Characterisation of the Hepatitis C Virus Genotype 3 Glycoproteins

by

Megan Louise Shaw

A thesis presented for the degree of Doctor of Philosophy

in

The Faculty of Medicine
at the
University of Glasgow

Institute of Virology
University of Glasgow
Church St
Glasgow
G11 5JR

December 2001

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

Acknowledgements

Firstly, I would like to thank Prof. Richard Elliott and Prof. Duncan McGeoch for providing research facilities in the Institute of Virology and the University of Glasgow and ORS awards scheme for funding me.

I would like to thank my supervisor, Dr Liz McCruden for her advice and encouragement throughout this project and especially for her assistance in proof-reading this thesis so promptly. I also wish to thank Prof. Richard Elliott for several helpful discussions during the course of this project.

I am indebted to Dr Arvind Patel and Dr John McLauchlan for providing numerous reagents, for their helpful advice and for their critical reading of parts of this thesis. Thanks also go to Dr Janisha Patel for providing the genotype 1 SFV constructs and for her technical assistance at the beginning of the project and to Dr Reg Clayton for his assistance with the confocal microscopy.

I would like to thank Dr Alex Cochrane and Mr Alex Orr for heading the PePHD study and Suzanne Barnard for analysing the activity of the interferon-alpha. Thank you to Dr Hovanessian for providing the PKR monoclonal antibody.

I wish to thank all past and present members of lab 311/301 and lab 300 for both technical assistance and the endless banter which relieved the monotony of an average day. Special thanks go to Alex Orr for showing me the ropes at the beginning and to Carol-Anne Smith for looking after the sera so expertly and for doing all the PCR for the CD81-binding experiment.

Many thanks to all my friends in the Institute and elsewhere in the university who have made my stay in Glasgow so enjoyable, especially Maria and Iain who have made the bad days seem not so bad and the good days seem even better. A big thank you to all past and present flatmates of 69E who have been my surrogate family, particularly Houda who has been on the receiving end of most of my moans and groans – you're a star.

Finally, I wish to thank Mum, Dad, Christopher and Jonathan for providing endless support and encouragement and for keeping me up to date with events on the other side of the world, especially the cricket score!

Contents

Chapter 5 – Interaction of HCV E2 with PKR and PERK.

Chapter 6 – Binding of genotypes 1 and 3 E2 to human CD81.

Chapter 7 – General Discussion.

References

List of Figures and Tables

Abbreviations

%	percentage
°C	degree Centigrade
β-ME	β-mercaptoethanol
µF	microfaraday
µg	microgram(s)
µl	microlitre(s)
^{35}S	sulphur isotope (35)
2YT broth	2 yeast tryptone broth
aa	amino acid(s)
APS	Ammonium persulphate
as	antisense
(d)ATP	(2'-deoxy)adenosine 5'-triphosphate
BHK	Baby hamster kidney cells
bp	base pair
BSA	Bovine serum albumin
C-terminal	Carboxy-terminal
$CaCl_2$	Calcium chloride
cDNA	Complementary DNA
Ci	Curie(s)
CIP	Calf intestinal phosphatase
cm	centimetre
CO_2	Carbon dioxide
(d)CTP	(2'-deoxy)cytidine 5'-triphosphate
dH_2O	Distilled water
DMEM	Dulbecco's modified Eagles Medium
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	dithiothreitol
E1	envelope glycoprotein 1
E2	envelope glycoprotein 2
E.coli	*Escherichia coli*
ECL	Enhanced chemiluminescence
EDTA	ethlenediaminetetra-acetic acid (disodium salt)
Endo H	endoglycosidase H
ELISA	Enzyme-linked immunosorbant assay
ER	Endoplasmic reticulum
FCS	Foetal calf serum
g	gram
GMEM	Glasgow minimal Eagles Medium

GNA	*Galanthus nivalis* agglutinin
GST	glutathione S-transferase
(d)GTP	(2'-deoxy)guanosine 5'-triphosphate
H_2O	Water
H_2SO_4	Sulphuric acid
HCl	Hydrochloric acid
HRP	Horseradish peroxidase
Huh7	Human hepatocellular carcinoma
IFN	Interferon
IgG	Immunoglobulin G
IP	Immunoprecipitation
IU	International units
K or kDa	kilodalton
kb	kilobase pair
KCl	Potassium chloride
KH_2PO_4	Potassium dihydrogen orthophosphate
KOAc	Potassium acetate
KOH	Potassium hydroxide
kV	kilovolts
L-broth	Luria Bertani medium
M	molarity
MAb	monoclonal antibody
$MgCl_2$	Magnesium chloride
$Mg(OAc)_2$	Magnesium acetate
$MgSO_4$	Magnesium sulphate
min	minute(s)
ml	millilitre(s)
mM	millimolar
mm	millimetre(s)
N-linked	Asparagine-linked
N-terminal	Amino-terminal
NaCl	Sodium chloride
Na_2HPO_4	di-sodium hydrogen orthophosphate anhydrous
NaOH	Sodium hydroxide
NBCS	New born calf serum
NEM	N-ethylmaleimide
$(NH_4)_2SO_4$	Ammonium sulphate
ng	nanogram(s)
nm	nanometre(s)
NP-40	nonidet P40
NS	non-structural
(d)NTP	(deoxy)nucleotide triphosphate
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline

PCR	Polymerase Chain Reaction
PCR 1	First round PCR
PCR 2	Second round PCR
PePHD	PKR and eIF2α phosphorylation homology domain
PKR	RNA-dependent protein kinase
pmoles	picomoles
PMSF	phenylmethylsulphonyl fluoride
PNGase F	peptide *N*-glycosidase F
PolyAb	polyclonal antibody
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
RNasin	Ribonuclease inhibitor
rpm	revolutions per minute
RT	Reverse transcription
RT-PCR	Reverse transcription PCR
s	sense
SDS	sodium dodecyl sulphate
sec	second(s)
SFV	Semliki forest virus
SGB	Stacking gel buffer
TBE	Tris-boric acid-EDTA buffer
TEMED	N,N,N'N-tetramethylethylenediamine
TMB	Tetramethyl benzidine
TP broth	Tryptose phosphate broth
(d)TTP	(2'-deoxy)thymidine 5'-triphosphate
U	unit
UTP	uridine 5'-triphosphate
UTR	untranslated region
UV	ultra violet
V	volts
v/v	volume per volume
w/v	weight per volume

Reference HCV isolates:

Genotype 1a	H77	Accession no: AF011751
	HCV-1a Gla (Gla)	R.M. Elliott, unpublished
Genotype 3a	NZL1	Accession no: D17763
	HPCHK6	Accession no: D28917
	HCVCENS1	Accession no: X76918
	AF046866	Accession no: AF046866

One and three letter abbreviations for amino acids

Amino acid	**Three letter code**	**One letter code**
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Summary

Hepatitis C virus (HCV) can be classified into six genotypes (1-6) which show 30% nucleotide sequence variability throughout the genome. HCV genotypes 1, 2 and 3 have a world-wide distribution but their prevalence differs from one geographical area to another. In Scotland there is an approximate 50/50 split between individuals infected with HCV genotype 1 and genotype 3. There is little evidence that disease progression or severity differs between the genotypes. On the other hand, one difference which has been consistently demonstrated is the better response of patients infected with genotypes 2 and 3 to interferon treatment than those infected with genotype 1. HCV encodes three structural proteins, core, E1 and E2. The two glycoproteins, E1 and E2, are believed to be the envelope proteins. The lack of a cell culture system for the production of HCV virions has meant that the proteins within the envelope of the HCV virion remain uncharacterised. However, recombinant forms of the E1 and E2 proteins have been shown to localise to the endoplasmic reticulum (ER) when expressed in cultured cells and to interact with one another to form a complex (Ralston *et al.*, 1993). E2 has also recently been shown to bind to the cell-surface molecule, CD81, which, as a result, has been proposed to be an HCV receptor (Pileri *et al.*, 1998). These characteristics of the HCV glycoproteins have been described using constructs derived from genotype 1 only. Comparison of the properties of E1 and E2 with those of other genotypes would identify both conserved features, which are possibly essential to the virus life cycle, and genotype-specific features. In this study, the E1 and E2 proteins of genotype 3 were compared with those of genotype 1 with respect to antibody recognition, subcellular localisation, complex formation, glycosylation status and CD81-binding. The role of E2 in mediating resistance to interferon (IFN) through a proposed interaction with PKR (Taylor *et al.*, 1999) was also investigated.

The genotype 3 structural genes were amplified by reverse-transcription polymerase chain reaction (RT-PCR) from the serum of an HCV genotype 3a-infected patient. This isolate was named HCV 3a-Gla (Gla-3a) and was used as the genotype 3 representative, whereas genotype 1 was represented by the H77 isolate. Sequence comparison of Gla-3a with four published genotype 3a isolates and the genotype 1 isolate revealed that the genotype 3 E2 protein was 6 amino acids longer and had one fewer cysteine residue than that of genotype 1. In terms of length and the number of cysteine residues, the E1 proteins of genotypes 1 and 3 were identical. E2 of Gla-3a (and three other genotype 3 isolates) had one fewer predicted glycosylation site compared to genotype 1, whereas the predicted glycosylation sites in E1 were identical between genotypes 3 and 1.

Expression of the core, E1 and E2 proteins was achieved using either the Semliki Forest Virus (SFV) expression system for production of large amounts of recombinant protein or the pcDNA system, in which recombinant gene expression is under control of a cytomegalovirus (CMV) promoter. None of the genotype 1 anti-E1 antibodies were able to recognise genotype 3 E1. Three genotype 1 anti-E2 monoclonal antibodies recognised genotype 3 E2 and they all recognised epitopes in a highly conserved sixteen amino acid region. The genotype 3 glycoproteins localised to the ER and showed evidence of both aggregate and native complex formation as described for those of genotype 1. The relative mobilities of untreated and glycosidase-treated proteins confirmed the predicted glycosylation status of the E2 proteins of genotypes 1 and 3 but the genotype 3 E1 protein appeared to have one additional glycan compared to genotype 1 E1. The deglycosylated genotype 1 E1 protein also seemed smaller than that of genotype 3. Neither of these observations was predicted from the sequence comparison. In contrast to genotype 1 E2, the recombinant genotype 3 E2 protein did not bind to human CD81, although interestingly, preliminary results on the ability of virions in human sera to bind CD81 suggested that genotype 3 virions did bind but that genotype 1 virions did not. The ability of the PKR-eIF2α phosphorylation homology domain (PePHD) in the E2 protein to predict response to IFN treatment was not validated for genotype 3 and the proposed interaction of genotype 1 E2 with PKR was not confirmed, either by co-localisation or

immunoprecipitation. Therefore the role of E2 in mediating IFN resistance remains unclear.

In conclusion, the glycoproteins of both genotypes 1 and 3 behave similarly with respect to ER localisation and complex formation in this experimental system. Glycosylation status of the glycoproteins varies between the genotypes and further investigation is required to determine the true glycosylation status of genotype 3 E1. The lack of CD81 binding of the genotype 3 E2 protein is intriguing. This will have to be confirmed with additional genotