteins. *Journal of Virology* **72**, 3851-3858.
- Clarke, B. (1997). Molecular virology of hepatitis C virus. *Journal of General Virology* **78**, 2397-2410.
- Cleaves, G. R., Ryan, T. E. & Schlesinger, R. W. (1981). Identification and characterization of type-2 dengue virus replicative intermediate and replicative form RNAs. *Virology* **111**, 73-83.

- Cocquerel, L., Meunier, J. C., Pillez, A., Wychowski, C. & Dubuisson, J. (1998). A retention signal necessary and sufficient for endoplasmic reticulum localization maps to the transmembrane domain of hepatitis C virus glycoprotein E2. *Journal of Virology* **72**, 2183-2191.
- Cocquerel, L., Duvet, S., Meunier, J. C., Pillez, A., Cacan, R., Wychowski, C. & Dubuisson, J. (1999). The transmembrane domain of hepatitis C virus glycoprotein E1 is a signal for static retention in the endoplasmic reticulum. *Journal of Virology* **73**, 2641-2649.
- Cole, N. B., Ellenberg, J., Song, J., DiEuliis, D. & Lippincott-Schwartz, J. (1998). Retrograde transport of Golgi-localized proteins to the ER. *Journal of Cell Biology* **140**, 1-15.
- Collett, M. S., Larson, R., Gold, C., Strick, D., Anderson, D. K. & Purchio, A. F. (1988). Molecular cloning and nucleotide sequence of the pestivirus bovine viral diarrhea virus. *Virology* **165**, 191-199.
- Collier, A. J., Tang, S. X. & Elliott, R. M. (1998). Translation efficiencies of the 5' untranslated region from representatives of the six major genotypes of hepatitis C virus using a novel bicistronic reporter assay system. *Journal of General Virology* **79**, 2359-2366.
- Cox, J. S., Shamu, C. E. & Walter, P. (1993). Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell* **73**, 1197-1206.
- Creighton, T. E. (1988). Disulfide bonds and protein stability. *Bioessays* **8**, 57-63.
- Davidson, F., Simmonds, P., Ferguson, J. C., Jarvis, L. M., Dow, B. C., Follett, E. A. C., Seed, C. R. G., Krusius, T., Lin, C., Medgyesi, G. A., Kiyokawa, H., Olim, G., Duraisamy, G., Cuypers, T., Saeed, A. A., Teo, D., Conradie, J., Kew, M. C., Lin, M., Nuchaprayoon, C., Ndimbie, O. K. & Yap, P. L. (1995). Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' noncoding region. *Journal of General Virology* **76**, 1197-1204.
- Davis, G. L., Balart, L. A., Schiff, E. R., Lindsay, K., Bodenheimer, H. C., Perrillo, R. P., Carey, W., Jacobson, I. M., Payne, J., Dienstag, J. L., vanThiel, D. H., Tamburro, C., Lefkowitz, J., Albrecht, J., Meschievitz, C., Ortego, T. J. & Gibas, A. (1989). Treatment of chronic hepatitis C with recombinant interferon alfa - a multicenter randomized, controlled trial. *New England Journal of Medicine* **321**, 1501-1506.
- Decastro, M., Sanchez, J., Herrera, J. F., Chaves, A., Duran, R., Garciabuey, L., Garciamonzon, C., Sequi, J. & Morenootero, R. (1993). Hepatitis C virus

- antibodies and liver disease in patients with porphyria cutanea tarda. *Hepatology* **17**, 551-557.
- DeFrancesco, R., Urbani, A., Nardi, M. C., Tomei, L., Steinkuhler, C. & Tramontano, A. (1996). A zinc binding site in viral serine proteinase. *Biochemistry* **41**, 13282-13287.
- Dejean, C., Berthoux, P., Cecillon, S. & Berthoux, F. (1999). Update on hepatitis G virus. *Presse Medicale* **28**, 1483-1487.
- Deleersnyder, V., Pillez, A., Wychowski, C., Blight, K., Xu, J., Hahn, Y. S., Rice, C. M. & Dubuisson, J. (1997). Formation of native hepatitis C virus glycoprotein complexes. *Journal of Virology* **71**, 697-704.
- Deng, R. T. & Brock, K. V. (1993). 5' and 3' untranslated regions of pestivirus genome - primary and secondary structure analyses. *Nucleic Acids Research* **21**, 1949-1957.
- Desilva, A., Braakman, I. & Helenius, A. (1993). Posttranslational folding of vesicular stomatitis virus G protein in the ER - involvement of noncovalent and covalent complexes. *Journal of Cell Biology* **120**, 647-655.
- Dhillon, A. P. & Dusheiko, G. M. (1995). Pathology of hepatitis C virus infection. *Histopathology* **26**, 297-309.
- DiBisceglie, A. M. (1997). Hepatitis C and hepatocellular carcinoma. *Hepatology* **26**, S34-S38.
- DiBisceglie, A. M. (1998). Hepatitis C. *Lancet* **351**, 351-355.
- DiBisceglie, A. M., Goodman, Z. D., Ishak, K. G., Hoofnagle, J. H., Melpolder, J. J. & Alter, H. J. (1991). Long term clinical and histopathological follow up of chronic posttransfusion hepatitis. *Hepatology* **14**, 969-974.
- Doms, R. W., Keller, D. S., Helenius, A. & Balch, W. E. (1987). Role For adenosine triphosphate in regulating the assembly and transport of vesicular stomatitis virus G protein trimers. *Journal of Cell Biology* **105**, 1957-1969.
- Doms, R. W., Lamb, R. A., Rose, J. K. & Helenius, A. (1993). Folding and assembly of viral membrane proteins. *Virology* **193**, 545-562.
- Dubuisson, J. & Rice, C. M. (1996). Hepatitis C virus glycoprotein folding: disulfide bond formation and association with calnexin. *Journal of Virology* **70**, 778-786.
- Dubuisson, J., Hsu, H. H., Cheung, R. C., Greenberg, H. B., Russell, D. G. & Rice, C. M. (1994). Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and Sindbis viruses. *Journal of Virology* **68**, 6147-6160.

- Duvet, S., Cocquerel, L., Pillez, A., Cacan, R., Verbert, A., Moradpour, D., Wychowski, C. & Dubuisson, J. (1998). Hepatitis C virus glycoprotein complex localization in the endoplasmic reticulum involves a determinant for retention and not retrieval. *Journal of Biological Chemistry* **273**, 32088-32095.
- Earl, P. L., Moss, B. & Doms, R. W. (1991). Folding, interaction with GRP78-Bip, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. *Journal of Virology* **65**, 2047-2055.
- Einfeld, D. A. & Hunter, E. (1997). Mutational analysis of the oligomer assembly domain in the transmembrane subunit of the Rous sarcoma virus glycoprotein. *Journal of Virology* **71**, 2383-2389.
- Eisenberg, D. (1999). Structural biology - how chaperones protect virgin proteins. *Science* **285**, 1021-1022.
- Elbers, K., Tautz, N., Becher, P., Stoll, D., Rumenapf, T. & Thiel, H. J. (1996). Processing in the pestivirus E2-NS2 region: identification of proteins p7 and E2p7. *Journal of Virology* **70**, 4131-4135.
- Elliott, J. G., Oliver, J. D. & High, S. (1997). The thiol-dependent reductase ERp57 interacts specifically with N-glycosylated integral membrane proteins. *Journal of Biological Chemistry* **272**, 13849-13855.
- Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Ogura, Y., Izumi, N., Marumo, F. & Sato, C. (1996). Mutations in the non-structural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *New England Journal of Medicine* **334**, 77-81.
- Erdoş, E. G. & Skidgel, R. A. (1989). Neutral endopeptidase 24.11 (enkephalinase) and related regulators of peptide hormones. *FASEB Journal* **3**, 145-151.
- Errington, W., Wardell, A. D., McDonald, S., Goldin, R. D. & McGarvey, M. J. (1999). Subcellular localisation of NS3 in HCV-infected hepatocytes. *Journal of Medical Virology* **59**, 456-462.
- Esteban, R. (1993). Epidemiology of hepatitis C virus infection. *Journal of Hepatology* **17**, S67-S71.
- Failla, C., Tomei, L. & DeFrancesco, R. (1994). Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus non-structural proteins. *Journal of Virology* **68**, 3753-3760.
- Fan, W. & Mason, P. W. (1990). Membrane association and secretion of the Japanese encephalitis virus NS1 protein from cells expressing NS1 cDNA. *Virology* **177**, 470-476.

- Feinstone, S. & Gust, D. (1997). Hepatitis A virus. In *Clinical Virology*, pp. 1049-1072. Edited by D. D. Richman, R.J. Whitely and F.G. Hayden. New York: Churchill Livingstone.
- Ferrari, C., Valli, A., Galati, L., Penna, A., Scaccaglia, P., Giuberti, T., Schianchi, C., Missale, G., Marin, M. G. & Fiaccadori, F. (1994). T-cell response to structural and non-structural hepatitis C virus antigens in persistent and self-limited hepatitis C virus infections. *Hepatology* **19**, 286-295.
- Ferrari, D. M. & Söling, H. D. (1999). The protein disulphide-isomerase family: unravelling a string of folds. *Biochemical Journal* **339**, 1-10.
- Fischer, G., Wittmannliedbold, B., Lang, K., Kiefhaber, T. & Schmid, F. X. (1989). Cyclophilin and peptidyl prolyl *cis-trans* isomerase are probably identical proteins. *Nature* **337**, 476-478.
- Flamand, M., Megret, F., Mathieu, M., Lepault, J., Rey, F.A. & Deubel, V. (1999). Dengue virus type 1 non-structural protein NS1 is secreted from mammalian cells as a soluble hexamer in a glycosylation-dependent fashion. *Journal of Virology* **73**, 6104-6110.
- Fleaton, M. N., Sheahan, B. J., Gould, E. A., Atkins, G. J. & Liljestrom, P. (1999). Recombinant Semliki Forest virus particles encoding the PrME or NS1 proteins of louping ill virus protect mice from lethal challenge. *Journal of General Virology* **80**, 1189-1198.
- Flint, M. & McKeating, J. A. (1999). The C-terminal region of the hepatitis C virus E1 glycoprotein confers localization within the endoplasmic reticulum. *Journal of General Virology* **80**, 1943-1947.
- Flint, M., Thomas, J. M., Maidens, C. M., Shotton, C., Levy, S., Barclay, W. S. & McKeating, J. A. (1999a). Functional analysis of cell surface-expressed hepatitis C virus E2 glycoprotein. *Journal of Virology* **73**, 6782-6790.
- Flint, M., Maidens, C. M., Loomis-Price, L. D., Shotton, C., Dubuisson, J., Monk, P., Higginbottom, A., Levy, S. & McKeating, J. A. (1999b). Characterisation of hepatitis C virus E2 glycoprotein interaction with a putative cellular receptor, CD81. *Journal of Virology* **73**, 6235-6244.
- Fonseca, B. A. L., Pincus, S., Shope, R. E., Paoletti, E. & Mason, P. W. (1994). Recombinant vaccinia viruses co-expressing dengue-1 glycoproteins PrM and E induce neutralizing antibodies in mice. *Vaccine* **12**, 279-285.
- Fournillier, A., Depla, E., Karayiannis, P., Vidalin, O., Maertens, G., Trepo, C. & Inchauspe, G. (1999). Expression of noncovalent hepatitis C virus envelope E1E2

- complexes is not required for the induction of antibodies with neutralizing properties following DNA immunization. *Journal of Virology* **73**, 7497-7504.
- Fournillier-Jacob, A., Cahour, A., Escriou, N., Girard, M. & Wychowski, C. (1996). Processing of the E1 glycoprotein of hepatitis C virus expressed in mammalian cells. *Journal of General Virology* **77**, 1055-1064.
- Frand, A. R. & Kaiser, C. A. (1998). The ERO1 gene of yeast is required for oxidation of protein dithiols in the endoplasmic reticulum. *Molecular Cell* **1**, 161-170.
- Frand, A. R. & Kaiser, C. A. (1999). Ero1p oxidises protein disulphide isomerase in a pathway for disulphide bond formation in the endoplasmic reticulum. *Molecular Cell* **4**, 469-477.
- Freedman, R. B., Hirst, T. R. & Tuite, M. F. (1994). Protein disulfide isomerase - building bridges in protein folding. *Trends in Biochemical Sciences* **19**, 331-336.
- Gale, M. J., Korth, M. J., Tang, N. M., Tan, S. L., Hopkins, D. A., Dever, T. E., Polyak, S. J., Gretch, D. R. & Katze, M. G. (1997). Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the non-structural 5A protein. *Virology* **230**, 217-227.
- Gale, M., Kwieciszewski, B., Dossett, M., Nakao, H. & Katze, M. G. (1999). Antiapoptotic and oncogenic potentials of hepatitis C virus are linked to interferon resistance by viral repression of the PKR protein kinase. *Journal of Virology* **73**, 6506-6516.
- Gallagher, P. J., Henneberry, J. M., Sambrook, J. F. & Gething, M. J. H. (1992). Glycosylation requirements for intracellular transport and function of the hemagglutinin of influenza virus. *Journal of Virology* **66**, 7136-7145.
- Gallinari, P., Brennan, D., Nardi, C., Brunetti, M., Tomei, L., Steinkuhler, C. & DeFrancesco, R. (1998). Multiple enzymatic activities associated with recombinant NS3 protein of hepatitis C virus. *Journal of Virology* **72**, 6758-6769.
- Gavrieli, Y., Sherman, Y. & Bensasson, S. A. (1992). Identification of programmed cell death *in situ* via specific labeling of nuclear DNA fragmentation. *Journal of Cell Biology* **119**, 493-501.
- Gilmore, R. (1991). The protein translocation apparatus of the rough endoplasmic reticulum, its associated proteins, and the mechanism of translocation. *Current Opinion in Cell Biology* **3**, 580-584.
- Gish, R. G. (1999). Standards of treatment in chronic hepatitis C. *Seminars in Liver Disease* **19**, 35-47.

- Glasgow, G. M., McGee, M. M., Tarbatt, C. J., Mooney, D. A., Sheahan, B. J. & Atkins, G. J. (1998). The Semliki Forest virus vector induces p53-independent apoptosis. *Journal of General Virology* **79**, 2405-2410.
- Gomez, J., Martell, M., Quer, J., Cabot, B. & Esteban, J. I. (1999). Hepatitis C viral quasispecies. *Journal of Viral Hepatitis* **6**, 3-16.
- Gonzalez-Peralta, R. P., Fang, J. W. S., Davis, G. L., Gish, R., Tsukiyamakohara, K., Kohara, M., Mondelli, M. U., Lesniewski, R., Phillips, M. I., Mizokami, M. & Lau, J. Y. N. (1994). Optimization for the detection of hepatitis C virus antigens in the liver. *Journal of Hepatology* **20**, 143-147.
- Gorbalenya, A. E., Donchenko, A. P., Koonin, E. V. & Blinov, V. M. (1989). N-terminal domains of putative helicases of flaviviruses and pestiviruses may be serine proteases. *Nucleic Acids Research* **17**, 3889-3897.
- Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M. & Rice, C. M. (1993a). Characterization of the hepatitis C virus-encoded serine proteinase - determination of proteinase-dependent polyprotein cleavage sites. *Journal of Virology* **67**, 2832-2843.
- Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M. & Rice, C. M. (1993b). Expression and identification of hepatitis C virus polyprotein cleavage products. *Journal of Virology* **67**, 1385-1395.
- Gray, E. W. & Nettleton, P. F. (1987). The ultrastructure of cell cultures infected with border disease and bovine virus diarrhea viruses. *Journal of General Virology* **68**, 2339-2346.
- Greenfield, J. P., Tsai, J., Gouras, G. K., Hai, B., Thinakaran, G., Checler, F., Sisodia, S. S., Greengard, P. & Xu, H. X. (1999). Endoplasmic reticulum and trans-Golgi network generate distinct populations of Alzheimer beta-amyloid peptides. *Proceedings of the National Academy of Sciences, USA* **96**, 742-747.
- Habersetzer, F., Fournillier, A., Dubuisson, J., Rosa, D., Abrignani, S., Wychowski, C., Nakano, I., Trepo, C., Desgranges, C. & Inchauspe, G. (1998). Characterization of human monoclonal antibodies specific to the hepatitis C virus glycoprotein E2 with *in vitro* binding neutralization properties. *Virology* **249**, 32-41.
- Hammond, C. & Helenius, A. (1995). Quality control in the secretory pathway. *Current Opinion in Cell Biology* **7**, 523-529.
- Hammond, C., Braakman, I. & Helenius, A. (1994). Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proceedings of the National Academy of Sciences, USA* **91**, 913-917.

- Hames, B. D. (1990). One-dimensional polyacrylamide gel electrophoresis. *Gel electrophoresis of protein - a practical approach*, pp1-139. Edited by B. D. Hames and D. Rickwood. Oxford University Press, Oxford.
- Han, J. H., Shyamala, V., Richman, K. H., Brauer, M. J., Irvine, B., Urdea, M. S., Tekampolson, P., Kuo, G., Choo, Q. L. & Houghton, M. (1991). Characterization of the terminal regions of hepatitis C viral RNA - identification of conserved sequences in the 5' untranslated region and poly(A) tails at the 3' end. *Proceedings of the National Academy of Sciences, USA* **88**, 1711-1715.
- Harding, H. P., Zhang, Y. H. & Ron, D. (1999). Protein translation and folding are coupled by an endoplasmic reticulum-resident kinase. *Nature* **398**, 90.
- Hebert, D. N., Zhang, J. X., Chen, W., Foellmer, B. & Helenius, A. (1997). The number and location of glycans on influenza hemagglutinin determine folding and association with calnexin and calreticulin. *Journal of Cell Biology* **139**, 613-623.
- Hedge, R. S. & Lingappa, V. R. (1997). Membrane protein biogenesis: regulated complexity at the endoplasmic reticulum. *Cell* **91**, 575-582.
- Heinz, F. X. (1986). Epitope mapping of flavivirus glycoproteins. *Advances in Virus Research* **31**, 103-123.
- Heinz, F. X. (1992). Comparative molecular biology of flaviviruses and hepatitis C virus. *Archives of Virology*, 163-171.
- Heinz, F. X., Stiasny, K., Puschnerauer, G., Holzmann, H., Allison, S. L., Mandl, C. W. & Kunz, C. (1994). Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the heterodimeric association with protein PrM. *Virology* **198**, 109-117.
- Heinz, F. X., Allison, S. L., Stiasny, K., Schalich, J., Holzmann, H., Mandl, C. W. & Kunz, C. (1995). Recombinant and virion-derived soluble and particulate immunogens for vaccination against tick-borne encephalitis. *Vaccine* **13**, 1636-1642.
- Helenius, A., Trombetta, E. S., Hebert, D. N. & Simons, J. F. (1997). Calnexin, calreticulin and the folding of glycoproteins. *Trends in Cell Biology* **7**, 193-200.
- Hellen, C. U. T. & Pestova, T. V. (1999). Translation of hepatitis C virus RNA. *Journal of Viral Hepatitis* **6**, 79-87.
- High, S. & Dobberstein, B. (1992). Mechanisms that determine the transmembrane disposition of proteins. *Current Opinion in Cell Biology* **4**, 581-586.
- Hijikata, M., Mizushima, H., Akagi, T., Mori, S., Kakiuchi, N., Kato, N., Tanaka, T., Kimura, K. & Shimotohno, K. (1993a). Two distinct proteinase activities required

- for the processing of a putative non-structural precursor protein of hepatitis C virus. *Journal of Virology* **67**, 4665-4675.
- Hijikata, M., Mizushima, H., Tanji, Y., Komoda, Y., Hirowatari, Y., Akagi, T., Kato, N., Kimura, K. & Shimotohno, K. (1993b). Proteolytic processing and membrane association of putative non-structural proteins of hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **90**, 10773-10777.
- Hiramatsu, N., Hayashi, N., Haruna, Y., Kasahara, A., Fusamoto, H., Mori, C., Fuke, I., Okayama, H. & Kamada, T. (1992). Immunohistochemical detection of hepatitis C virus-infected hepatocytes in chronic liver disease with monoclonal antibodies to core, envelope and NS3 region of hepatitis C virus genome. *Hepatology* **16**, 306-310.
- Hobman, T. C., Lemon, H. F. & Jewell, K. (1997). Characterization of an endoplasmic reticulum retention signal in the rubella virus E1 glycoprotein. *Journal of Virology* **71**, 7670-7680.
- Hobman, T. C., Woodward, L. & Farquhar, M. G. (1993). The rubella virus E2 and E1 spike glycoproteins are targeted to the Golgi complex. *Journal of Cell Biology* **121**, 269-281.
- Holst, B., Tachibana, C. & Winther, J. R. (1997). Active site mutations in yeast protein disulphide isomerase cause dithiothreitol sensitivity and a reduced rate of protein folding in the endoplasmic reticulum. *Journal of Cell Biology* **138**, 1229-1238.
- Honda, M., Rijnbrand, R., Abell, G., Kim, D. S. & Lemon, S. M. (1999). Natural variation in translational activities of the 5' nontranslated RNAs of hepatitis C virus genotypes 1a and 1b: evidence for a long range RNA-RNA interaction outside of the internal ribosomal entry site. *Journal of Virology* **73**, 4941-4951.
- Hope, R. G. & McLauchlan, J. (1999). Manuscript submitted for publication.
- Hosein, B., Fang, C. T., Popovsky, M. A., Ye, J., Zhang, M. L. & Wang, C. Y. (1991). Improved serodiagnosis of hepatitis C virus infection with synthetic peptide antigen from capsid protein. *Proceedings of the National Academy of Sciences, USA* **88**, 3647-3651.
- Houghton, M. (1996). Hepatitis C virus. *Fields Virology*, pp. 1035-1058. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. New York: Raven Press.
- Houghton, M., Weiner, A., Han, J., Kuo, G. & Choo, Q. L. (1991). Molecular biology of the hepatitis C viruses - implications for diagnosis, development and control of viral disease. *Hepatology* **14**, 381-388.

- Hsieh, T. Y., Matsumoto, M., Chou, H. C., Schneider, R., Hwang, S. B., Lee, A. S. & Lai, M. M. C. (1998). Hepatitis C virus core protein interacts with heterogeneous nuclear ribonucleoprotein K. *Journal of Biological Chemistry* **273**, 17651-17659.
- Huber, M. T. & Compton, T. (1999). Intracellular formation and processing of the heterotrimeric gH-gL-gO (gCIII) glycoprotein envelope complex of human cytomegalovirus. *Journal of Virology* **73**, 3886-3892.
- Hulst, M. M. & Moormann, R. J. M. (1997). Inhibition of pestivirus infection in cell culture by envelope proteins E-rns and E2 of classical swine fever virus: E-rns and E2 interact with different receptors. *Journal of General Virology* **78**, 2779-2787.
- Hulst, M. M., Panoto, F. E., Hoekman, A., vanGennip, H. G. P. & Moormann, R. J. M. (1998). Inactivation of the RNase activity of glycoprotein E-rns of classical swine fever virus results in a cytopathogenic virus. *Journal of Virology* **72**, 151-157.
- Hussy, P., Schmid, G., Mous, J. & Jacobsen, H. (1996). Purification and *in vitro*-phospholabeling of secretory envelope proteins E1 and E2 of hepatitis C virus expressed in insect cells. *Virus Research* **45**, 45-57.
- Hwang, C., Sinsky, A. J. & Lodish, H. F. (1992). Oxidized redox state of glutathione in the endoplasmic reticulum. *Science* **257**, 1496-1502.
- Idilman, R., DeMaria, N., Colantoni, A. & vanThiel, D. H. (1998). Pathogenesis of hepatitis B and C-induced hepatocellular carcinoma. *Journal of Viral Hepatitis* **5**, 285-299.
- Inchauspe, G. (1999). DNA vaccine strategies for hepatitis C. *Journal of Hepatology* **30**, 339-346.
- Inchauspe, G., Zebedee, S., Lee, D. H., Sugitani, M., Nasoff, M. & Prince, A. M. (1991). Genomic structure of the human prototype strain H of hepatitis C virus - comparison with American and Japanese isolates. *Proceedings of the National Academy of Sciences, USA* **88**, 10292-10296.
- Ishak, R., Tovey, D. G. & Howard, C. R. (1988). Morphogenesis of yellow fever virus 17d in infected cell cultures. *Journal of General Virology* **69**, 325-335.
- Isoyama, T., Kamoshita, N., Yasui, K., Iwai, A., Shiroki, K., Toyoda, H., Yamada, A., Takasaki, Y. & Nomoto, A. (1999). Lower concentration of La protein required for internal ribosome entry on hepatitis C virus RNA than on poliovirus RNA. *Journal of General Virology* **80**, 2319-2327.
- Ito, T. & Lai, M. M. C. (1997). Determination of the secondary structure of and cellular protein binding to the 3'-untranslated region of the hepatitis C virus RNA genome. *Journal of Virology* **71**, 8698-8706.

- Ito, T. & Lai, M. M. C. (1999). An internal polypyrimidine-tract-binding protein-binding site in the hepatitis C virus RNA attenuates translation, which is relieved by the 3'-untranslated sequence. *Virology* **254**, 288-296.
- Jackson, M. R., Nilsson, T. & Peterson, P. A. (1990a). Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO Journal* **9**, 3153-3162.
- Jackson, R. J., Howell, M. T. & Kaminski, A. (1990b). The novel mechanism of initiation of picornavirus RNA translation. *Trends in Biochemical Sciences* **15**, 477-483.
- Janknecht, R., Demartynoff, G., Lou, J., Hipskind, R. A., Nordheim, A. & Stunnenberg, H. G. (1991). Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Proceedings of the National Academy of Sciences, USA* **88**, 8972-8976.
- Johnson, R. J., Gretch, D. R., Yamabe, H., Hart, J., Bacchi, C. E., Hartwell, P., Couser, W. G., Corey, L., Wener, M. H., Alpers, C. E. & Willson, R. (1993). Membranoproliferative glomerulonephritis associated with hepatitis C virus infection. *New England Journal of Medicine* **328**, 465-470.
- Jons, A., Dijkstra, J. M. & Mettenleiter, T. C. (1998). Glycoproteins M and N of pseudorabies virus form a disulfide-linked complex. *Journal of Virology* **72**, 550-557.
- Kaito, M., Watanabe, S., Tsukiyamakohara, K., Yamaguchi, K., Kobayashi, Y., Konishi, M., Yokoi, M., Ishida, S., Suzuki, S. & Kohara, M. (1994). Hepatitis C virus particle detected by immunoelectron microscopic study. *Journal of General Virology* **75**, 1755-1760.
- Kao, J. H., Chen, P. J., Lei, M. Y., Wang, T. H. & Chen, D. S. (1993). Sexual transmission of HCV. *Lancet* **342**, 626.
- Kapoor, M., Zhang, L. W., Ramachandra, M., Kusukawa, J., Ebner, K. E. & Padmanabhan, R. (1995). Association between NS3 and NS5 proteins of dengue virus type 2 in the putative RNA replicase is linked to differential phosphorylation of NS5. *Journal of Biological Chemistry* **270**, 19100-19106.
- Kato, N., Ohkoshi, S. & Shimotohno, K. (1989). Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variations from the original isolate in the USA. *Proceedings of the Japan Academy Series B-Physical and Biological Sciences* **65**, 219-223.
- Kato, N., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T. & Shimotohno, K. (1990). Molecular cloning of the human hepatitis C virus genome

- from Japanese patients with non-A, non-B hepatitis. *Proceedings of the National Academy of Sciences, USA* **87**, 9524-9528.
- Kato, N., Ootsuyama, Y., Tanaka, T., Nakagawa, M., Nakazawa, T., Muraiso, K., Ohkoshi, S., Hijikata, M. & Shimotohno, K. (1992). Marked sequence diversity in the putative envelope proteins of hepatitis C viruses. *Virus Research* **22**, 107-123.
- Kato, N., Sekiya, H., Ootsuyama, Y., Nakazawa, T., Hijikata, M., Ohkoshi, S. & Shimotohno, K. (1993). Humoral immune response to hypervariable region 1 of the putative envelope glycoprotein (gp70) of hepatitis C virus. *Journal of Virology* **67**, 3923-3930.
- Kato, N., Ootsuyama, Y., Sekiya, H., Ohkoshi, S., Nakazawa, T., Hijikata, M. & Shimotohno, K. (1994). Genetic drift in hypervariable region 1 of the viral genome in persistent hepatitis C virus infection. *Journal of Virology* **68**, 4776-4784.
- Katze, M. G. (1995). Regulation of the interferon-induced PKR: can viruses cope? *Trends in Microbiology* **3**, 75-78.
- Kim, D. W., Gwack, Y., Han, J. H. & Choe, J. (1997a). Towards defining a minimal functional domain for NTPase and RNA helicase activities of the hepatitis C virus NS3 protein. *Virus Research* **49**, 17-25.
- Kim, D. W., Kim, J., Gwack, Y., Han, J. H. & Choe, J. (1997b). Mutational analysis of the hepatitis C virus RNA helicase. *Journal of Virology* **71**, 9400-9409.
- Kim, J. L., Morgenstern, K. A., Lin, C., Fox, T., Dwyer, M. D., Landro, J. A., Chambers, S. P., Markland, W., Lepre, C. A., Omalley, E. T., Harbeson, S. L., Rice, C. M., Murcko, M. A., Caron, P. R. & Thomson, J. A. (1996). Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* **87**, 343-355.
- Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A., Murcko, M. A., Lin, C. & Caron, P. R. (1998). Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure* **6**, 89-100.
- Kim, J. E., Song, W. K., Chung, K. M., Back, S. H. & Jang, S. K. (1999). Subcellular localization of hepatitis C viral proteins in mammalian cells. *Archives of Virology* **144**, 329-343.
- Kita, H., Moriyama, T., Kaneko, T., Harase, I., Nomura, M., Miura, H., Nakamura, I., Yazaki, Y. & Imawari, M. (1993). HLa B44-restricted cytotoxic T lymphocytes recognizing an epitope on hepatitis C virus nucleocapsid protein. *Hepatology* **18**, 1039-1044.

- Koch, J. O. & Bartenschlager, R. (1999). Modulation of hepatitis C virus NS5A hyperphosphorylation by non-structural proteins NS3, NS4A, and NS4B. *Journal of Virology* **73**, 7138-7146.
- Kocisko, D. A., Priola, S. A., Raymond, G. J., Chesebro, B., Lansbury, P. T. & Caughey, B. (1995). Species specificity in the cell-free conversion of prion protein to protease-resistant forms - a model for the scrapie species barrier. *Proceedings of the National Academy of Sciences, USA* **92**, 3923-3927.
- Koike, K., Moriya, K., Ishibashi, K., Matsuura, Y., Suzuki, T., Saito, I., Iino, S., Kurokawa, K. & Miyamura, T. (1995). Expression of hepatitis C virus envelope proteins in transgenic mice. *Journal of General Virology* **76**, 3031-3038.
- Kolykhalov, A. A., Agapov, E. V., Blight, K. J., Mihalik, K., Feinstone, S. M. & Rice, C. M. (1997). Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. *Science* **277**, 570-574.
- Kolykhalov, A. A., Feinstone, S. M. & Rice, C. M. (1996). Identification of a highly conserved sequence element at the 3' terminus of hepatitis C virus genome RNA. *Journal of Virology* **70**, 3363-3371.
- Konishi, E. & Mason, P. W. (1993). Proper maturation of the Japanese encephalitis virus envelope glycoprotein requires cosynthesis with the premembrane protein. *Journal of Virology* **67**, 1672-1675.
- Koonin, E. V. (1991a). The phylogeny of RNA-dependent RNA polymerases of positive strand RNA viruses. *Journal of General Virology* **72**, 2197-2206.
- Koonin, E. V. (1991b). Similarities in RNA helicases. *Nature* **352**, 290.
- Kornfeld, R. & Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. *Annual Review of Biochemistry* **54**, 631-664.
- Kuo, G., Choo, Q. L., Alter, H. J., Gitnick, G. L., Redeker, A. G., Purcell, R. H., Miyamura, T., Dienstag, J. L., Alter, M. J., Stevens, C. E., Tegtmeier, G. E., Bonino, F., Colombo, M., Lee, W. S., Kuo, C., Berger, K., Shuster, J. R., Overby, L. R., Bradley, D. W. & Houghton, M. (1989). An assay for circulating antibodies to a major etiologic virus of human non-A, non-B-hepatitis. *Science* **244**, 362-364.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 570-574.
- Lagging, L. M., Meyer, K., Owens, R. J. & Ray, R. (1998). Functional role of hepatitis C virus chimeric glycoproteins in the infectivity of pseudotyped virus. *Journal of Virology* **72**, 3539-3546.

- Lanford, R. E., Notvall, L., Chavez, D., White, R., Frenzel, G., Simonsen, C. & Kim, J. (1993). Analysis of hepatitis C virus capsid, E1, and E2/NS1 proteins expressed in insect cells. *Virology* **197**, 225-235.
- Lang, K., Schmid, F. X. & Fischer, G. (1987). Catalysis of protein folding by prolyl isomerase. *Nature* **329**, 268-270.
- Large, M. K., Kittlesen, D. J. & Hahn, Y. S. (1999). Suppression of host immune response by the core protein of hepatitis C virus: possible implications for hepatitis C virus persistence. *Journal of Immunology* **162**, 931-938.
- Laude, H. (1977). Improved method for the purification of hog cholera virus grown in tissue culture. *Archives of Virology* **54**; 41-51.
- Lavanchy, D., Purcell, R., Hollinger, F. B., Howard, C., Alberti, A., Kew, M., Dusheiko, G., Alter, M., Ayoola, E., Beutels, P., Bloomer, R., Ferret, B., Decker, R., Esteban, R., Fay, O., Fields, H., Fuller, E. C., Grob, P., Houghton, M., Leung, N., Locarnini, S. A., Margolis, H., Meheus, A., Miyamura, T., Mohamed, M. K., Tandon, B., Thomas, D., Head, H. T., Toukan, A. U., vanDamme, P., Zanetti, A., Arthur, R., Couper, M., Damelio, R., Emmanuel, J. C., Esteves, K., Gavinio, P., Griffiths, E., Hallaj, Z., Heuck, C. C., Heymann, D. L., Holck, S. E., Kane, M., Martinez, L. J., Meslin, F., Mochny, I. S., Ndikuyeze, A., Padilla, A. M., Rodier, G. R. M., Roure, C., Savage, F. & Vercauteren, G. (1999). Global surveillance and control of hepatitis C. *Journal of Viral Hepatitis* **6**, 35-47.
- Lee, C. & Chen, L. B. (1988). Dynamic behaviour of endoplasmic reticulum in living cells. *Cell* **54**, 37-46.
- Lee, K. J., Suh, Y. A., Cho, Y. G., Cho, Y. S., Ha, G. W., Chung, K. H., Hwang, J. H., Yun, Y. D., Lee, D. S., Kim, C. M. & Sung, Y. C. (1997). Hepatitis C virus E2 protein purified from mammalian cells is frequently recognized by E2-specific antibodies in patient sera. *Journal of Biological Chemistry* **272**, 30040-30046.
- Lesniewski, R. R., Boardway, K. M., Casey, J. M., Desai, S. M., Devare, S. G., Leung, T. K. & Mushahwar, I. K. (1993). Hypervariable 5'-terminus of hepatitis C virus E2/NS1 encodes antigenically distinct variants. *Journal of Medical Virology* **40**, 150-156.
- Li, K. J. & Garoff, H. (1996). Production of infectious recombinant Moloney murine leukemia virus particles in BHK cells using Semliki Forest virus-derived RNA expression vectors. *Proceedings of the National Academy of Sciences, USA* **93**, 11658-11663.

- Liberman, E., Fong, Y. L., Selby, M. J., Choo, Q. L., Cousens, L., Houghton, M. & Yen, T. S. B. (1999). Activation of the GRP78 and GRP94 promoters by hepatitis C virus E2 envelope protein. *Journal of Virology* **73**, 3718-3722.
- Liljestrom, P. & Garoff, H. (1991a). Internally located cleavable signal sequences direct the formation of Semliki Forest virus membrane proteins from a polyprotein precursor. *Journal of Virology* **65**, 147-154.
- Liljestrom, P. & Garoff, H. (1991b). A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. *Bio-Technology* **9**, 1356-1361.
- Liljestrom, P., Lusa, S., Huylebroeck, D. & Garoff, H. (1991). *In vitro* mutagenesis of a full-length cDNA clone of Semliki Forest virus - the small 6,000 molecular weight membrane protein modulates virus release. *Journal of Virology* **65**, 4107-4113.
- Lin, C., Lindenbach, B. D., Pragai, B. M., McCourt, D. W. & Rice, C. M. (1994). Processing in the hepatitis C virus E2-NS2 region - identification of p7 and 2 distinct E2-specific products with different C termini. *Journal of Virology* **68**, 5063-5073.
- Lin, C., Thomson, J. A. & Rice, C. M. (1995). A central region in the hepatitis C virus NS4A protein allows formation of an active NS3-NS4A serine proteinase complex *in vivo* and *in vitro*. *Journal of Virology* **69**, 4373-4380.
- Lin, C., Wu, J. W., Hsiao, K. & Su, M. S. S. (1997). The hepatitis C virus NS4A protein: interactions with the NS4B and NS5A proteins. *Journal of Virology* **71**, 6465-6471.
- Liu, Q. Y., Tackney, C., Bhat, R. A., Prince, A. M. & Zhang, P. (1997). Regulated processing of hepatitis C virus core protein is linked to subcellular localization. *Journal of Virology* **71**, 657-662.
- Lo, S. Y., Masiarz, F., Hwang, S. B., Lai, M. M. C. & Ou, J. H. (1995). Differential subcellular localization of hepatitis C virus core gene products. *Virology* **213**, 455-461.
- Lo, S. Y., Selby, M. J. & Ou, J. H. (1996). Interaction between hepatitis C virus core protein and E1 envelope protein. *Journal of Virology* **70**, 5177-5182.
- Lohmann, V., Korner, F., Herian, U. & Bartenschlager, R. (1997). Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *Journal of Virology* **71**, 8416-8428.
- Lohmann, V., Korner, F., Koch, J. O., Herian, U., Theilmann, L. & Bartenschlager, R. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* **285**, 110-113.

- Long, D., Wilcox, W. C., Abrams, W. R., Cohen, G. H. & Eisenberg, R. J. (1992). Disulphide bond structure of glycoprotein D of herpes simplex virus type-1 and type-2. *Journal of Virology* **66**, 6668-6685.
- Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw, E. W., Adachi, T. & Hostomska, Z. (1996). The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell* **87**, 331-342.
- Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw, E. W., Adachi, T., Margosiak, S., Dagostino, E. & Hostomska, Z. (1998). The conformation of hepatitis C virus NS3 proteinase with and without NS4A: a structural basis for the activation of the enzyme by its cofactor. *Clinical and Diagnostic Virology* **10**, 151-156.
- Maass, D. R. & Atkinson, P. H. (1994). Retention by the endoplasmic reticulum of rotavirus VP7 is controlled by 3 adjacent amino terminal residues. *Journal of Virology* **68**, 366-378.
- Macejak, D. G. & Sarnow, P. (1991). Internal initiation of translation mediated by the 5' leader of a cellular messenger RNA. *Nature* **353**, 90-94.
- Machamer, C. E. (1993). Targeting and retention of Golgi membrane proteins. *Current Opinion in Cell Biology* **5**, 606-612.
- Machamer, C. E. & Rose, J. K. (1988a). Influence of new glycosylation sites on expression of the vesicular stomatitis virus G protein at the plasma membrane. *Journal of Biological Chemistry* **263**, 5948-5954.
- Machamer, C. E. & Rose, J. K. (1988b). Vesicular stomatitis virus G proteins with altered glycosylation sites display temperature-sensitive intracellular transport and are subject to aberrant intermolecular disulfide bonding. *Journal of Biological Chemistry* **263**, 5955-5960.
- Machamer, C. E., Grim, M. G., Esquela, A., Chung, S. W., Rolls, M., Ryan, K. & Swift, A. M. (1993). Retention of a *cis* Golgi protein requires polar residues on one face of a predicted alpha helix in the transmembrane domain. *Molecular Biology of the Cell* **4**, 695-704.
- Macpherson, I. & Stoker, M. (1962). Polyoma transformation of hamster cell clones: an investigation of genetic factors affecting cell competence. *Virology* **16**, 147-151.
- Magrin, S., Craxi, A., Fabiano, C., Simonetti, R. G., Fiorentino, G., Marino, L., Diquattro, O., Dimarco, V., Loiacono, O., Volpes, R., Almasio, P., Urdea, M. S., Neuwald, P., Sanchezpescador, R., Detmer, J., Wilber, J. C. & Pagliaro, L. (1994). Hepatitis C

- viremia in chronic liver disease - relationship to interferon alpha or corticosteroid treatment. *Hepatology* **19**, 273-279.
- Mamiya, N. & Worman, H. J. (1999). Hepatitis C virus core protein binds to a DEAD box RNA helicase. *Journal of Biological Chemistry* **274**, 15751-15756.
- Mandl, C. W., Holzmann, H., Kunz, C. & Heinz, F. X. (1993). Complete genomic sequence of Powassan virus - evaluation of genetic elements in tick-borne versus mosquito-borne flaviviruses. *Virology* **194**, 173-184.
- Margottin, F., Bour, S. P., Durand, H., Selig, L., Benichou, S., Richard, V., Thomas, D., Strebel, K. & Benarous, R. (1998). A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Molecular Cell* **1**, 565-574.
- Markovic, I., Pulyaeva, H., Sokoloff, A. & Chernomordik, L. V. (1998). Membrane fusion mediated by baculovirus gp64 involves assembly of stable gp64 trimers into multiprotein aggregates. *Journal of Cell Biology* **143**, 1155-1166.
- Marusawa, H., Hijikata, M., Chiba, T. & Shimotohno, K. (1999). Hepatitis C virus core protein inhibits *Fas*- and tumor necrosis factor alpha-mediated apoptosis via NF-kappa B activation. *Journal of Virology* **73**, 4713-4720.
- Mason, P. W., Pincus, S., Fournier, M. J., Mason, T. L., Shope, R. E. & Paoletti, E. (1991). Japanese encephalitis virus-vaccinia recombinants produce particulate forms of the structural membrane proteins and induce high levels of protection against lethal JEV infection. *Virology* **180**, 294-305.
- Matsumoto, M., Hwang, S. B., Jeng, K. S., Zhu, N. L. & Lai, M. M. C. (1996). Homotypic interaction and multimerization of hepatitis C virus core protein. *Virology* **218**, 43-51.
- Matsumoto, M., Hsieh, T. Y., Zhu, N. L., vanArsdale, T., Hwang, S. B., Jeng, K. S., Gorbalenya, A. E., Lo, S. Y., Ou, J. H., Ware, C. F. & Lai, M. M. C. (1997). Hepatitis C virus core protein interacts with the cytoplasmic tail of lymphotoxin-beta receptor. *Journal of Virology* **71**, 1301-1309.
- Matsuura, Y., Suzuki, T., Suzuki, R., Sato, M., Aizaki, H., Saito, I. & Miyamura, T. (1994). Processing of E1 and E2 glycoproteins of hepatitis C virus expressed in mammalian and insect Cells. *Virology* **205**, 141-150.
- Maynard, J. E. (1990). Hepatitis B - global importance and need for control. *Vaccine* **8**, S18-S20.
- McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., McNab, D., Perry, L. J., Scott, J. E. & Taylor, P. (1988). The complete DNA sequence of the

- long unique region in the genome of herpes simplex virus type 1. *Journal of General Virology* **69**, 1531-1574.
- Meerovitch, K., Svitkin, Y. V., Lee, H. S., Lejbkowitz, F., Kenan, D. J., Chan, E. K. L., Agol, V. I., Keene, J. D. & Sonenberg, N. (1993). La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *Journal of Virology* **67**, 3798-3807.
- Meunier, J. C., Fournillier, A., Choukhi, A., Cahour, A., Cocquerel, L., Dubuisson, J. & Wychowski, C. (1999). Analysis of the glycosylation sites of hepatitis C virus (HCV) glycoprotein E1 and the influence of E1 glycans on the formation of the HCV glycoprotein complex. *Journal of General Virology* **80**, 887-896.
- Meyer, D. I., Krause, E. & Dobberstein, B. (1982). Secretory protein translocation across membranes - the role of the docking protein. *Nature* **297**, 647-650.
- Michalak, J. P., Wychowski, C., Choukhi, A., Meunier, J. C., Ung, S., Rice, C. M. & Dubuisson, J. (1997). Characterization of truncated forms of hepatitis C virus glycoproteins. *Journal of General Virology* **78**, 2299-2306.
- Miller, R. H. & Purcell, R. H. (1990). Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. *Proceedings of the National Academy of Sciences, USA* **87**, 2057-2061.
- Mizushima, H., Hijikata, M., Asabe, S-I., Hirota, M., Kimura, K. & Shimotohno, K. (1994). Two hepatitis C virus glycoprotein E2 products with different C termini. *Journal of Virology* **68**, 6215-6222.
- Monazahian, M., Bohme, I., Bonk, S., Koch, A., Scholz, C., Grethe, S. & Thomssen, R. (1999). Low density lipoprotein receptor as a candidate receptor for hepatitis C virus. *Journal of Medical Virology* **57**, 223-229.
- Moormann, R. J. M. & Hulst, M. M. (1988). Hog cholera virus - identification and characterization of the viral RNA and the virus-specific RNA synthesized in infected swine kidney cells. *Virus Research* **11**, 281-291.
- Moradpour, D., Wakita, T., Wands, J. R. & Blum, H. E. (1998). Tightly regulated expression of the entire hepatitis C virus structural region in continuous human cell lines. *Biochemical and Biophysical Research Communications* **246**, 920-924.
- Mori, H. & Christensen, A. K. (1980). Morphometric analysis of Leydig cells in the normal rat testis. *Journal of Cell Biology* **87**, 340-354.

- Moriya, K., Yotsuyanagi, H., Shintani, Y., Fujie, H., Ishibashi, K., Matsuura, Y., Miyamura, T. & Koike, K. (1997). Hepatitis C virus core protein induces hepatic steatosis in transgenic mice. *Journal of General Virology* **78**, 1527-1531.
- Moriya, K., Fujie, H., Shintani, Y., Yotsuyanagi, H., Tsutsumi, T., Ishibashi, K., Matsuura, Y., Kimura, S., Miyamura, T. & Koike, K. (1998). Hepatitis C virus core protein induces hepatocellular carcinoma in transgenic mice. *Nature Medicine* **4**, 1065-1067.
- Moss, B. (1996). *Poxviridae: The viruses and their replication*. In *Fields Virology*, pp. 2637-2671. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. New York: Raven Press.
- Munro, S. (1991). Sequences within and adjacent to the transmembrane segment of alpha-2,6-sialyltransferase specify Golgi retention. *EMBO Journal* **10**, 3577-3588.
- Muylaert, I. R., Chambers T. J., Galler, R. & Rice, C. M. (1996). Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence. *Virology* **222**, 159-168.
- Muylaert, I. R., Galler, R. & Rice, C. M. (1997). Genetic analysis of the yellow fever virus NS1 protein: identification of a temperature-sensitive mutation which blocks RNA accumulation. *Journal of Virology* **71**, 291-298.
- Nilsson, T. & Warren, G. (1994). Retention and retrieval in the endoplasmic reticulum and the Golgi apparatus. *Current Opinion in Cell Biology* **6**, 517-521.
- Nouri Aria, K.T., Sallie, R. & Sanger, D. (1993). Detection of genomic and intermediate replicative strands of hepatitis C virus in liver tissue by *in situ* hybridisation. *Journal of Clinical Investigation* **91**, 2226-2234.
- Nowak, T., Farber, P. M. & Wengler, G. (1989). Analyses of the terminal sequences of West Nile virus structural proteins and of the *in vitro* translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. *Virology* **169**, 365-376.
- Nuovo, G. J., Lidonnici, K., MacConnell, P. & Lane, B. (1993). Intracellular localisation of polymerase chain amplification of hepatitis C cDNA. *American Journal of Surological Pathology* **17**, 683-690.
- Ogata, N., Alter, H. J., Miller, R. H. & Purcell, R. H. (1991). Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **88**, 3392-3396.

- Oh, J. W., Ito, T. & Lai, M. M. C. (1999). A recombinant hepatitis C virus RNA-dependent RNA polymerase capable of copying the full-length viral RNA. *Journal of Virology* **73**, 7694-7702.
- Ohto, H., Terazawa, S., Sasaki, N., Hino, K., Ishiwata, C., Kako, M., Ujiie, N., Endo, C., Matsui, A., Okamoto, H., Mishiro, S., Kojima, M., Aikawa, T., Shimoda, K., Sakamoto, M., Akahane, Y., Yoshizawa, H., Tanaka, T., Tokita, H. & Tsuda, F. (1994). Transmission of hepatitis C virus from mothers to infants. *New England Journal of Medicine* **330**, 744-750.
- Okamoto, H., Kojima, M., Sakamoto, M., Iizuka, H., Hadiwandowo, S., Suwignyo, S., Miyakawa, Y. & Mayumi, M. (1994a). The entire nucleotide sequence and classification of a hepatitis C virus isolate of a novel genotype from an Indonesian patient with chronic liver-disease. *Journal of General Virology* **75**, 629-635.
- Okamoto, H., Mishiro, S., Tokita, H., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1994b). Superinfection of chimpanzees carrying hepatitis C virus of genotype Ii/1b with that of genotype Iii/2a or I/1a. *Hepatology* **20**, 1131-1136.
- Opstelten, D. J. E., Wallin, M. & Garoff, H. (1998). Moloney murine leukemia virus envelope protein subunits, gp70 and Pr15E, form a stable disulfide-linked complex. *Journal of Virology* **72**, 6537-6545.
- Owsianka, A. M. & Patel, A. H. (1999). Hepatitis C virus core protein interacts with a human DEAD box protein DDX3. *Virology* **257**, 330-340.
- Park, Y. N., Abe, K., Li, H., Hsuih, T., Thung, S. N. & Zhang, D. Y., (1996). Detection of hepatitis C virus RNA using ligation-dependent polymerase chain reaction in formalin-fixed, paraffin-embedded liver tissues. *American Journal of Pathology* **149**, 1485-1491.
- Parodi, A. J. (1999). Reglucosylation of glycoproteins and quality control of glycoprotein folding in the endoplasmic reticulum of yeast cells. *Biochimica et Biophysica Acta-General Subjects* **1426**, 287-295.
- Pease, R.J. and Leiper, J.M. (1996). Regulation of hepatic apolipoprotein-B-containing lipoprotein secretion. *Current Opinion in Lipidology* **7**; 132-138
- Pelham, H. R. B. (1995). Sorting and retrieval between the endoplasmic reticulum and Golgi apparatus. *Current Opinion in Cell Biology* **7**, 530-535.
- Persson, R. & Pettersson, R. F. (1991). Formation and intracellular transport of a heterodimeric viral spike protein complex. *Journal of Cell Biology* **112**, 257-266.
- Pestova, T. V., Shatsky, I. N. & Hellen, C. U. T. (1996). Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G

- subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Molecular and Cellular Biology* **16**, 6870-6878.
- Pileri, P., Uematsu, Y., Campagnoli, S., Galli, G., Falugi, F., Petracca, R., Weiner, A. J., Houghton, M., Rosa, D., Grandi, G. & Abrignani, S. (1998). Binding of hepatitis C virus to CD81. *Science* **282**, 938-941.
- Pincus, S., Mason, P. W., Konishi, E., Fonseca, B. A. L., Shope, R. E., Rice, C. M. & Paoletti, E. (1992). Recombinant vaccinia virus producing the PrM and E proteins of yellow fever virus protects mice from lethal yellow fever encephalitis. *Virology* **187**, 290-297.
- Plempner, R. K. & Wolf, D. H. (1999). Retrograde protein translocation: ERADication of secretory proteins in health and disease. *Trends in Biochemical Sciences* **24**, 266-270.
- Pollard, M. G., Travers, K. J. & Weissman, J. S. (1998). Ero1p: a novel and ubiquitous protein with an essential role in oxidative protein folding in the endoplasmic reticulum. *Molecular Cell* **1**, 171-182.
- Poole, T. L., Wang, C. Y., Popp, R. A., Potgieter, L. N. D., Siddiqui, A. & Collett, M. S. (1995). Pestivirus translation initiation occurs by internal ribosome entry. *Virology* **206**, 750-754.
- Porath, J., Carlsson, J., Olsson, I. & Berfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598-599.
- Poynard, T., Leroy, V., Cohard, M., Thevenot, T., Mathurin, P., Opolon, P. & Zarski, J. P. (1996). Meta-analysis of interferon randomized trials in the treatment of viral hepatitis C: effects of dose and duration. *Hepatology* **24**, 778-789.
- Prabakaran, D., Kim, P. S., Dixit, V. M. & Arvan, P. (1996). Oligomeric assembly of thrombospondin in the endoplasmic reticulum of thyroid epithelial cells. *European Journal of Cell Biology* **70**, 134-141.
- Prince, A. M., Brotman, B. & Grady G. (1974). Long incubation post-transfusion hepatitis without serological evidence of exposure to hepatitis B virus. *Lancet* **2**, 241-246.
- Prince, A. M., HuimaByron, T., Parker, T. S. & Levine, D. M. (1996). Visualization of hepatitis C virions and putative defective interfering particles isolated from low-density lipoproteins. *Journal of Viral Hepatitis* **3**, 11-17.
- Pryor, M. J., Gualano, R. C., Lin, B., Davidson, A. D. & Wright, P. J. (1998). Growth restriction of dengue virus type 2 by site-specific mutagenesis of virus-encoded glycoproteins. *Journal of General Virology* **79**, 2631-2639.

- Puntoriero, G., Meola, A., Lahm, A., Zucchelli, S., Ercole, B. B., Tafi, R., Pezzanera, M., Mondelli, M. U., Cortese, R., Tramontano, A., Galfre, G. & Nicosia, A. (1998). Towards a solution for hepatitis C virus hypervariability: mimotopes of the hypervariable region 1 can induce antibodies cross-reacting with a large number of viral variants. *EMBO Journal* **17**, 3521-3533.
- Ralston, R., Thudium, K., Berger, K., Kuo, C., Gervase, B., Hall, J., Selby, M., Kuo, G., Houghton, M. & Choo, Q. L. (1993). Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. *Journal of Virology* **67**, 6753-6761.
- Randolph, V.B., Winkler, G. & Stollar, V. (1990). Acidotropic amines inhibit proteolytic processing of flavivirus PrM protein. *Virology* **174**, 450-458.
- Rapoport, T. A. (1992). Transport of proteins across the endoplasmic reticulum membrane. *Science* **258**, 931-936.
- Rathmell, J. C. & Thompson, C. B. (1999). The central effectors of cell death in the immune system. *Annual Review of Immunology* **17**, 781-828.
- Ray, R. B., Lagging, L. M., Meyer, K. & Ray, R. (1996). Hepatitis C virus core protein cooperates with *ras* and transforms primary rat embryo fibroblasts to tumorigenic phenotype. *Journal of Virology* **70**, 4438-4443.
- Reddy, P. S. & Corley, R. B. (1998). Assembly, sorting, and exit of oligomeric proteins from the endoplasmic reticulum. *Bioessays* **20**, 546-554.
- Reed, K. E., Grakoui, A. & Rice, C. M. (1995). Hepatitis C virus-encoded NS2-3 protease - cleavage site mutagenesis and requirements for bimolecular cleavage. *Journal of Virology* **69**, 4127-4136.
- Reid, A. & Dienstag, J. (1997). Viral hepatitis. In *Clinical Virology*, pp. 69-86. Edited by D. D. Richman, R.J. Whitely and F.G. Hayden. New York: Churchill Livingstone.
- Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C. & Harrison, S. C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2 angstrom resolution. *Nature* **375**, 291-298.
- Reyes, G. R., Purdy, M. A., Kim, J. P., Luk, K. C., Young, L. M., Fry, K. E. & Bradley, D. W. (1990). Isolation of a cDNA from the virus responsible for enterically transmitted non-A, non-B hepatitis. *Science* **247**, 1335-1339.
- Reynolds, J. E., Kaminski, A., Kettinen, H. J., Grace, K., Clarke, B. E., Carroll, A. R., Rowlands, D. J. & Jackson, R. J. (1995). Unique features of internal initiation of hepatitis C virus RNA translation. *EMBO Journal* **14**, 6010-6020.

- Rice, C. M. (1996). *Flaviviridae: the viruses and their replication*. In *Fields Virology*, pp 931-959. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott - Raven Publishers.
- Rice, P. S., Smith, D. B., Simmonds, P. & Holmes, E. (1993). Heterosexual transmission of hepatitis C virus. *Lancet* **342**, 1052-1053.
- Rietsch, A. & Beckwith, J. (1998). The genetics of disulfide bond metabolism. *Annual Review of Genetics* **32**, 163-184.
- Robertson, B., Myers, G., Howard, C., Brettin, T., Bukh, J., Gaschen, B., Gojobori, T., Maertens, G., Mizokami, M., Nainan, O., Netesov, S., Nishioka, K., Shin-i, T., Simmonds, P., Smith, D., Stuyver, L. & Weiner, A. (1998). Classification, nomenclature, and database development of hepatitis C virus (HCV) and related viruses: proposals for standardization. *Archives of Virology* **143**, 2493-2503.
- Rolls, M. M., Webster, P., Balba, N. H. & Rose, J. K. (1994). Novel infectious particles generated by expression of the vesicular stomatitis virus glycoprotein from a self-replicating RNA. *Cell* **79**, 497-506.
- Rosa, D., Campagnoli, S., Moretto, C., Guenzi, E., Cousens, L., Chin, M., Dong, C., Weiner, A. J., Lau, J. Y. N., Choo, Q. L., Chien, D., Pileri, P., Houghton, M. & Abrignani, S. (1996). A quantitative test to estimate neutralizing antibodies to the hepatitis C virus: cytofluorimetric assessment of envelope glycoprotein 2 binding to target cells. *Proceedings of the National Academy of Sciences, USA* **93**, 1759-1763.
- Ruggieri, A., Harada, T., Matsuura, Y. & Miyamura, T. (1997). Sensitization to Fas-mediated apoptosis by hepatitis C virus core protein. *Virology* **229**, 68-76.
- Rumenapf, T., Unger, G., Strauss, J. H. & Thiel, H. J. (1993). Processing of the envelope glycoproteins of pestiviruses. *Journal of Virology* **67**, 3288-3294.
- Ryan, M. D., Monaghan, S. & Flint, M. (1998). Virus-encoded proteinases of the *Flaviviridae*. *Journal of General Virology* **79**, 947-959.
- Saeed, A. A., Fairclough, D., Aladmawi, A. M., Bacchus, R., Alrased, A. M. & Waller, D. K. (1991). High prevalence of HCV antibody among Egyptian blood donors. *Annals of Saudi Medicine* **11**, 591-592.
- Saito, T., Ishiguro, K., Onuki, R., Nagai, Y., Kishimoto, T. & Hisanaga, S. (1998). Okadaic acid-stimulated degradation of p35, an activator of CDK5, by proteasome in cultured neurons. *Biochemical and Biophysical Research Communications* **252**, 775-778.

- Santolini, E., Migliaccio, G. & Lamonica, N. (1994). Biosynthesis and biochemical properties of the hepatitis C virus core protein. *Journal of Virology* **68**, 3631-3641.
- Santolini, E., Pacini, L., Fipaldini, C., Migliaccio, G. & Lamonica, N. (1995). The NS2 protein of hepatitis C virus is a transmembrane polypeptide. *Journal of Virology* **69**, 7461-7471.
- Scales, S. J., Pepperkok, R. & Kreis, T. E. (1997). Visualization of ER-to-Golgi transport in living cells reveals a sequential mode of action for COPII and COPI. *Cell* **90**, 1137-1148.
- Schneider, R., Unger, G., Stark, R., Schneiderscherzer, E. & Thiel, H. J. (1993). Identification of a structural glycoprotein of an RNA virus as a ribonuclease. *Science* **261**, 1169-1171.
- Schreiber, G. B., Busch, M. P., Kleinman, S. H. & Korelitz, J. J. (1996). The risk of transfusion-transmitted viral infections. *New England Journal of Medicine* **334**, 1685-1690.
- Schweizer, A., Rohrer, J., Slot, J. W., Geuze, H. J. & Kornfeld, S. (1995). Reassessment of the subcellular localization of p63. *Journal Of Cell Science* **108**, 2477-2485.
- Scott, A., Leary, T. P., Simons, J. N., *et al.*, (1995). Genomic organisation of GB viruses A and B: two new members of the Flaviviridae associated with GB agent hepatitis. *Journal of Virology* **69**, 5621.
- Selby, M. J., Glazer, E., Masiarz, F. & Houghton, M. (1994). Complex processing and protein-protein interactions in the E2-NS2 region of HCV. *Virology* **204**, 114-122.
- Shamu, C. E., Cox, J. S. & Walter, P. (1994). The unfolded-protein-response pathway in yeast. *Trends in Cell Biology* **4**, 56-60.
- Shenkman, M., Ayalon, M. & Lederkremer, G. Z. (1997). Endoplasmic reticulum quality control of asialoglycoprotein receptor H2a involves a determinant for retention and not retrieval. *Proceedings of the National Academy of Sciences, USA* **94**, 11363-11368.
- Shih, C. M., Lo, S. C. J., Miyamura, T., Chen, S. Y. & Lee, Y. H. W. (1993). Suppression of hepatitis B virus expression and replication by hepatitis C virus core protein in Huh-7 cells. *Journal of Virology* **67**, 5823-5832.
- Shih, C. M., Chen, C. M., Chen, S. Y. & Lee, Y. H. W. (1995). Modulation of the trans-suppression activity of hepatitis C virus core protein by phosphorylation. *Journal of Virology* **69**, 1160-1171.

- Shimizu, Y. K., Iwamoto, A., Hijikata, M., Purcell, R. H. & Yoshikura, H. (1992). Evidence for *in vitro* replication of hepatitis C virus genome in a human T-cell line. *Proceedings of the National Academy of Sciences, USA* **89**, 5477-5481.
- Shimizu, Y. K., Hijikata, M., Iwamoto, A., Alter, H. J., Purcell, R. H. & Yoshikura, H. (1994). Neutralizing antibodies against hepatitis C virus and the emergence of neutralization escape mutant viruses. *Journal of Virology* **68**, 1494-1500.
- Shimizu, Y. K., Feinstone, S. M., Kohara, M., Purcell, R. H. & Yoshikura, H. (1996). Hepatitis C virus: detection of intracellular virus particles by electron microscopy. *Hepatology* **23**, 205-209.
- Shimotohno, K. & Feinstone, S. M. (1997). Hepatitis C virus and hepatitis G virus. In *Clinical Virology*, 1st Edition edn, pp. 1187-1215. Edited by D. D. Richman, R.J. Whitley and F.G. Hayden. New York: Churchill Livingstone Inc.
- Sidrauski, C., Chapman, R. & Walter, P. (1998). The unfolded protein response: an intracellular signalling pathway with many surprising features. *Trends in Cell Biology* **8**, 245-249.
- Siemoneit, K., Cardoso, M. D., Koerner, K., Wolpl, A. & Kubanek, B. (1995). Human monoclonal antibodies for the immunological characterization of a highly conserved protein domain of the hepatitis C virus glycoprotein E1. *Clinical and Experimental Immunology* **101**, 278-283.
- Sifers, R. N. (1995). Defective protein-folding as a cause of disease. *Nature Structural Biology* **2**, 355-357.
- Simmonds, P. (1995). Variability of hepatitis C virus. *Hepatology* **21**, 570-583.
- Smedile, A., Niro, G. & Rizzetto, M. (1997). Hepatitis D virus. In *Clinical Virology*, pp. 1273-1284. Edited by D. D. Richman, R.J. Whitley and F.G. Hayden. New York: Churchill Livingstone.
- Smith, G.W. & Wright, P.J. (1985). Synthesis of proteins and glycoproteins in dengue type 2 virus-infected Vero and *Aedes albopictus* cells. *Journal of General Virology* **66**, 559-571.
- Smith, S. & Blobel, G. (1993). The first membrane spanning region of the lamin B receptor is sufficient for sorting to the inner nuclear membrane. *Journal of Cell Biology* **120**, 631-637.
- Song, J., Fujii, M., Wang, F., Itoh, M. & Hotta, H. (1999). The NS5A protein of hepatitis C virus partially inhibits the antiviral activity of interferon. *Journal of General Virology* **80**, 879-886.

- Spaete, R. R., Alexander, D., Rugroden, M. E., Choo, Q. L., Berger, K., Crawford, K., Kuo, C., Leng, S., Lee, C., Ralston, R., Thudium, K., Tung, J. W., Kuo, G. & Houghton, M. (1992). Characterization of the hepatitis C virus E2/NS1 gene product expressed in mammalian cells. *Virology* **188**, 819-830.
- Speed, M. A., Wang, D. I. C. & King, J. (1995). Multimeric intermediates in the pathway to the aggregated inclusion body state for p22 tailspike polypeptide chains. *Protein Science* **4**, 900-908.
- Speed, M. A., Wang, D. I. C. & King, J. (1996). Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. *Nature Biotechnology* **14**, 1283-1287.
- Stadler, K., Allison, S. L., Schlich, J. & Heinz, F. X. (1997). Proteolytic activation of tick-borne encephalitis virus by furin. *Journal of Virology* **71**, 8475-8481.
- Steinkuhler, C., Urbani, A., Tomei, L., Biasiol, G., Sardana, M., Bianchi, E., Pessi, A. & DeFrancesco, R. (1996). Activity of purified hepatitis C virus protease NS3 on peptide substrates. *Journal of Virology* **70**, 6694-6700.
- Stewart, S. R. & Semler, B. L. (1997). RNA determinants of picornavirus cap-independent translation initiation. *Seminars in Virology* **8**, 242-255.
- Stiasny, K., Allison, S. L., Marchler-Bauer, A., Kunz, C. & Heinz, F. X. (1996). Structural requirements for low-pH-induced rearrangements in the envelope glycoprotein of tick-borne encephalitis virus. *Journal of Virology* **70**, 8142-8147.
- Suzuki, T., Yan, Q. & Lennarz, W. J. (1998). Complex, two-way traffic of molecules across the membrane of the endoplasmic reticulum. *Journal of Biological Chemistry* **273**, 10083-10086.
- Tai, C. L., Chi, W. K., Chen, D. S. & Hwang, L. H. (1996). The helicase activity associated with hepatitis C virus non-structural protein 3 (NS3). *Journal of Virology* **70**, 8477-8484.
- Takamizawa, A., Mori, C., Fuke, I., Manabe, S., Murakami, S., Fujita, J., Onishi, E., Andoh, T., Yoshida, I. & Okayama, H. (1991). Structure and organization of the hepatitis C virus genome isolated from human carriers. *Journal of Virology* **65**, 1105-1113.
- Tanaka, T., Kato, N., Cho, M. J. & Shimotohno, K. (1995). A novel sequence found at the 3'-terminus of hepatitis C virus genome. *Biochemical and Biophysical Research Communications* **215**, 744-749.

- Tanaka, Y., Enomoto, N., Kojima, S., Tang, L., Goto, M., Marumo, F. & Sato, C. (1993). Detection of hepatitis C virus RNA in the liver by *in situ* hybridization. *Liver* **13**, 203-208.
- Tanji, Y., Kaneko, T., Satoh, S. & Shimotohno, K. (1995). Phosphorylation of hepatitis C virus-encoded non-structural protein NS5A. *Journal of Virology* **69**, 3980-3986.
- Tatu, U. & Helenius, A. (1995). Analyzing calnexin and its substrate complexes in the ER of CHO cells. *Molecular Biology of the Cell* **6**, 377.
- Tatu, U. & Helenius, A. (1997). Interactions between newly synthesized glycoproteins, calnexin and a network of resident chaperones in the endoplasmic reticulum. *Journal of Cell Biology* **136**, 555-565.
- Tatu, U., Braakman, I. & Helenius, A. (1993). Membrane glycoprotein folding, oligomerization and intracellular transport - effects of dithiothreitol in living cells. *EMBO Journal* **12**, 2151-2157.
- Taylor, D. R., Shi, S. T., Romano, P. R., Barber, G. N. & Lai, M. M. C. (1999). Inhibition of the interferon-inducible protein kinase PKR by HCV E2 protein. *Science* **285**, 107-110.
- Thaler, M. M., Park, C. K., Landers, D. V., Wara, D. W., Houghton, M., Veeremanwauters, G., Sweet, R. L. & Han, J. H. (1991). Vertical transmission of hepatitis C virus. *Lancet* **338**, 17-18.
- Thiel, H. J., Stark, R., Weiland, E., Rumenapf, T. & Meyers, G. (1991). Hog cholera virus - molecular composition of virions from a pestivirus. *Journal of Virology* **65**, 4705-4712.
- Towbin, H., Staehelin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Science, USA* **76**, 4350-4354.
- Townsley, F. M., Wilson, D.W. & Pelham, H.R.B. (1993). Mutational analysis of the human KDEL receptor - distinct structural requirements for Golgi retention, ligand binding and retrograde transport. *EMBO Journal* **12**, 2821-2829.
- Tsukiyamakohara, K., Iizuka, N., Kohara, M. & Nomoto, A. (1992). Internal ribosome entry site within hepatitis C virus RNA. *Journal of Virology* **66**, 1476-1483.
- van Meer, G. (1989). Lipid traffic in animal cells. *Annual Review of Cell Biology* **5**, 247-275.
- Verde, C., Pascale, M. C., Martire, G., Lotti, L. V., Torrisi, M. R., Helenius, A. & Bonatti, S. (1995). Effect of ATP depletion and DTT on the transport of membrane

- proteins from the endoplasmic reticulum and the intermediate compartment to the Golgi complex. *European Journal of Cell Biology* **67**, 267-274.
- Vogel, F., Hartmann, E., Gorlich, D. & Rapoport, T.A. (1990). Segregation of the signal sequence receptor protein in the rough endoplasmic reticulum membrane. *European Journal of Cell Biology* **53**, 197-202.
- Walker, F. M., Dazza, M. C., Dauge, M. C., Boucher, O., Bedel, C., Henin, D. & Lehy, T. (1998). Detection and localization by *in situ* molecular biology techniques and immunohistochemistry of hepatitis C virus in livers of chronically infected patients. *Journal of Histochemistry & Cytochemistry* **46**, 653-660.
- Wang, Y. H., Trowbridge, R. & Gowans, E. J. (1997). Expression and interaction of the hepatitis C virus structural proteins and the 5' untranslated region in baculovirus infected cells. *Archives of Virology* **142**, 2211-2223.
- Wang, S. L., He, R. T. & Anderson, R. (1999). PrM- and cell-binding domains of the dengue virus E protein. *Journal of Virology* **73**, 2547-2551.
- Wardell, A. D., Errington, W., Ciaramella, G., Merson, J. & McGarvey, M. J. (1999). Characterization and mutational analysis of the helicase and NTPase activities of hepatitis C virus full-length NS3 protein. *Journal of General Virology* **80**, 701-709.
- Weiland, E., Stark, R., Haas, B., Rumenapf, T., Meyers, G. & Thiel, H. J. (1990). Pestivirus glycoprotein which induces neutralizing antibodies forms part of a disulfide-linked heterodimer. *Journal of Virology* **64**, 3563-3569.
- Weiland, E., Ahl, R., Stark, R., Weiland, F. & Thiel, H. J. (1992). A 2nd envelope glycoprotein mediates neutralization of a pestivirus, hog cholera virus. *Journal of Virology* **66**, 3677-3682.
- Weiner, A. J., Brauer, M. J., Rosenblatt, J., Richman, K. H., Tung, J., Crawford, K., Bonino, F., Saracco, G., Choo, Q. L., Houghton, M. & Han, J. H. (1991). Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins. *Virology* **180**, 842-848.
- Weiner, A. J., Thaler, M. M., Crawford, K., Ching, K., Kansopon, J., Chien, D. Y., Hall, J. E., Hu, F. & Houghton, M. (1993). A unique, predominant hepatitis C virus variant found in an infant born to a mother with multiple variants. *Journal of Virology* **67**, 4365-4368.
- Weiner, A., Erickson, A. L., Kansopon, J., Crawford, K., Muchmore, E., Hughes, A. L., Houghton, M. & Walker, C. M. (1995). Persistent hepatitis C virus infection in a

- chimpanzee is associated with emergence of a cytotoxic T-lymphocyte escape variant. *Proceedings of the National Academy of Sciences, USA* **92**, 2755-2759.
- Weisz, O. A., Swift, A. M. & Machamer, C. E. (1993). Oligomerization of a membrane protein correlates with its retention in the Golgi complex. *Journal of Cell Biology* **122**, 1185-1196.
- Wengler, G. & Wengler, G. (1989). Cell-associated West Nile flavivirus is covered with E+Pre-M protein heterodimers which are destroyed and reorganized by proteolytic cleavage during virus release. *Journal of Virology* **63**, 2521-2526.
- Westaway, E. G. (1987). Flavivirus replication strategy. *Advances in Virus Research* **33**, 45-90.
- Winkler, G., Maxwell, S.E., Ruemmler, C. & Stollar V. (1989). Newly synthesized dengue-2 virus non-structural protein NS1 is a soluble protein but becomes partially hydrophobic and membrane-associated after dimerization. *Virology* **171**, 302-305.
- Wood, J. & Patel, A. H. (1999). Manuscript submitted for publication. .
- Wright, T. L., Terrault, N.A. & Ganem, D. (1997). Hepatitis B virus. In *Clinical Virology*, pp. 633-681. Edited by D. D. Richman, R.J. Whitely and F.G. Hayden. New York: Churchill Livingstone.
- Wu, Z., Yao, N. H., Le, H. V. & Weber, P. C. (1998). Mechanism of autoproteolysis at the NS2-NS3 junction of the hepatitis C virus polyprotein. *Trends in Biochemical Sciences* **23**, 92-94.
- Yamada, N., Tanihara, K., Mizokami, M., Ohba, K., Takada, A., Tsutsumi, M. & Date, T. (1994). Full length sequence of the genome of hepatitis C virus type 3a - comparative study with different genotypes. *Journal of General Virology* **75**, 3279-3284.
- Yan, B. S., Tam, M. H. & Syu, W. J. (1998). Self-association of the C-terminal domain of the hepatitis C virus core protein. *European Journal of Biochemistry* **258**, 100-106.
- Yanagi, M., Purcell, R. H., Emerson, S. U., & Bukh, J. (1997). Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 8738-8743.
- Yang, M., Ellenberg, J., Bonifacino, J. S. & Weissman, A. M. (1997). The transmembrane domain of a carboxyl-terminal anchored protein determines localization to the endoplasmic reticulum. *Journal of Biological Chemistry* **272**, 1970-1975.

- Yano, M., Kumada, H., Kage, M., Ikeda, K., Shimamatsu, K., Inoue, O., Hashimoto, E., Lefkowitz, J. H., Ludwig, J. & Okuda, K. (1996). The long-term pathological evolution of chronic hepatitis C. *Hepatology* **23**, 1334-1340.
- Yap, S. H., Willems, M., Vandenoord, J., Habets, W., Middeldorp, J. M., Hellings, J. A., Nevens, F., Moshage, H., Desmet, V. & Fevery, J. (1994). Detection of hepatitis C virus antigen by immunohistochemical staining - a histological marker of hepatitis C virus infection. *Journal of Hepatology* **20**, 275-281.
- Yasui, K., Wakita, T., Tsukiyama-Kohara, K., Funahashi, S.-I., Ichikawa, M., Kajita, T., Moradpour, D., Wands, J. R. & Kohara, M. (1998). The native form and maturation process of hepatitis C virus core protein. *Journal of Virology* **72**, 6048-6055.
- Yi, M. K., Nakamoto, Y., Kaneko, S., Yamashita, T. & Murakami, S. (1997a). Delineation of regions important for heteromeric association of hepatitis C virus E1 and E2. *Virology* **231**, 119-129.
- Yi, M. Y., Kaneko, S., Yu, D. Y. & Murakami, S. (1997b). Hepatitis C virus envelope proteins bind lactoferrin. *Journal of Virology* **71**, 5997-6002.
- Yoshida, M., Dehara, K., Inoue, K., Okamoto, H. & Mayumi, M. (1994). Contribution of hepatitis C virus to non-A, non-B fulminant hepatitis in Japan. *Hepatology* **19**, 829-835.
- Young, R. S., Lee, C. H. B. & Im, D. S. (1998). Characterization of the structural proteins of hepatitis C virus expressed by an adenovirus recombinant. *Virus Research* **55**, 177-185.
- Yuasa, T., Ishikawa, G., Manabe, S., Sekiguchi, S., Takeuchi, K. & Miyamura, T. (1991). The particle size of hepatitis C virus estimated by filtration through microporous regenerated cellulose fiber. *Journal of General Virology* **72**, 2021-2024.
- Zhou, X., Berglund, P., Rhodes, G., Parker, S. E., Jondal, M. & Liljestrom, P. (1994). Self-replicating Semliki Forest virus RNA as recombinant vaccine. *Vaccine* **12**, 1510-1514.
- Zhu, N. L., Khoshnan, A., Schneider, R., Matsumoto, M., Dennert, G., Ware, C. & Lai, M. M. C. (1998). Hepatitis C virus core protein binds to the cytoplasmic domain of tumor necrosis factor (TNF) receptor 1 and enhances TNF-induced apoptosis. *Journal of Virology* **72**, 3691-3697.
- Zibert, A., Schreier, E. & Roggendorf, M. (1995). Antibodies in human sera specific to hypervariable region 1 of hepatitis C virus can block viral attachment. *Virology* **208**, 653-661.

Appendix 1 – Strain Glasgow Core, E1 and E2 Nucleic Acid and Amino Acid Sequence

ApaI

CCGTGCACC ATG AGC ACG AAT CCT AAA CCT CAA AGA AAA ACC AAA CGT AAC ACC AAC CGT 393
M S T N P K P Q R K T K R N T N R
+10

OGC CCA CAG GAC GTT AAG TTC COG GGT GGC GGT CAG ATC GTT GGT GGA GTT TAC TTG TTG 453
R P Q D V K F P G G G Q I V G G V Y L L
+20 +30

COG CGC AGG GGC OCT AGA TTG GGT GTG CGC GCG ACG AGG AAG ACT TOC GAG CGG TCG CAA 513
P R R G P R L G V R A T R K T S E R S Q
+40 +50

XhoI

CCT CGA GGT AGA OGT CAG OCT ATC CCC AAG GCA OGT CGG CCC AAG GGC AGG AAC TGG GCT 573
P R G R R Q P I P K A R R P K G R N W A
+60 +70

CAG CCC GGG TAT OCT TGG CCC CTC TAT GGC AAT GAG GGG TTC GGG TGG GCG GGA TGG CTC 633
Q P G Y P W P L Y G N E G F G W A G W L
+80 +90

COG TOC CCC OGT GGC TCT CGG OCT AGT TGG GGC CCC AAC GAC CCC CGA OGT AGG TOG CGC 693
P S P R G S R P S W G P N D P R R R S R
+100 +110

ClaI

AAT TTG GGT AAG GTC ATC GAT ACC CTT ACG TGC GGC TTC GTC GAT CTC ATG GGG TAC ATA 753
N L G K V I D T L T C G F V D L M G Y I
+120 +130

COG CTC GTC GGC GGC OCT CTT AGA GGC GCT GGC AGG GGC CTG GCG CAT GGC GTC OGG GTT 813
P L V G A P L R G A A R A L A H G V R V
+140 +150

CTG GAA GAC GGT GTG AAC TAT GCA ACA GGT AAC CTT OCT GGT TGC TCT TTC TCT ATC TTC 873
L E D G V N Y A T G N L P G C S F S I F
+160 +170

FspI

CTT CTG GCC CTG CTC TCT TGC CTG ACT GTG CCC GCT TCA GCC TAC CAA GTG OGC AAC TOC 933
L L A L L S C L T V P A S A Y Q V R N S
+180 +190

XhoI

ACG GGG CTT TAC CAT GTC ACC AAT GAT TGC CCT AAC TOG AGT AIT GTG TAC GAG GCG GTC 993
T G L Y H V T N D C P N S S I V Y E A V
+200 +210

XhoI

GAT GCC ATC TTG CAC ACT COG GGG TGT GTC CCT TGC GTT OGC GAG GGT AAC GCC TOG AGG 1053
D A I L H T P G C V P C V R E G N A S R
+220 +230

MscI

TGT TGG GTG GCG ATG ACC CCC ACG GTG GCC ACC AGG GAC GGC AGA CTC CCC ACA ACG CAG 1113
C W V A M T P T V A T R D G R L P T T Q
+240 +250

ClaI

CTT CGA OGT CAC ATC GAT CTG CTG GTC GGG AGT GCC ACC CTT TGT TOG GCC CTT TAC GTG 1173
L R R H I D L L V G S A T L C S A L Y V
+260 +270

GGG GAC CTG TGC GGG TCT GTC TTC CTT GTC GGC CAA CTG TTT ACT TTC TOC CCC AGA CGC 1233
 G D L C G S V F L V G Q L F T F S P R R
 +280 +290

MmI

CAC TGG ACG ACG CAA GGC TGC AAT TGT TOC ATC TAT CCC GGC CAT ATA ACG GGT CAT CGC 1293
 H W T T Q G C N C S I Y P G H I T G H R
 +300 +310

ATG GCA TGG GAT ATG ATG ATG AAC TGG TOC CCC ACG ACG GCG TTG GTG GTA GCT CAG CTG 1353
 M A W D M M M N W S P T A L V V A Q L
 +320 +330

BamHI

CTC CGG ATC CCA CAA GCC ATC TTG GAC ATG ATC GCT GGT GCT CAC TGG GGA GTC CTG GCG 1413
 L R I P O A I L D M I A G A H W G V L A
 +340 +350

GGC ATG GCG TAT TTC TOC ATG GTG GGG AAC TGG GCG AAG GTC CTG GCA GTG CTG CTG CTA 1473
 G M A Y F S M V G N W A K V L A V L L L
 +360 +370

SalI/HincII

TTT GCC GGC GTC GAC GCG GAG ACC CAC GTC ACC GGG GGA GCT GCC GCC CGC AGC ACG CTT 1533
 F A G V D A E T H V T G G A A A R S T L
 +380 +390

BclI

CAG CTT GCC GGT CTC TTT CAA CCA GGC GCC AAG CAG AAC GTC CAG TTG ATC AAC ACC AAC 1593
 Q L A G L F Q P G A K Q N V Q L I N T N
 +400 +410

GGC AGC TGG CAT GTC AAT CGC ACG GCC TTG AAC TGC AAT GAT AGC CTT AAC ACC GGC TGG 1653
 G S W H V N R T A L N C N D S L N T G W
 +420 +430

MscI

ATA GCA GGG CTT TTC TAT TAC CAC GGA TTC AAT TCT TCA GGC TGT TOC GAG AGG TTG GCC 1713
 I A G L F Y Y H G F N S S G C S E R L A
 +440 +450

AGC TGC CGA TOC CTT ACC GAC TTT GAC CAA GGT TGG GGC OCT ATC AGT TAT GCC GGC GGA 1773
 S C R S L T D F D Q G W G P I S Y A G G
 +460 +470

GGC GGC CCC GAC CAT CGC CCC TAT TGC TGG CAC TAC CCC CCA AAG CCC TGC GGT ATT GTG 1833
 G G P D H R P Y C W H Y P P K P C G I V
 +480 +490

CCC GCA AAG AGC GTG TGT GGC CCG GTA TAT TGT TTC ACT CCC AGC CCC GTG GTG GTG GGA 1893
 P A K S V C G P V Y C F T P S P V V V G
 +500 +510

AscI

ACG ACC GAC AGG TCG GGC GCG OCT ACC TAC AGC TGG GGT GCA GAT GAT ACG GAC GTC TTC 1953
 T T D R S G A P T Y S W G A D D T D V F
 +520 +530

MmI

GTC CTT AAT AAC ACC AGG CCA CCG CTG GGC AAT TGG TTC GGT TGT ACC TGG ATG AAC TCA 2013
 V L N N T R P P L G N W F G C T W M N S
 +540 +550

ACT GGA TTC ACC AAA GTG TGC GGA GCG CCC CCT TGT GTC ATC GGA GGG GTG GGC AAT AAC 2073
 T G F T K V C G A P P C V I G G V G N N
 +560 +570

ACC TTG CAC TGC CCC ACT GAT TGT TTC CGC AAG CAT CCG GAA GCC ACA TAC TCT CGG TGC 2133
 T L H C P T D C F R K H P E A T Y S R C
 +580 +590

SalI/HincII

GGC TOC GGT CCC TGG CTT ACA CCC AGG TGC CTG GTC GAC TAC CCG TAT AGG CTT TGG CAT 2193
 G S G P W L T P R C L V D Y P Y R L W H
 +600 +610

TAC CCT TGT ACC ATC AAC CAC AGC ATA TTC AAA GTC AGG ATG TAC GTG GGA GGG GTC GAG 2253
 Y P C T I N H S I F K V R M Y V G G V E
 +620 +630

MmiI

CAC AGG CTG GAC GCT GOC TGC AAT TGG ACG CGG GGC GAA CGT TGT GAT CTG GAA GAC AGG 2313
 H R L D A A C N W T R G E R C D L E D R
 +640 +650

GAC AGG TOC GAG CTT AGC CCG TTG CTG CTG TOC ACC ACA CAG TGG CAG GTC CTT CCG TGT 2373
 D R S E L S P L L L S T T Q W Q V L P C
 +660 +670

TOC TTC ACG ACC CTG CCA GOC TTG TOC ACC GGC CTC ATC CAC CTC CAC CAG AAC AAT GTG 2433
 S F T T L P A L S T G L I H L H Q N I V
 +680 +690

ScaI

GAC GTG CAG TAC TTG TAC GGG GTG GGG TCA AGC ATC GCG TOC TGG GOC ATC AAG TGG GAA 2493
 D V Q Y L Y G V G S S I A S W A I K W E
 +700 +710

TAC GTC GTT CTC CTG TTC CTT CTG CTT GCA GAC GCG CGC GTC TGC TOC TGC TTG TGG ATG 2553
 Y V V L L F L L L A D A R V C S C L W M
 +720 +730

ScaI

ATG TTA CTC ATA TOC CAA GCG GAG GCG GCT TTG GAA AAC CTC GTA GTA CTC AAT GCA GCA 2613
M L L I S Q A E R A L E N L V V L N A A
 +740 +750

TOC CTG GOC GGG ACG CAT GGT CTT GTG TOC TTC CTC GTG TTC TTC TGC TTT GCG TGG TTT 2673
 S L A S T H G L V S F L V F F C F A W F
 +760 +770

CTG AGG GGT AAG TGG GTG CCC GGA GCG GTC TAC GOC CTC TAC GGG ATG TGG OCT CTC CTC 2733
 L R S K W V P S A V Y A L Y G H W P L L
 +780 +790

CTG CTC CTG TTG GCG TTG CCC CAG CGG GCA TAC GCG CTG GAC ACG GAG GTA GCC GCG TCG 2793
 L L L L A L P Q R A Y A L D T E V A A S
 +800 +810

Eco47III

TGT GGC GGC GTT GTT CTC GTC GGG TTA ATG GOC CTA ACT CTG TCA CCA TAT TAC AAG GGC 2853
 C G G V V L V G L M A L T L S P Y Y K R
 +820 +830

TAT ATC AGC TGG TGC TTG TGG TGG CTG CAG TAT TTT CTG ACC 2856
 Y I S W C L W W L Q T F L T
 +840 +850

Key

- C** Cysteine residue
- N-X-S/T** Potential N-linked glycosylation site
- L P G** Predicted single sequence to direct protein into the ER
- Q A I** Predicted hydrophobic/transmembrane domain

M S T	Core or NS2 protein amino acid sequence
Y Q V	E1 glycoprotein amino acid sequence
E T H	E2 glycoprotein amino acid sequence
E I H	E2 glycoprotein hypervariable regions
A A L	p7 amino acid sequence
<i>GTGCAC</i>	Endonuclease restriction site

Appendix 2 - Amino Acid Sequence Alignment of Strain Glasgow and H77 Core, E1 and E2

	1				50
Cele2aline.Con	MSTNPKPQRK	TKRNTNRRPQ	DVKFPGGGQI	VGGVYLLPRR	GPRLGVRATR
Glacelle2.Frg	MSTNPKPQRK	TKRNTNRRPQ	DVKFPGGGQI	VGGVYLLPRR	GPRLGVRATR
H77cele2.Frg	MSTNPKPQRK	TKRNTNRRPQ	DVKFPGGGQI	VGGVYLLPRR	GPRLGVRATR
	51				100
Cele2aline.Con	KTSEERSQPRG	RRQPIPKARR	P GR WAQPG	YPWPLYGNEG	.GWAGWL.SP
Glacelle2.Frg	KTSEERSQPRG	RRQPIPKARR	PKGRNWAQPG	YPWPLYGNEG	FGWAGWLPSP
H77cele2.Frg	KTSEERSQPRG	RRQPIPKARR	PEGRTWAQPG	YPWPLYGNEG	CGWAGWLLSP
	101				150
Cele2aline.Con	RGSRPSWGP.	DPRRRSRNLG	KVIDTLTCGF	.DLMGYIPLV	GAPL.GAARA
Glacelle2.Frg	RGSRPSWGPN	DPRRRSRNLG	KVIDTLTCGF	VDLMGYIPLV	GAPLRGAARA
H77cele2.Frg	RGSRPSWGP	DPRRRSRNLG	KVIDTLTCGF	ADLMGYIPLV	GAPLGGARA
	151				200
Cele2aline.Con	LAHGVRVLED	GVNYATGNLP	GCSFSIFLLA	LLSCLTVPAS	AYQVRNS.GL
Glacelle2.Frg	LAHGVRVLED	GVNYATGNLP	GCSFSIFLLA	LLSCLTVPAS	AYQVRNSTGL
H77cele2.Frg	LAHGVRVLED	GVNYATGNLP	GCSFSIFLLA	LLSCLTVPAS	AYQVRNSSGL
	201				250
Cele2aline.Con	YHVTNDPCNS	SIVYEA.DAI	LHTPGCVPCV	REGNASRCWV	A.TPTVATRD
Glacelle2.Frg	YHVTNDPCNS	SIVYEAVD	LHTPGCVPCV	REGNASRCWV	AMTPTVATRD
H77cele2.Frg	YHVTNDPCNS	SIVYEAADAI	LHTPGCVPCV	REGNASRCWV	AVTPTVATRD
	251				300
Cele2aline.Con	G.LPTTQLRR	HIDLLVGSAT	LCSALYVGDL	CGSVFLVGQL	FTFSPRRHWT
Glacelle2.Frg	GRLPTTQLRR	HIDLLVGSAT	LCSALYVGDL	CGSVFLVGQL	FTFSPRRHWT
H77cele2.Frg	GKLPTTQLRR	HIDLLVGSAT	LCSALYVGDL	CGSVFLVGQL	FTFSPRRHWT
	301				350
Cele2aline.Con	TQ.CNCISIYP	GHITGHRMAW	DMMNWSPT.	ALVVAQLLRI	PQAI.DMIAG
Glacelle2.Frg	TQGCNCISIYP	GHITGHRMAW	DMMNWSPTT	ALVVAQLLRI	PQAILDMIAG
H77cele2.Frg	TQDCNCISIYP	GHITGHRMAW	DMMNWSPTA	ALVVAQLLRI	PQAIMDMIAG
	351				400
Cele2aline.Con	AHWGVLG.A	YFSMVGWAK	VL.VLLLFAG	VDAETHVTGG	.A.R.T..L.
Glacelle2.Frg	AHWGVLGMA	YFSMVGWAK	VLAVLLLFAG	VDAETHVTGG	AAARSTLQLA
H77cele2.Frg	AHWGVLGIA	YFSMVGWAK	VLVLLLFAG	VDAETHVTGG	NAGRRTAGLV
	401				450
Cele2aline.Con	GL..PGAKQN	.QLINTNGSW	H.N.TALNCN	.SLNTGW.AG	LFY.H.FNSS
Glacelle2.Frg	GLFQPGAKQN	VQLINTNGSW	HVNRTALNCN	DSLNTGWIAG	LFYYHGFNSS
H77cele2.Frg	GLLTPGAKQN	IQLINTNGSW	HINSTALNCN	ESLNTGWLAG	LFYQHGFNSS
	451				500
Cele2aline.Con	GC.ERLASC	.LTDF.QGWG	PISYA.G.G.	D.RPYCWHYP	P.PCGIVPAK
Glacelle2.Frg	GCSERLASC	SLTDFDQGWG	PISYAGGGGP	DHRPYCWHYP	PKPCGIVPAK
H77cele2.Frg	GCPERLASC	RLTDFAQGWG	PISYANGSGL	DERPYCWHYP	PRPCGIVPAK
	501				550
Cele2aline.Con	SVCGPVYCF	PSPVVVGTTD	RSGAPTYSWG	A.DTDVFLVN	NTRPPLGNWF
Glacelle2.Frg	SVCGPVYCF	PSPVVVGTTD	RSGAPTYSWG	ADTDVFLVN	NTRPPLGNWF
H77cele2.Frg	SVCGPVYCF	PSPVVVGTTD	RSGAPTYSWG	ANDTDVFLVN	NTRPPLGNWF

	551				600
Cele2aline.Con	GCTWMN ST GF	TKVCGAPPCV	IGGVGN NTL .	CPTDCFRKHP	EATYSRCGSG
Glacelle2.Frg	GCTWMN ST GF	TKVCGAPPCV	IGGVGN NTLH	CPTDCFRKHP	EATYSRCGSG
H77cele2.Frg	GCTWMN ST GF	TKVCGAPPCV	IGGVGN NTLL	CPTDCFRKHP	EATYSRCGSG
	601				650
Cele2aline.Con	PW.TPRC.VD	YPYRLWHYPC	TIN..IFKVR	MYVGGVEHRL	.AACN W TRGE
Glacelle2.Frg	PWLTPRC L VD	YPYRLWHYPC	TIN HS IFKVR	MYVGGVEHRL	DAAC NW TRGE
H77cele2.Frg	PWITPRC M VD	YPYRLWHYPC	TIN Y TIFKVR	MYVGGVEHRL	EAAC NW TRGE
	651				700
Cele2aline.Con	RCDLEDRDRS	ELSP LL LSTT	QWQVLP S FT	TLPALSTGLI	HLHQ N IVDVQ
Glacelle2.Frg	RCDLEDRDRS	ELSP LL LSTT	QWQVLP S FT	TLPALSTGLI	HLHQ N IVDVQ
H77cele2.Frg	RCDLEDRDRS	ELSP LL LSTT	QWQVLP S FT	TLPALSTGLI	HLHQ N IVDVQ
	701				750
Cele2aline.Con	YLYGVGSSIA	SWAIKWEYV V	LLFLL L LADAR	VC S CLW M MLL	ISQAE AA LEN
Glacelle2.Frg	YLYGVGSSIA	SWAIKWEYV V	LLFLL L LADAR	VC S CLW M MLL	ISQAE AA LEN
H77cele2.Frg	YLYGVGSSIA	SWAIKWEYV V	LLFLL L LADAR	VC S CLW M MLL	ISQAE AA LEN
	751				800
Cele2aline.Con	LV.LRAASLA	STRGLV S FLV	FE CF AW L LG	K WV PC AV T AL	Y CH W P LL L LL
Glacelle2.Frg	LVVLRRAASLA	STRGLV S FLV	FE CF AW L LG	K WV PC AV T AL	Y CH W P LL L LL
H77cele2.Frg	LVILRAASLA	STRGLV S FLV	FE CF AW L LG	K WV PC AV T AL	Y CH W P LL L LL
	801			838	
Cele2aline.Con	LAL PQ RAYAL	DTEVAASCGG	VVLVGLMALT	LSPY Y KRY	
Glacelle2.Frg	LAL PQ RAYAL	DTEVAASCGG	VVLVGLMALT	LSPY Y KRY	
H77cele2.Frg	LAL PQ RAYAL	DTEVAASCGG	VVLVGLMALT	LSPY Y KRY	

Key

C	Cysteine residue
<u>N-X-S/T</u>	Potential N-linked glycosylation site
M S T	Core or NS2 protein amino acid sequence
Y Q V	E1 glycoprotein amino acid sequence
E T H	E2 glycoprotein amino acid sequence
E T H	p7 amino acid sequence
..	Amino acid sequence changes

Appendix 3 – Strain H77 Core, E1, and E2 Nucleic Acid and Amino Acid Sequence

ApaLI

CCG**TGCACC** ATG AGC ACG AAT CCT AAA CCT CAA AGA AAA ACC AAA CGT AAC ACC AAC CGT 392
 M S T N P K P Q R K T K R N T N R
 +10

CGC CCA CAG GAC GTC AAG TTC CCG GGT GGC GGT CAG ATC GTT GGT GGA GTT TAC TTG TTG 452
 R P Q D V K F P G G G Q I V G G V Y L L
 +20 +30

CCG CGC AGG GGC OCT AGA TTG GGT GTG CCG GCG ACG AGG AAG ACT TOC GAG CGG TCG CAA 512
 P R R G P R L G V R A T R K T S E R S Q
 +40 +50

XhoI

CCT CCA GGT AGA CGT CAG OCT ATC CCC AAG GCA CGT CCG CCC GAG GGC AGG ACC TGG GCT 572
 P R G R R Q P I P K A R R P E G R T W A
 +60 +70

CAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAT GAG GGT TGC GGG TGG GCG GGA TGG CTC 632
 Q P G Y P W P L Y G N E G C G W A G W L
 +80 +90

CTG TCT CCC CGT GGC TCT CCG OCT AGC TGG GGC CCC ACA GAC CCC CCG CGT AGG TCG CGC 692
 L S P R G S R P S W G P T D P R R R S R
 +100 +110

CluI

AAT TTG GGT AAG GTC **ATC GAT** ACC CTT ACG TGC GGC TTC GCC GAC CTC ATG GGG TAC ATA 752
 N L G K V I D T L T C G F A D L M G Y I
 +120 +130

CCG CTC GTC GGC GGC OCT CTT GGA GGC GCT GCC AGG GCC CTG GCG CAT GGC GTC CCG GTT 812
 P L V G A P L G G A A R A L A H G V R V
 +140 +150

CTG GAA GAC GGC GTG AAC TAT GCA ACA GGG AAC CTT OCT GGT TGC TCT TTC TCT ATC TTC 872
 L E D G V N Y A T G N L P G C S F S I F
 +160 +170

FspI

CTT CTG GCC CTG CTC TCT TGC CTG ACT GTG CCC GCT TCA GCC TAC CAA **GTG CCG AAT TOC** 932
L L A L L S C L T V P A S A Y Q V R N S
 +180 +190

XhoI

EagI

TCG GGG CTT TAC CAT GTC ACC AAT GAT TGC CCT AAC **TCG AGT** ATT GTG TAC GAG **GCG GCC** 992
S G L Y H V T N D C P N S S I V Y E A A
 +200 +210

XhoI

GAT GCC ATC CTG CAC ACT CCG GGG TGT GTC CCT TGC GTT CCG GAG GGT AAC **GCC TCG AGG** 1052
 D A I L H T P G C V P C V R E G N A S R
 +220 +230

BstEII

TGT TGG GTG **GCG GTG ACC** CCC ACG GTG GCC ACC AGG GAC GGC AAA CTC CCC ACA ACG CAG 1112
 C W V A V T P T V A T R D G K L P T T Q
 +240 +250

ClaI

CTT CGA CGT CAT ATC GAT CTG CTT GTC GGG AGC GCC ACC CTC TGC TCG GCC CTC TAC GTG 1172
 L R R H I D L L V G S A T L C S A L Y V
 +260 +270

GGG GAC CTG TGC GGG TCT GTC TTT CTT GTT GGT CAA CTG TTT ACC TTC TCT CCC AGG CGC 1232
 G D L C G S V F L V G Q L F T F S P R R
 +280 +290

MmiI

CAC TGG ACG ACG CAA GAC TGC AAT TGT TCT ATC TAT CCC GGC CAT ATA ACG GGT CAT CGC 1292
 H W T T Q D C N C S I Y P G H I T G H R
 +300 +310

ATG GCA TGG GAT ATG ATG ATG AAC TGG TOC CCT ACG GCA GCG TTG GTG GTA GCT CAG CTG 1352
 M A W D M M M N W S P T A A L V V A Q L
 +320 +330

BamHI

CTC CGG ATC CCA CAA GCC ATC ATG GAC ATG ATC GCT GGT GCT CAC TGG GGA GTC CTG GCG 1412
 L R I P Q A I M D M I A G A H W G V L A
 +340 +350

GGC ATA GCG TAT TTC TOC ATG GTG GGG AAC TGG GCG AAG GTC CTG GTA GTG CTG CTG CTA 1472
 G I A Y F S M V G N W A K V L V V L L L
 +360 +370

SalI

TTT GCC GGC GTC GAC GCG GAA ACC CAC GTC ACC GGG GGA AAT GCC GGC CGC ACC ACG GCT 1532
 F A G V D A E T H V T G G N A G R T T A
 +380 +390

GGG CTT GTT GGT CTC CTT ACA CCA GGC GCC AAG CAG AAC ATC CAA CTG ATC AAC ACC AAC 1592
 G L V G L L T P G A K Q N I Q L I N T N
 +400 +410

GGC AGT TGG CAC ATC AAT AGC ACG GCC TTG AAT TGC AAT GAA AGC CTT AAC ACC GGC TGG 1652
 G S W H I N S T A L N C N E S L N T G W
 +420 +430

TTA GCA GGG CTC TTC TAT CAA CAC AAA TTC AAC TCT TCA GGC TGT OCT GAG AGG TTG GCC 1712
 L A G L F Y Q H K F N S S G C P E R L A
 +440 +450

AGC TGC CGA CGC CTT ACC GAT TTT GCC CAG GGC TGG GGT OCT ATC AGT TAT GCC AAC GGA 1772
 S C R R L T D F A Q G W G P I S Y A N G
 +460 +470

AGC GGC CTC GAC GAA CGC CCC TAC TGC TGG CAC TAC OCT CCA AGA OCT TGT GGC ATT GTG 1832
 S G L D E R P Y C W H Y P P R P C G I V
 +480 +490

CCC GCA AAG AGC GTG TGT GGC CCG GTA TAT TGC TTC ACT CCC AGC CCC GTG GTG GTG GGA 1892
 P A K S V C G P V Y C F T P S P V V V G
 +500 +510

AscI

ACG ACC GAC AGG TOG GGC GCG OCT ACC TAC AGC TGG GGT GCA AAT GAT ACG GAT GTC TTC 1952
 T T D R S G A P T Y S W G A N D T D V F
 +520 +530

MmiI

GTC CTT AAC AAC ACC AGG CCA CCG CTG GGC AAT TGG TTC GGT TGT ACC TGG ATG AAC TCA 2012
 V L N N T R P P L G N W F G C T W M N S
 +540 +550

ACT GGA TTC ACC AAA GTG TGC GGA GCG CCC OCT TGT GTC ATC GGA GGG GTG GGC AAC AAC 2072
 T G F T K V C G A P P C V I G G V G N N
 +560 +570

ACC TTG CTC TGC OCC ACT GAT TGC TTC CGC AAA CAT COG GAA GCC ACA TAC TCT OGG TGC 2132
T L L C P T D C F R K H P E A T Y S R C
+580 +590

SalI

GGC TOC GGT OCC TGG ATT ACA OCC AGG TGC ATG *GTC GAC* TAC COG TAT AGG CTT TGG CAC 2192
G S G P W I T P R C M V D Y P Y R L W H
+600 +610

TAT OCT TGT ACC ATC AAT TAC ACC ATA TTC AAA GTC AGG ATG TAC GTG GGA GGG GTC GAG 2252
Y P C T I N Y T I F K V R M Y V G G V E
+620 +630

CAC AGG CTG GAA GCG GCC TGC AAC TGG ACG CGG GGC GAA CGC TGT GAT CTG GAA GAC AGG 2312
H R L E A A C N W T R G E R C D L E D R
+640 +650

GAC AGG TOC GAG CTC AGC COG TTG CTG CTG TOC ACC ACA CAG TGG CAG GTC CTT COG TGT 2372
D R S E L S P L L L S T T Q W Q V L P C
+660 +670

TCT TTC ACG ACC CTG CCA GCC TTG TOC ACC GGC CTC ATC CAC CTC CAC CAG AAC ATT GTG 2432
S F T T L P A L S T G L I H L H Q N I V
+680 +690

ScaI

GAC GTG *CAG TAC* TTG TAC GGG GTA GGG TCA AGC ATC GCG TOC TGG GCC ATT AAG TGG GAG 2492
D V Q Y L Y G V G S S I A S W A I K W E
+700 +710

TAC GTC GTT CTC CTG TTC CTT CTG CTT GCA GAC GCG CGC GTC TGC TOC TGC TTG TGG ATG 2552
Y V V L L F L L L A D A R V C S C L W M
+720 +730

ATG TTA CTC ATA TOC CAA GCG GAG GCG GCT TTG GAG AAC CTC GTA ATA CTC AAT GCA GCA 2612
M L L I S Q A E A A L E H L V I L N A A
+740 +750

TOC CTG GCC GGG ACG CAC GGT CTT GTG TOC TTC CTC GTG TTC TTC TGC TTT GCG TGG TAT 2672
S L A S T H S L V S F L V F F C E A W Y
+760 +770

CTG AAG GGT AGG TGG GTG OCC GGA GCG GTC TAC GGC CTC TAC GGG ATG TGG OCT CTC CTC 2732
L K S R W V P S A V Y A L Y S H W P L L
+780 +790

CTG CTC CTG CTG GCG TTG OCT CAG CGG GCA TAC GCA CTG GAC ACG GAG GTG GCC GCG TCG 2792
L L L L A L P Q R A Y A L D T E V A A S
+800 +810

Eco47III

TGT GGC GGC GTT GTT CTT GTC GGG TTA ATG GCG CTG ACT CTG TCG OCA TAT TAC *AAG CGC* 2852
C G G V V L V G L M A L T L S P Y Y K R
+820 +830

TAT 2855
Y

Key

C Cysteine residue

N-X-S/T Potential N-linked glycosylation site

L P G Predicted single sequence for insertion into the ER

<u>Q A I</u>	Predicted transmembrane domain
M S T	Core or NS2 protein amino acid sequence
Y Q V	E1 glycoprotein amino acid sequence
E T H	E2 glycoprotein amino acid sequence
E I H	E2 glycoprotein hypervariable regions
Λ Λ L	p7 amino acid sequence
<i>GTGCAC</i>	Endonuclease restriction site

Appendix 4 - HSV-1 Glycoprotein D Nucleic Acid and Amino Acid Sequence

CGTTCCGGT ATG GGG GGG GCT GCC GCC AGG TTG GGG GCC GTG ATT TTG TTT GTC GTC ATA 138469
M G G A A A R L G A V I L F V V I
-20 -10

GTG GGC CTC CAT GGG GTC CGC AGC AAA TAT GCC TTG GTG GAT GCC TCT CTC AAG ATG GCC 138529
V G L H G V R S K Y A L V D A S L K M A
-1 +1 +10 *EaeI*

GAC OCC AAT CGC TTT CGC GGC AAA GAC CTT COG GTC CTG GAC *CAG CTG* ACC GAC CCT COG 138589
D P N R F R G K D L P V L D Q L T D P P
+20 +30 *PvuII*

GGG GTC OGG CGC GTG TAC CAC ATC CAG GCG GGC CTA COG GAC COG TTC CAG CCC OCC AGC 138649
G V R R V Y H I Q A G L P D P F Q P P S
+40 +50

CTC COG ATC ACG GTT TAC TAC GCC GTG TTG GAG CGC GCC TGC CGC AGC GTG CTC CTA AAC 138709
L P I T V Y Y A V L E R A C R S V L L N
+60 +70

GCA COG TCG GAG GCC COC CAG ATT GTC CGC GGG GOC TOC GAA GAC GTC COG AAA CAA COC 138769
A P S E A P Q I V R G A S E D V R K Q P
+80 +90

TAC AAC CTG ACC ATC GCT TGG TTT COG ATG GGA GGC AAC TGT GCT ATC COC ATC ACG GTC 138829
Y N L T I A W F R M G G N C A I P I T V
+100 +110

ATG GAG TAC ACC GAA TGC TOC TAC AAC AAG TCT CTG GGG GOC TGT COC ATC CGA ACG CAG 138889
M E Y T E C S Y N K S L G A C P I R T Q
+120 +130

COC CGC TGG AAC TAC TAT GAC AGC TTC AGC GOC GTC AGC GAG GAT AAC CTG GGG TTC CTG 138949
P R W N Y Y D S F S A V S E D N L G F L
+140 +150

ATG CAC GOC COC GCG TTT GAG ACC GOC GGC ACG TAC CTG COG CTC GTG AAG ATA AAC GAC 139009
M H A P A F E T A G T Y L R L V K I N D
+160 +170

TGG ACG GAG ATT ACA CAG TTT ATC CTG GAG CAC CGA GOC AAG GGC TOC TGT AAG TAC GOC 139069
W T E I T Q F I L E H R A K G S C K Y A
+180 +190

CTC COG CTG CGC ATC COC COG TCA GOC TGC CTC TOC COC CAG GOC TAC CAG CAG GGG GTG 139129
L P L R I P P S A C L S P Q A Y Q Q G V
+200 +210

ACG GTG GAC AGC ATC GGG ATG CTG COC CGC TTC ATC COC GAG AAC CAG CGC ACC GTC GOC 139189
T V D S I G M L P R F I P E N Q R T V A
+220 +230

GTA TAC AGC TTG AAG ATC GOC GGG TGG CAC GGG COC AAG GOC CCA TAC ACG AGC ACC CTG 139249
V Y S L K I A G W H G P K A P Y T S T L
+240 +250

CTG COC COG GAG CTG TOC GAG ACC COC AAC GOC ACG CAG CCA GAA CTC GOC COG GAA GAC 139309
L P P E L S E T P N A T Q P E L A P E D
+260 +270

COC GAG GAT TOG GOC CTC TTG GAG GAC COC GTG GGG ACG GTG GCG COG CAA ATC CCA CCA 139369
P E D S A L L E D P V G T V A P Q I P P
+280 +290

EaeI

AAC TGG CAC ATA CCG TGG ATC CAG GAC GGC GGG ACG OCT TAC CAT CCC COG GGC ACC CCG 139429
 N W H I P S I Q D A A T P Y H P P A T P
 +300 +310

AAC AAC ATG GGC CTG ATC GGC GGC GGG GTG GGC GGC AGT CTC CTG GCA GGC CTG GTC ATT 139489
 N N M G L I A G A V G G S L L A A L V I
 +320 +330

TGC GGA ATT GTG TAC TGG ATG GGC CGC CAC ACT CAA AAA GGC CCA AAG CGC ATA CGC CTC 139549
C G I V Y W M R R H T Q K A P K R I R L
 +340 +350

CCC CAC ATC CCG GAA GAC GAC CAG CCG TOC TGG CAC CAG CCC TTG TTT TAC TAG ATACCC 139609
 P H I R E D D Q P S S H Q P L F Y *
 +360 +370

Key

C Cysteine residue

N-X-S/T Potential N-linked glycosylation site

L P G Predicted single sequence for insertion into the ER

Q A I Predicted transmembrane domain

M S T Glycoprotein D amino acid sequence

A A L Cytoplasmic domain amino acid sequence

GTGCAC Endonuclease restriction site

**Appendix 5 – Characteristics of E1 and E2 Specific Antibodies
Generated by Dr A Patel**

Antibody	Type	Epitope	Specificity	Reactivity to Strain	
				H77	Glasgow
R141	rabbit polyclonal	Linear	E2 (aa 384-410)	-	+++
R528	rabbit polyclonal	Linear	E1	+++	+++
ALP98	mouse monoclonal	Linear	E2 (aa 610-699)	+++	+++
AP21.010	mouse monoclonal	Linear	E1 (aa 191-236)	+++	+++
AP497	mouse monoclonal	Conformational	E1 (aa 191-236)	+++	+++

Appendix 6 – Restriction Enzymes Employed to Generate DNA Clones

Restriction enzyme	Recognition sequence (/ indicates site of cleavage)
<i>ApaLI</i>	G/TGCAC
<i>AscI</i>	GG/CGCGCC
<i>BamHI</i>	G/GATCC
<i>BclI</i>	T/GATCA
<i>BglII</i>	A/GATCT
<i>BstEII</i>	G/GTNACC
<i>ClaI</i>	AT/CGAT
<i>EaeI</i>	Y/GGCCR
<i>EagI</i>	C/GGCCG
<i>Eco47III</i>	AGC/GCT
<i>EcoRI</i>	G/AATTC
<i>FspI</i>	TGC/GCA
<i>HincII</i>	GTY/RAC
<i>HindIII</i>	A/AGCTT
<i>KpnI</i>	GGTAC/C
<i>MscI</i>	TGG/CCA
<i>MunI</i>	C/AATTG
<i>NsiI</i>	ATGCA/T
<i>PvuII</i>	CAG/CTG
<i>SalI</i>	G/TCGAC
<i>ScaI</i>	AGT/ACT
<i>XbaI</i>	T/CTAGA
<i>XhoI</i>	C/TCGAG

Covalent interactions are not required to permit or stabilize the non-covalent association of hepatitis C virus glycoproteins E1 and E2

Shishu Patil, Arvind H. Patel and John McLauchlan

Department of Virology Unit, Division of Virology, University of Glasgow, Church Street, Glasgow G11 5JR, UK

Hepatitis C virus (HCV) encodes two glycoproteins, E1 and E2, which are thought to locate to the envelope of virus particles. These proteins form two complexes in tissue culture systems, a high molecular mass aggregate that contains intermolecular covalent bonds and a native complex in which E1 and E2 associate by non-covalent interactions. The contribution of either complex to the structures of the proteins on virus particles is not known. Using dithiothreitol to reduce inter- and intramolecular disulphide bonds *in situ*, we have studied the nature of the interactions within the aggregate and the role of covalent bonds in the early stages of E1–E2 association. Results with two HCV type 1a strains, Glasgow and H77, showed that the aggregate contains not only covalent interactions but also non-covalent associations between E1 and E2. These non-covalent associations are complex since deletion mutant analysis failed to identify any single region which was required for non-covalent interaction. Complex formation by *de novo* synthesized proteins was not arrested under reducing conditions which prevented the production of inter- and intramolecular disulphide bonds. Moreover, a conformation-specific antibody continued to recognize the E2 protein in reduced complexes, indicating that covalent bonds do not stabilize certain structures of E2 that can interact with E1. These data suggest that disulphide bonds are not required either to allow association between the proteins or to stabilize E1–E2 complexes.

Introduction

Chronic infection with hepatitis C virus (HCV) is the major cause of non-A, non-B viral hepatitis. The virus is a member of the *Flaviviridae* and has a positive-sense, single-stranded genome of about 9.6 kb which encodes a polyprotein of between 3008 and 3037 amino acids (reviewed in Clarke, 1997). Three proteins generated by host cell signalase cleavage in the N-terminal portion of the polyprotein, termed core, E1 and E2, are the proposed proteinaceous components of the virus envelope. Numerous studies have reported that E1 and E2 are glycosylated and associate to form a heteromeric complex (Grakoui *et al.*, 1993; Ralston *et al.*, 1993; Matsuura *et al.*, 1994; Selby *et al.*, 1994). The nature of the authentic association between E1 and E2 is the subject of debate since both covalent and non-covalent interactions have

been identified (Grakoui *et al.*, 1993; Ralston *et al.*, 1993). Although direct analysis of HCV virions has not been possible, it is considered that the non-covalent interactions are authentic while covalent associations may represent aggregation of misfolded proteins (Dubuisson *et al.*, 1994; Deleersnyder *et al.*, 1997). From studies on the formation of the E1–E2 complex, E2 acquires disulphide bonds rapidly but this process is slower with E1 (Dubuisson & Rice, 1996). Since the two proteins appear to associate soon after synthesis, disulphide bond formation within E2 may be important for interaction with E1 and, moreover, may induce conformational changes in E1, including disulphide bond formation.

A convenient method to determine the role of disulphide bonds in proteins which cycle through the endoplasmic reticulum (ER) is addition of the reducing agent dithiothreitol (DTT) to cells during protein synthesis (Braakman *et al.*, 1992). In this way, proteins can be reversibly reduced and reoxidized within a cellular environment. DTT does not seriously impair translation, translocation, *N*-glycosylation or signal sequence removal and the secretory pathway from the intermediate

Author for correspondence: John McLauchlan

© +44 141 337 2236. e-mail j.mclauchlan@vir.gla.ac.uk

compartment and Golgi complex also continues to function although export from the ER may be reduced (Verde *et al.*, 1995). This methodology also has been used to identify DTT-resistant forms of the E1 glycoprotein in Sindbis virus which arise upon heteromeric complex formation with the precursor E2 protein and are considered to represent species that have achieved a particular stage of the maturation process (Carleton & Brown, 1996, 1997).

Using DTT to create a reducing environment within cells, we have examined the role of disulphide bonds in E1–E2 complexes formed by two type 1a strains of HCV, a local isolate called strain Glasgow (M. McElwee & R. M. Elliott, unpublished data) and H77 (Ogata *et al.*, 1991). This has allowed us to determine whether disulphide bonds are necessary for stabilizing interactions between the proteins and to study their involvement in the initial stages of complex formation.

Methods

Construction of plasmids. Details of the regions used in this study to express core, E1 and E2 from both H77 and Glasgow strains are given in Fig. 1. Cloned cDNA fragments from HCV strain Glasgow, which has a 1a genotype, were kindly provided by M. McElwee and R. M. Elliott. A plasmid, pCV-H77C (kindly supplied by J. Bukh), carrying the infectious full-length cDNA sequence of HCV strain H77 (Yanagi *et al.*, 1997), provided the fragment for the H77 construct. All fragments were initially cloned into vector pGEM-1 (Promega) and then flanked by *Bgl*III sites. For expression purposes, *Bgl*III fragments containing HCV sequences from these plasmids were introduced into the unique *Bam*HI site of the Semliki Forest virus expression vector, pSFV1 (Liljestrom & Garoff, 1991). E2 deletion mutants in Fig. 1 were derived from pSFV/E1_{his}E2 and therefore contain the histidine tag in E1. Constructs were made by standard methods using restriction enzyme fragments along with synthetic oligonucleotides to combine fragments with incompatible termini. PCR was used only to construct the 3'-terminal HCV sequences in pGEM/Δ609–698. The nucleotide sequences of all regions that were modified as a consequence of cloning procedures were determined before insertion into pSFV1.

Maintenance of tissue culture cells. BHK C13 cells were grown and maintained in Glasgow minimal Eagle's Medium supplemented with 10% newborn calf serum (NCS), 4% tryptose phosphate broth and 100 IU/ml penicillin/streptomycin (ETC10).

In vitro transcription and electroporation of SFV RNA into cells. Recombinant SFV plasmids were linearized by *Spe*I digestion and then used as templates for *in vitro* transcription essentially as described by Liljestrom & Garoff (1991). Typically, 25 μl transcription reactions contained 40 mM Tris–HCl, pH 7.9, 6 mM MgCl₂, 2 mM spermidine, 10 mM NaCl, 5 mM DTT, 1 mM each of ATP, CTP and UTP, 500 μM GTP, 1 mM cap analogue m₇G(5')ppp(5')G (Boehringer Mannheim), 40 units RNasin and 50 units SP6 RNA polymerase (apart from cap analogue, all reagents were supplied by Pharmacia). Reactions were incubated at 37 °C for 2 h.

For electroporation, cells were grown to confluence, trypsinized and resuspended in ETC10 medium. Cells were pelleted by centrifugation at 400 g at room temperature for 5 min, resuspended in PBSA and repelleted as before. The cell pellet was finally resuspended in PBSA to give

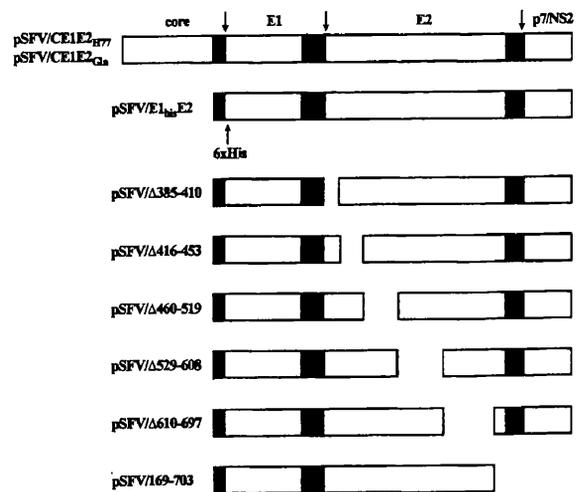


Fig. 1. Details of the regions of HCV expressed using the SFV system. The positions of cleavage sites between core/E1, E1/E2 and E2/p7/NS2 are arrowed. With the exception of pSFV/169–703, all constructs terminated translation at amino acid residue 837 of the HCV polyprotein. In pSFV/CE1E2_{H77} and pSFV/CE1E2_{Gla}, translation was initiated at residue 1 of the HCV polyprotein, whereas in the remaining constructs translation began at HCV residue 169. The histidine tag, which introduced six histidine residues into the E1 sequences in pSFV/E1_{his}E2, was present in the E2 deletion mutants also. For the deletion mutants, constructs are named according to the residues which were removed. Filled boxes denote the hydrophobic domains which contain the signal sequences that direct cleavage within the polyprotein.

10⁷ cells/ml. Cell suspension (800 μl) was mixed by inversion with 25 μl of *in vitro* transcription reactions and the mixture electroporated (Bio-Rad Gene Pulser II) with two consecutive pulses at 1.2 kV and 25 μF at room temperature. Electroporated cells were diluted 1/20 in ETC10 and seeded onto 35 mm tissue culture dishes.

Metabolic labelling of proteins. For extended labelling times, cells were washed and incubated in BHK medium with reduced concentrations of NCS (2%) and methionine (1/5 normal concentration) at 3.5 h post-electroporation. After incubation at 37 °C for 30 min, proteins were radiolabelled by addition of 10 μCi/ml [³⁵S]methionine (1175 Ci/mmol; Amersham) and incubation at 37 °C was continued for a further 8 h before harvesting.

In pulse–chase experiments, cells were incubated at 37 °C for 11.5 h following electroporation and then incubated in media with reduced concentrations of NCS (1%), methionine and cysteine (1/5 normal concentration) for 30 min prior to labelling. Cells were radiolabelled with 150 μCi Promix (> 1000 Ci/mmol; Amersham) in 500 μl PBS containing 1% NCS for the appropriate length of time and chased in ETC10 supplemented with 0.5 mg/ml methionine and cysteine.

DTT treatment of cells. In experiments analysing the effects of reducing agent on preformed complexes, cells which had been radiolabelled were firstly incubated in 200 μg/ml cycloheximide to block translation followed by addition of DTT at concentrations and for times indicated in the text. In pulse–chase experiments, 5 mM DTT was added to cells 15 min prior to and maintained at that concentration during the pulse–chase period.

For harvesting, cells were washed twice in cold PBS containing 20 mM *N*-ethylmaleimide (NEM) and harvested in either 500 μl lysis buffer (25 mM Tris–HCl, pH 8.0, 40 mM imidazole, 300 mM NaCl, 1% Triton X-100, 20 mM NEM, 1 mM PMSF) for purification by Ni–NTA

agarose or 500 μ l immunoprecipitation buffer (25 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1% Triton X-100, 20 mM NEM, 1 mM EDTA, 1 mM PMSF). Cells were lysed for 10 min on ice before centrifugation at 13 000 g to remove insoluble cell debris.

■ **Endo H digestion of glycoproteins.** Lysates from cells were denatured in 0.5% SDS, 1% β -mercaptoethanol at 100 °C for 10 min. Sodium citrate was added to 50 mM followed by 2000 units of endo H₁ (New England Biolabs) and the reaction was incubated at 37 °C for 1 h.

■ **Ni-NTA agarose purification and immunoprecipitation of E1-E2 complexes.** For immunoprecipitation, cell lysates were incubated overnight in 1–5 μ l of the appropriate antibody (diluted 1/500) at 4 °C, followed by addition of 50 μ l Protein A-Sepharose (Sigma) and incubation at 4 °C for a further 1 h. Three anti-E2 immunological reagents were used: two monoclonal antibodies, H53 (Cocquerel *et al.*, 1998; kindly provided by J. Dubuisson) and ALP98 (raised against strain Glasgow E2), and a rabbit antiserum, R141 (raised against the hypervariable region of strain Glasgow E2). Immune complexes attached to Sepharose were pelleted at 10 000 g for 30 s and washed three times with immunoprecipitation buffer. Proteins bound to Sepharose were removed by addition of 200 mM Tris-HCl, pH 6.8, 0.5% SDS, 10% glycerol and 0.01% bromophenol blue.

Purification on Ni-NTA agarose was accomplished by adding 50 μ l of equilibrated Ni-NTA resin to cell lysates and rotating the mixture at 4 °C for 1 h. Resin was pelleted at 10 000 g for 30 s and washed three times with lysis buffer containing 50 mM imidazole. Proteins were eluted using the buffer described above.

■ **PAGE and Western blot analysis.** Samples were prepared for electrophoresis by heating to 95 °C. For electrophoresis under reducing conditions, DTT was added to samples to a final concentration of 20 mM before heating. Samples were cooled and electrophoresed on 10 or 12% polyacrylamide gels cross-linked with 2.5% (w/w) *N,N'*-methylene bisacrylamide (Laemmli, 1970). Polypeptides were detected by autoradiography using XS-1 X-ray film (Kodak).

For Western blot analysis, proteins were separated on polyacrylamide gels and transferred to nitrocellulose membrane (Towbin *et al.*, 1979). After blocking with 3% gelatin, 4 mM Tris-HCl, pH 7.4, 100 mM NaCl, the membrane was incubated with a monoclonal antibody specific for the histidine tag (Penta-His antibody; Qiagen) for 4 h in 1% gelatin, 4 mM Tris-HCl, pH 7.4, 100 mM NaCl, 0.05% Tween 20. After washing, bound antibody was detected using a horseradish peroxidase-conjugated secondary antibody followed by enhanced chemiluminescence (Amersham).

Results

The complexes formed by E1 and E2 from HCV strain Glasgow contain intermolecular disulphide bonds

Expression of the structural proteins from HCV strains Glasgow and H77 was achieved using the Semliki Forest virus (SFV) expression system. The parent plasmid used to construct derivatives for expression of HCV proteins was pSFV1, which carries the coding regions for the SFV replication proteins but lacks any structural genes. Details of the regions expressed using the SFV system are shown in Fig. 1. To facilitate purification of E1 and any associated E2, a histidine tag was introduced between amino acids 195 and 196 in the strain

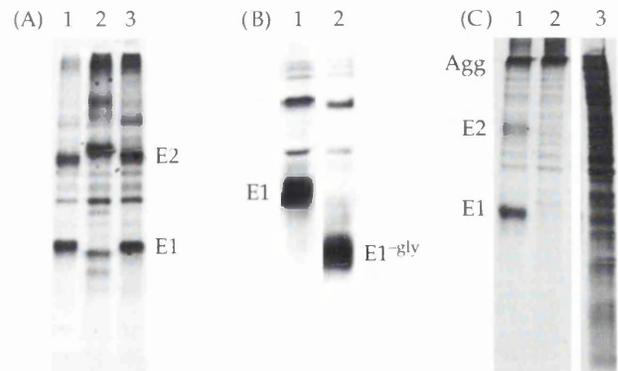


Fig. 2. (A) Comparison of the relative mobilities of E1 and E2 made by strains Glasgow and H77. Cells were electroporated with RNA and labelled between 4 and 12 h after electroporation. After lysis, the glycoproteins were immunoprecipitated from cell extracts using ALP98. Immunoprecipitates were electrophoresed on a polyacrylamide gel under reducing conditions. Samples were derived from cells electroporated with RNA from the following constructs: lane 1, pSFV/E1_{his}E2; lane 2, pSFV/CE1E2_{H77}; lane 3, pSFV/CE1E2_{Gla}. The positions of E1 and E2 are indicated. (B) Sensitivity of glycoprotein E1 to endo H digestion. Cells were electroporated with RNA from pSFV/E1_{his}E2 and radiolabelled as described in (A). After lysis, a portion (1/50) of the cell extract was digested with endo H. Digested (lane 2) and undigested (lane 1) samples were electrophoresed on a polyacrylamide gel under reducing conditions and then examined by Western blot analysis using the Penta-His antibody, which recognizes the histidine tag in E1. The positions of glycosylated (E1) and de-glycosylated forms (E1^{-gly}) of E1 are indicated. (C) Analysis of the E1-E2 complex formed by strain Glasgow under reducing and non-reducing conditions. Cells were electroporated with pSFV/E1_{his}E2 and labelled as described in (A). A cell extract was prepared and subjected to Ni-NTA affinity chromatography. Bound proteins were electrophoresed on a polyacrylamide gel under reducing (lane 1) and non-reducing (lane 2) conditions. Lane 3 shows the crude extract applied to the Ni-NTA agarose column. The positions of E1, E2 and a high molecular mass aggregate (Agg) which fails to enter the resolving gel are shown.

Glasgow sequence, four amino acids from the cleavage site which defines the N terminus of E1; this modification gave a construct called pSFV/E1_{his}E2.

Transcripts synthesized *in vitro* were introduced by electroporation into BHK C13 cells to achieve expression of E1 and E2. Expression of E1 and E2 was reproducibly detected in 60–90% of electroporated cells as determined by indirect immunofluorescence (data not shown). Extracts from cells analysed by Western blot analysis using E1- and E2-specific antibodies indicated that the strain Glasgow glycoproteins had molecular masses of about 35 kDa (for E1) and 65 kDa (for E2; data not shown). Comparison of the mobilities of strain Glasgow glycoproteins with their counterparts in strain H77 was analysed using an E2-specific antibody which immunoprecipitates not only E2 but also E1 that is complexed to E2 (Fig. 2A). This revealed that E2 from strain Glasgow has a greater mobility than E2 from H77 while the converse is the situation for the apparent molecular masses of the corresponding E1 proteins made by the two strains (Fig. 2A, compare lanes 1 and 2). The difference in mobilities for the E1 glycoproteins is not a consequence of the presence of the histidine tag since untagged E1 made by strain Glasgow also

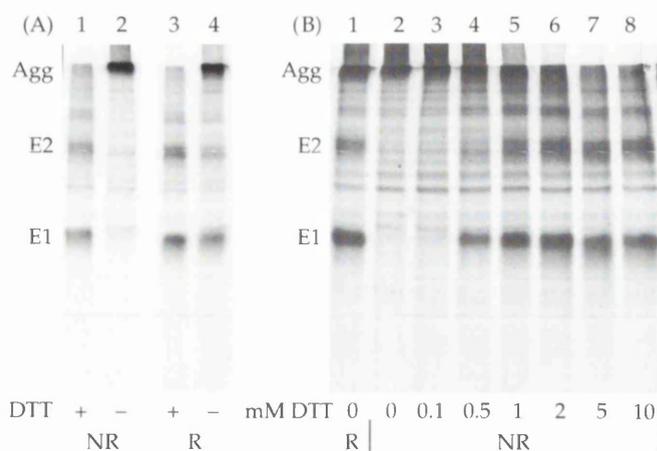


Fig. 3. Effect of DTT on E1–E2 complexes formed by strain Glasgow. (A) Cells were electroporated with pSFV/E1_{his}E2 RNA, radiolabelled and treated with cycloheximide and DTT as described in the text. Cell extracts were subjected to Ni–NTA affinity chromatography and bound proteins were electrophoresed on a polyacrylamide gel under non-reducing (NR; lanes 1 and 2) and reducing (R; lanes 3 and 4) conditions. Samples derived from cells treated with (+) or without DTT (–) are indicated. (B) Experimental details are as described for (A) except that the concentration of DTT added to cells was varied. The concentration of DTT added to cells is shown below each lane. Following Ni–NTA affinity chromatography, samples were electrophoresed on a polyacrylamide gel under reducing (R; lane 1) and non-reducing conditions (NR; lanes 2 to 7). The positions of E1, E2 and high molecular mass aggregates (Agg) which fail to enter the resolving gel are shown in (A) and (B).

has a greater apparent molecular mass than H77 E1 (Fig. 2A, lanes 2 and 3). From comparison of the amino acid sequences for the two strains, E2 from strain Glasgow lacks two predicted glycosylation sites which are present in the H77 sequence (Ogata *et al.*, 1991; M. McElwee & R. M. Elliott, unpublished data). However, the difference in the apparent mobilities of E1 between the strains does not appear to result from variations in either the predicted molecular masses or glycosylation patterns. Hence, at present the basis for the different mobilities of E1 from strain Glasgow and H77 is unclear. It was noted also that the E2 species for both strains often resolved into two bands (see Fig. 2A and Fig. 3A, B). These may arise from inefficient cleavage at the E2/p7 site. Such species have been identified previously for strain H (Dubuisson *et al.*, 1994; Lin *et al.*, 1994), which is highly homologous to strain H77. In strain Glasgow, however, we cannot rule out the possibility that they result from differential glycosylation of E2. For simplicity, these species will be referred to as E2. From indirect immunofluorescence studies, both proteins made by the Glasgow strain were localized to and retained in the ER (data not shown). ER retention was confirmed by endo H cleavage of extracts which showed that E1 was sensitive to enzyme digestion (Fig. 2B, lanes 1 and 2); E2 displayed similar sensitivity to endo H cleavage (data not shown). These data are consistent with previous findings (Cocquerel *et al.*, 1998).

The above data clearly demonstrate that, in accordance with the glycoproteins made by other HCV strains, E1 and E2 from the Glasgow strain form a complex. The nature of the

interaction between the proteins was further studied by comparing their mobilities under reducing and non-reducing electrophoresis conditions. On this occasion, extracts from cells electroporated with pSFV/E1_{his}E2 RNA were mixed with Ni–NTA agarose from which bound proteins were eluted. Examination of bound material under reducing electrophoresis conditions revealed the presence of two major species of 35 and 65 kDa (Fig. 2C, lane 1) which were confirmed to be histidine-tagged E1 and E2 respectively by Western blot analysis (data not shown). However, little monomeric E1 and E2 was detected under non-reducing conditions and most of the Ni–NTA bound material failed to migrate into the resolving component of the gel (Fig. 2C, lane 2), indicative that the complex contains intermolecular disulphide bonds. This property for the E1–E2 complex made by strain Glasgow was not a consequence of the histidine tag at the N-terminal end of E1 since complexes composed of untagged E1 and E2 behaved identically under non-reducing electrophoresis conditions (data not shown). Other studies have shown a similar behaviour for the E1–E2 complex although the proportion of E1 and E2 present as a disulphide-linked complex as compared to a non-covalently associated complex varies from study to study (Grakoui *et al.*, 1993; Ralston *et al.*, 1993; Dubuisson *et al.*, 1994). While these differences may reflect in part the expression system and cell type used, the high proportion of disulphide-linked complexes formed by the Glasgow strain E1 and E2 proteins are likely to result from properties inherent to their sequences.

Treatment of cells with DTT reveals non-covalent interactions between E1 and E2

It has been proposed that the disulphide-linked complexes formed by E1 and E2 represent aggregates of misfolded proteins (Dubuisson *et al.*, 1994; Deleersnyder *et al.*, 1997); however, their contribution to virion structure and the assembly process are not known since direct analysis of HCV particles has not been possible. To further examine the contribution of disulphide bonds to E1–E2 complex formation, we analysed the properties of the complex under reducing conditions *in situ*. From previous studies, treating cells *in situ* with DTT reversibly reduces disulphide bonds within proteins (Braakman *et al.*, 1992). Therefore, we devised a series of experiments employing this method to analyse the effect of such treatment on the strain Glasgow E1–E2 complexes. Initially, we investigated the effect of DTT addition on the stability of the preformed complexes. Cells were electroporated with pSFV/E1_{his}E2 RNA and proteins were radiolabelled with [³⁵S]methionine 4 h after electroporation. Labelling was continued for a further 8 h at which point cycloheximide was added to block translation. Following incubation for 30 min in the presence of cycloheximide, DTT was added to cells at a final concentration of 5 mM and cells were harvested after a further 30 min incubation; in a parallel culture, cells were

treated with cycloheximide but not with DTT. Extracts were then prepared in the presence of NEM, an alkylating agent which blocks free sulphhydryl groups. Following purification by Ni-NTA affinity chromatography, bound proteins were analysed under both reducing and non-reducing conditions.

The E1-E2 complexes purified from cells that had not been treated with DTT behaved in an identical manner to those shown in Fig. 2(C); E1 and E2 migrated as monomers under reducing conditions (Fig. 3A, lane 4) but mainly as a high molecular mass complex in the absence of reducing agent (lane 2). From the extracts of cells treated with DTT, E2 continued to co-purify with histidine-tagged E1 as both proteins are found in gels run under reducing conditions (Fig. 3A, lane 3). Significantly, under non-reducing conditions, E1 and E2 no longer migrated as a high molecular mass complex but rather as monomeric proteins (Fig. 3A, lane 1). From these data, we conclude that the intermolecular disulphide bonds within the E1-E2 complex have been reduced. However, disruption of these covalent interactions did not dissociate the pre-formed E1-E2 complex. Hence, splitting the covalent interactions uncovered non-covalent associations which are present in E1-E2 complexes that are apparently aggregated.

Following this observation, we sought to further characterize the effect of DTT on E1-E2 complexes. In a subsequent experiment, the result of varying the concentration of DTT on the high molecular mass material and the non-covalent E1-E2 interactions was examined. Again, radiolabelled complexes bound to Ni-NTA were analysed under non-reducing conditions to monitor the loss of high molecular mass material. At lower concentrations of DTT (0.5 and 1.0 mM; Fig. 3B, lanes 4 and 5), there was evidence that reduction of E1 and E2 had occurred *in situ*; however, a considerable amount of high molecular mass material remained. At higher concentrations of DTT (5 and 10 mM), the high molecular mass material had disappeared and the predominant species present were monomeric E1 and E2 (Fig. 3B, lanes 7 and 8). Thus, even at concentrations of DTT above 5 mM, the non-covalently linked E1 and E2 complex is stable. In subsequent experiments, DTT was added to cells at a final concentration of 5 mM, a concentration which is consistent with that used in other studies (Braakman *et al.*, 1992). Analysis of the time required to fully reduce the high molecular mass complex in 5 mM DTT showed that reduction is observed by 5 min following the addition of DTT to cells and is complete by 20 min (data not shown). Therefore, our results provide strong evidence that the non-covalent interactions between E1 and E2 are not stabilized by intermolecular disulphide bonds present in aggregates of the complex.

The high molecular mass material rapidly reforms upon removal of DTT from cells

To determine whether reduction of the disulphide bonds in E1 and E2 mediated a change in either the conformation of the

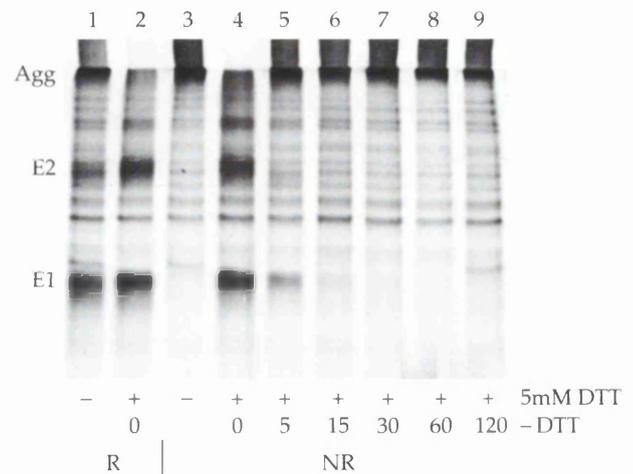


Fig. 4. Effect of removal of DTT on E1-E2 complexes. Cells were electroporated with pSFV/E1_{his}E2 and radiolabelled between 4 and 12 h after electroporation. Cycloheximide and DTT were added to cells as described for Fig. 3 (A). Samples treated with 5 mM DTT for 30 min (+) or without DTT (-) as indicated. DTT-treated samples were washed with media lacking the reducing agent and then incubated at 37 °C in the absence of DTT for the times indicated (-DTT). Cell extracts were subjected to Ni-NTA affinity chromatography and bound proteins were electrophoresed on a polyacrylamide gel under reducing (R; lanes 1 and 2) and non-reducing conditions (NR; lanes 3-9).

E1 and E2 proteins or the complex which they form, the effect of removal of DTT was examined. Cells expressing E1 and E2 synthesized from pSFV/E1_{his}E2 RNA were treated with cycloheximide and DTT, after which the cultures were washed with medium lacking DTT and then harvested at various times following removal of the reducing agent. We found that high molecular mass material could be detected within 5 min of DTT removal (Fig. 4, compare lanes 4 and 5) and it was the predominant species by 15 min (lane 6). Hence, the effect of DTT is reversible and disulphide bond formation recurs upon its removal. However, non-aggregated forms of the complex are not recovered, suggesting that even transient release from possibly incorrect intermolecular covalent linkages does not allow re-ordering of the non-covalent interactions to generate complexes which do not aggregate.

Disulphide bond formation is not required for the initial interactions between E1 and E2

The data described thus far examined the function of disulphide bonds in pre-existing complexes. To address any role for such covalent linkages in initial interactions between E1 and E2, we analysed complexes formed by *de novo* synthesized E1 and E2 in the presence and absence of DTT. Electroporated cells were pulse-labelled for 2 min with [³⁵S]methionine/cysteine and then chased for up to 10 min. To prevent disulphide bond formation during the pulse-chase period, DTT was added to one set of cells 15 min prior to addition of the radiolabelled amino acids and was maintained

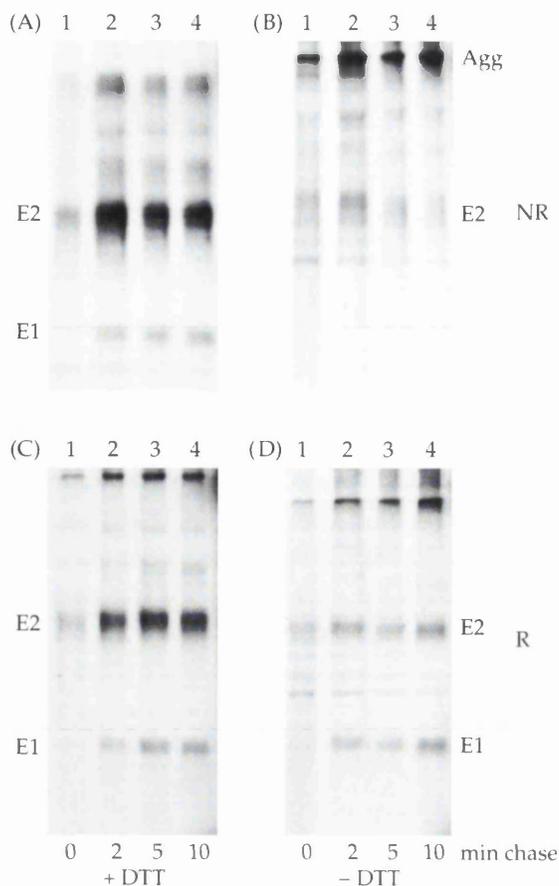


Fig. 5. Formation of *de novo* synthesized E1–E2 complexes in the presence and absence of DTT. Cells were electroporated with pSFV/E1_{his}E2 and incubated at 37 °C for 12 h. Cells were labelled for 2 min and then chased for the times indicated. In panels (A) and (C) samples are from cells incubated in the presence of DTT from 15 min prior to labelling until they were harvested. Immunoprecipitations were performed on cell extracts with anti-E2 antiserum R141 and bound proteins were electrophoresed on polyacrylamide gels under non-reducing (NR; panels A and B) and reducing (R; panels C and D) conditions.

until the cells were harvested. Immunoprecipitation with E2-specific antiserum R141 showed that E1 was complexed to E2 at 2 min after the pulse in both DTT-treated and non-treated samples (Fig. 5C and D, lane 2). The presence of monomeric E1 and E2 in non-reducing gels verified that intermolecular disulphide bonds had not formed in proteins synthesized in the presence of DTT (Fig. 5A, lanes 1–4). By comparison, in samples not treated with DTT, only small quantities of E2 and no E1 could be detected in non-reducing gels and the abundance of monomeric E2 protein decreased at longer chase times (compare lanes 1–4 in Fig. 5B). The lack of monomeric E1 and the reduced amount of monomeric E2 result from the formation of high molecular mass material which barely enters the resolving gel. Similar data were obtained by Ni–NTA chromatography of E1–E2 complexes (data not shown). From these data, we conclude that non-covalent association of E1 and E2 occurs rapidly following synthesis and disulphide bond formation is not necessary for such interactions.

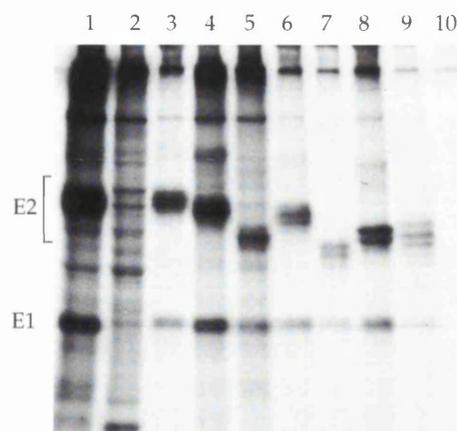


Fig. 6. Complexes formed by E2 deletion mutants in the presence of DTT. Cells were electroporated with RNA from each construct and, at 12 h post-electroporation, they were treated with DTT 15 min prior to labelling. Cells were pulse labelled for 2 min and then chased for 10 min, all in the presence of DTT. Immunoprecipitations were performed with anti-E2 antiserum R141 and bound proteins were electrophoresed on a polyacrylamide gel under non-reducing conditions. Lanes 1 and 2 contain crude extracts from two cellular preparations and the remaining lanes show bound proteins following immunoprecipitation. Samples were derived from cells electroporated with RNA from the following constructs: lanes 1 and 3, pSFV/E1_{his}E2; lanes 2 and 10, pSFV1; lane 4, pSFV/Δ385–410; lane 5, pSFV/Δ416–453; lane 6, pSFV/Δ460–519; lane 7, pSFV/Δ529–608; lane 8, pSFV/Δ610–697; lane 9, pSFV/169–703.

Effect of DTT on complexes formed by E2 deletion mutants

It was anticipated that reduction of E1 and E2 *in situ* may provide a means of distinguishing between regions in the proteins which interact by either covalent or non-covalent interactions. Hence, a series of deletion mutants was derived from pSFV/E1_{his}E2 which sequentially removed regions of E2 (Fig. 1). From several experiments with these mutants using either immunoprecipitation or Ni–NTA affinity chromatography, we found that E2 mutant polypeptides retained the capacity to interact with E1 both in the presence and absence of DTT. A typical set of data obtained is shown in Fig. 6. Here, complex formation has been assessed by Ni–NTA affinity chromatography of radiolabelled extracts prepared from DTT-treated cells electroporated with the E2 mutants shown in Fig. 1. Each mutant E2 species co-purified with E1 suggesting that E1 and E2 can non-covalently associate at multiple sites.

Effect of DTT treatment on E1–E2 complexes formed by HCV strain H77

During the course of our studies, it was reported that RNAs synthesized from cDNAs composed of the consensus sequence for HCV strain H77 were infectious in chimpanzees (Kolykhalov *et al.*, 1997; Yanagi *et al.*, 1997). Published data also showed that both covalently and non-covalently linked E1–E2 complexes could be identified with the HCV H strain, which was isolated from a chimpanzee infected with strain H77 and has very high homology to H77 (Inchauspe *et al.*, 1991;

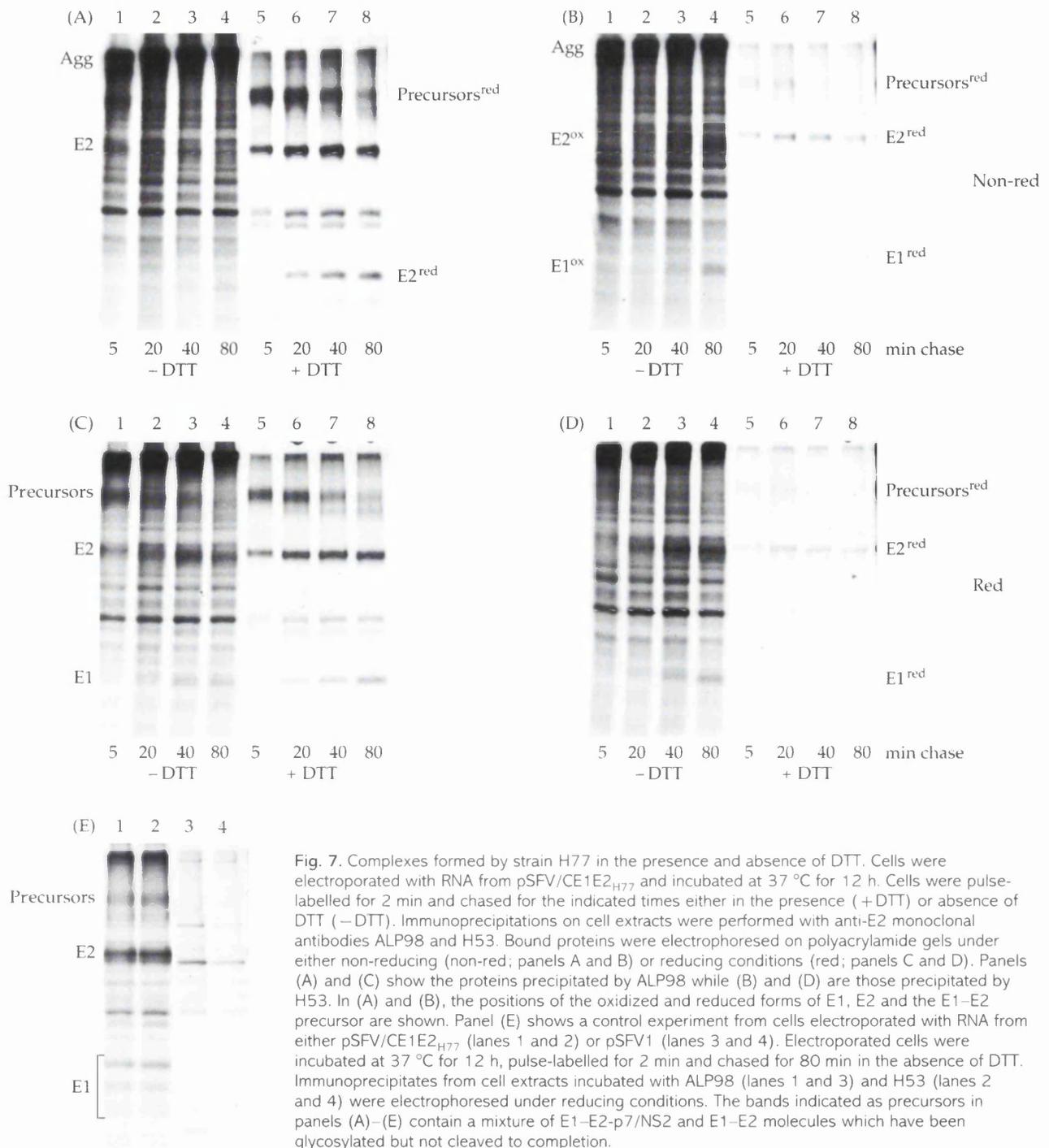


Fig. 7. Complexes formed by strain H77 in the presence and absence of DTT. Cells were electroporated with RNA from pSFV/CE1E2_{H77} and incubated at 37 °C for 12 h. Cells were pulse-labelled for 2 min and chased for the indicated times either in the presence (+DTT) or absence of DTT (-DTT). Immunoprecipitations on cell extracts were performed with anti-E2 monoclonal antibodies ALP98 and H53. Bound proteins were electrophoresed on polyacrylamide gels under either non-reducing (non-red; panels A and B) or reducing conditions (red; panels C and D). Panels (A) and (C) show the proteins precipitated by ALP98 while (B) and (D) are those precipitated by H53. In (A) and (B), the positions of the oxidized and reduced forms of E1, E2 and the E1-E2 precursor are shown. Panel (E) shows a control experiment from cells electroporated with RNA from either pSFV/CE1E2_{H77} (lanes 1 and 2) or pSFV1 (lanes 3 and 4). Electroporated cells were incubated at 37 °C for 12 h, pulse-labelled for 2 min and chased for 80 min in the absence of DTT. Immunoprecipitates from cell extracts incubated with ALP98 (lanes 1 and 3) and H53 (lanes 2 and 4) were electrophoresed under reducing conditions. The bands indicated as precursors in panels (A)–(E) contain a mixture of E1-E2-p7/NS2 and E1-E2 molecules which have been glycosylated but not cleaved to completion.

Dubuisson *et al.*, 1994). Moreover, the non-covalently linked complex could be recognized by conformation-specific antibodies (Deleersnyder *et al.*, 1997; Michalak *et al.*, 1997; Cocquerel *et al.*, 1998). For comparative purposes, we examined the effect of DTT on E1-E2 complex formation in strain H77. Cells were electroporated with RNA from pSFV/CE1E2_{H77} and a pulse-chase experiment performed on cultures either in the presence or absence of DTT. Immunoprecipitations were carried out with two antibodies: ALP98, which detects a linear epitope in E2, and H53, which recognizes a conformation-

dependent native form of E2 (Cocquerel *et al.*, 1998). These antibodies precipitated both E2 and E1, which often migrated as a series of differentially glycosylated species (Fig. 7E, compare lanes 1 and 2 with lanes 3 and 4). The identity of E1 was confirmed by Western blot analysis using an E1-specific antibody (data not shown).

In the absence of DTT, H53 recognized a native form of the complex, as deduced from electrophoresis under non-reducing conditions, more efficiently than ALP98 (compare lanes 1–4 in Fig. 7A and B). This complex was most apparent at later times

in the pulse-chase experiment, in agreement with previously reported data (Deleersnyder *et al.*, 1997). We also found that, in these experiments, H53 precipitated some aggregated material (Fig. 7B and D, lanes 1–4). In the presence of DTT, both antibodies precipitated E1 as well as E2 even at early chase times (see lanes 5 and 6 in Fig. 7A and B), although ALP98 was more efficient than H53; this may reflect the specificity of H53 for a particular conformation of E2. Nonetheless, treatment with DTT does not disrupt the conformation detected by H53 and, in agreement with our earlier data, does not prevent E1 association with E2. It is also of interest to note that H53 recognized E1–E2 precursor proteins in DTT-treated samples (Fig. 7B and D, lanes 5 and 6) but with markedly reduced efficiency in untreated samples (Fig. 7B and D, lanes 1–4). By contrast, ALP98 recognized these precursors with greater efficiency than H53 in both DTT-treated and untreated samples (Fig. 7A and C, lanes 1–8). Coupled with the ability of H53 to precipitate E2 soon after synthesis (Fig. 7B and D, lane 5), it would appear that the conformation detected by the antibody can form rapidly in the presence of reducing agent.

From comparison of the relative mobilities of E1 and E2 in Fig. 7(B), it is evident that oxidized forms of both E1 and E2 are present in untreated samples (lane 4) while only the reduced forms are detected in those samples treated with DTT (lanes 6–8). This indicates that intramolecular disulphide bonds do not form in the presence of DTT, further highlighting that E1 and E2 association does not rely upon the formation of covalent linkages.

Discussion

We have examined the effect of creating reducing conditions within cells on E1–E2 complex formation. In agreement with a previous report which examined E2–NS2 cleavage (Dubuisson & Rice, 1996), we find that addition of DTT to cells does not impair cleavage and processing of the glycoproteins from precursor molecules. However, we have considerably expanded this finding to examine the wider role of disulphide bonds in E1–E2 complex formation. Given the lack of a system for efficient propagation of HCV and the difficulties associated with recovering virus particles from infected sera, it is difficult to formulate a pathway for E1–E2 interactions and the folding which occurs *in vivo*. To date, two complexes have been identified: an aggregate in which intermolecular covalent interactions occur and a native complex composed of non-covalently associated E1 and E2 (Deleersnyder *et al.*, 1997; Michalak *et al.*, 1997). While there is growing belief that non-covalent associations represent the authentic interactions between E1 and E2, a contribution of covalent interactions to E1–E2 complexes remains a possibility. Hence, analysis of the role of covalent interactions may shed light on the intermediates which could occur during E1–E2 association.

E1 and E2 from HCV strain H associate to give both types of complex. The native complex can be detected by conformation-specific antibodies and forms more slowly than the aggregate (Deleersnyder *et al.*, 1997; Michalak *et al.*, 1997). Using the SFV system and a conformation-specific antibody, we also have detected both types of E1–E2 complex with strain H77, which is highly homologous to strain H. Aggregates containing associated E1 and E2 occur soon after synthesis and processing of the polypeptides while the native form appears later. With strain Glasgow, there was reduced accumulation of the native complex. We conclude that this occurs as a result of sequence differences between the two strains. Comparison of the predicted amino acid sequences from both strains showed that the cysteine residues are conserved in both proteins as are the predicted *N*-glycosylation sites in E1. However, E2 from strain Glasgow is predicted to have two fewer *N*-glycosylation sites. The loss of these modification sites could at least partially account for the reduced ability of the Glasgow strain to form native E1–E2 complexes. E1 glycoprotein from strain Glasgow also has lower mobility on gels as compared to its counterpart in H77. Since the predicted molecular masses of the non-glycosylated species are very similar and the predicted glycosylation patterns of the two proteins are identical, there is no obvious reason for the different mobilities. Moreover, the endo H digestion products of H77 and strain Glasgow E1 do not co-migrate (data not shown); thus the observed molecular mass difference is not due to differences in glycosylation. Our observations are similar to those noted by Fournillier-Jacob *et al.* (1996), in which the mobility of E1 was dependent on HCV sequences flanking the polypeptide. We suggest that there may be additional processing to the E1 protein from strain Glasgow which, as yet, has not been identified, and this also may contribute to the reduced production of a native complex by this strain.

The occurrence of non-covalent interactions between E1 and E2 within the aggregated complex has not been previously reported. From our results, covalent interactions within the aggregate can be disrupted by DTT *in situ* yet E1 remains non-covalently attached to E2. One interpretation of these data is that reduction of covalent bonds by DTT exposes regions of E1 and E2 which can then interact non-covalently. Alternatively, E1 and E2 may form homo-oligomers which retain the ability to interact non-covalently with each other. We have expressed separately both E1 and E2 and find that both proteins multimerize through covalent interactions (data not shown). A third option is that covalent and non-covalent associations can occur simultaneously between individual E1 and E2 molecules. Under such circumstances, it would be predicted that segments in E1 and E2 which form either type of interaction would be separable and thus identified by deletion analysis. However, we have been unable to identify such regions from our E2 mutants. This suggests that the interactions which can occur between E1 and E2 are complex.

It is not possible to state whether the non-covalent interactions which are present in the native complex also participate in the formation of the aggregate. If there are two sets of interactions which are mutually exclusive, each specific for either native or aggregated complexes, then caution should be exercised on the authenticity of regions which have been identified as necessary for E1–E2 association. This may be particularly the case where methods such as Far-Western blot analysis have been used to map sequences required for E1–E2 binding but which do not directly analyse the nature of the complex (Yi *et al.*, 1997).

Our data with E2 deletion mutants failed to identify any regions in the protein whose removal prevented complex formation with E1. This included a truncated form of E2 lacking the C-terminal transmembrane domain. Previous results using the vaccinia virus expression system have indicated that deletion of the equivalent region in E2 abolishes interaction with E1 (Selby *et al.*, 1994; Michalak *et al.*, 1997). However, unlike the vaccinia virus system, E2 truncated at residue 703 was not secreted in our SFV system; similar data have been observed also on expression of truncated forms of E2 with a Sindbis virus vector (Michalak *et al.*, 1997). Thus, truncated E2 produced by SFV may reside within an intracellular compartment where it retains the capacity to interact with E1 through the E2 ectodomain. By contrast, vaccinia virus expression of truncated E2 may direct the protein to a different compartment as part of the secretory process, thereby reducing the ability to associate with E1. We note also that our truncated form of E2 removes sequences which would encode p7. Since E1 and E2 continue to form a complex in the absence of these sequences, we conclude that p7 is not critical for their interaction.

We have used a conformation-specific antibody which recognizes the native complex and find that this reagent does not lose specificity with *de novo* synthesized complexes produced under reducing conditions. Therefore, the conformation for E2 required for antibody recognition is not dependent on disulphide bonds for stability. By contrast, a recent report showed that, following reduction, a truncated form of E2 was detected less efficiently by antibodies in approximately half of the infected sera tested (Lee *et al.*, 1997). Thus, covalent bonds may function to stabilize certain conformations of E2 but our data would suggest that they are not required for E1–E2 interactions.

In conclusion, we have shown that intra- and intermolecular covalent interactions are not required for non-covalent association of E1 and E2. Since E1 and E2 can form both native and aggregated complexes through non-covalent bonding, this highlights a difficulty in identifying authentic interactions which occur during virus assembly and maturation.

We wish to thank Professor R. M. Elliott, Drs J. Bukh, J. Dubuisson and M. McElwee for supplying materials and Dr S. Graham for assistance in the production of antibodies. We are grateful also to Professor D. J.

McGeoch for valuable comments on the manuscript and Dr S. Butcher for helpful discussions.

References

- Braakman, I., Helenius, J. & Helenius, A. (1992).** Manipulating disulphide bond formation and protein folding in the endoplasmic reticulum. *EMBO Journal* **11**, 1717–1722.
- Carleton, M. & Brown, D. T. (1996).** Disulphide bridge-mediated folding of Sindbis virus glycoproteins. *Journal of Virology* **70**, 5541–5547.
- Carleton, M. & Brown, D. T. (1997).** The formation of intramolecular disulphide bridges is required for induction of the Sindbis virus mutant ts23 phenotype. *Journal of Virology* **71**, 7696–7703.
- Clarke, B. E. (1997).** Molecular virology of hepatitis C virus. *Journal of General Virology* **78**, 2397–2410.
- Cocquerel, L., Meunier, J. C., Pillez, A., Wychowski, C. & Dubuisson, J. (1998).** A retention signal necessary and sufficient for endoplasmic reticulum localization maps to the transmembrane domain of hepatitis C virus glycoprotein E2. *Journal of Virology* **72**, 2183–2191.
- Deleersnyder, V., Pillez, A., Wychowski, C., Blight, K., Xu, J., Hahn, Y. S., Rice, C. M. & Dubuisson, J. (1997).** Formation of native hepatitis C virus glycoprotein complexes. *Journal of Virology* **71**, 697–704.
- Dubuisson, J. & Rice, C. M. (1996).** Hepatitis C virus glycoprotein folding: disulphide bond formation and association with calnexin. *Journal of Virology* **70**, 778–786.
- Dubuisson, J., Hsu, H. H., Cheung, R. C., Greenberg, H. B., Russell, D. G. & Rice, C. M. (1994).** Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and Sindbis viruses. *Journal of Virology* **68**, 6147–6160.
- Fournillier-Jacob, A., Cahour, A., Escriou, N., Girard, M. & Wychowski, C. (1996).** Processing of the E1 glycoprotein of hepatitis C virus expressed in mammalian cells. *Journal of General Virology* **77**, 1055–1064.
- Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M. & Rice, C. M. (1993).** Expression and identification of hepatitis C virus polyprotein cleavage products. *Journal of Virology* **67**, 1385–1395.
- Inchauspe, G., Zebedee, S., Lee, D. H., Sugitani, M., Nasoff, M. & Prince, A. M. (1991).** Genomic structure of the human prototype strain H of hepatitis C virus – comparison with American and Japanese isolates. *Proceedings of the National Academy of Sciences, USA* **88**, 10292–10296.
- Kolykhalov, A. A., Agapov, E. V., Blight, K. J., Mihalik, K., Feinstone, S. M. & Rice, C. M. (1997).** Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. *Science* **277**, 570–574.
- Laemmli, U. K. (1970).** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Lee, K. J., Suh, Y.-A., Cho, Y. G., Cho, Y. S., Ha, G. W., Chung, K.-H., Hwang, J. H., Yun, Y. D., Lee, D. S., Kim, C. M. & Sung, Y.-C. (1997).** Hepatitis C virus E2 protein purified from mammalian cells is frequently recognized by E2-specific antibodies in patient sera. *Journal of Biological Chemistry* **272**, 30040–30046.
- Liljestrom, P. & Garoff, H. (1991).** A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. *Biotechnology* **9**, 1356–1361.
- Lin, C., Lindenbach, B. D., Pragai, D., McCourt, D. W. & Rice, C. M. (1994).** Processing of the hepatitis C virus E2–NS2 region: identification of p7 and two distinct E2-specific products with different C-termini. *Journal of Virology* **68**, 5063–5073.
- Matsuura, Y., Suzuki, T., Suzuki, R., Sato, M., Aizaki, H., Saito, I. & Miyamura, T. (1994).** Processing of E1 and E2 glycoproteins of hepatitis C virus expressed in mammalian and insect cells. *Virology* **205**, 141–150.

Michalak, J. P., Wychowski, C., Choukhi, A., Meunier, J. C., Ung, S., Rice, C. M. & Dubuisson, J. (1997). Characterization of truncated forms of hepatitis C virus glycoproteins. *Journal of General Virology* **78**, 2299–2306.

Ogata, N., Alter, H. J., Miller, R. H. & Purcell, R. H. (1991). Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **88**, 3392–3396.

Ralston, R., Thudium, K., Berger, K., Kuo, C., Gervase, B., Hall, J., Selby, M., Kuo, G., Houghton, M. & Choo, Q. L. (1993). Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. *Journal of Virology* **67**, 6753–6761.

Selby, M. J., Glazer, E., Masiarz, F. & Houghton, M. (1994). Complex processing and protein–protein interactions in the E2–NS2 region of HCV. *Virology* **204**, 114–122.

Towbin, H., Staehelin, T & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure

and some applications. *Proceedings of the National Academy of Sciences, USA* **76**, 4350–4354.

Verde, C., Pascale, M. C., Martire, G., Lotti, L. V., Torrisi, M. R., Helenius, A. & Bonatti, S. (1995). Effect of ATP depletion and DTT on the transport of membrane proteins from the endoplasmic reticulum and the intermediate compartment to the Golgi complex. *European Journal of Cell Biology* **67**, 267–274.

Yanagi, M., Purcell, R. H., Emerson, S. U. & Bukh, J. (1997). Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. *Proceedings of the National Academy of Sciences, USA* **94**, 8738–8743.

Yi, M. K., Nakamoto, Y., Kaneko, S., Yamashita, T. & Murakami, S. (1997). Delineation of regions important for heteromeric association of hepatitis C virus E1 and E2. *Virology* **231**, 119–129.

Received 12 January 1999; Accepted 6 April 1999

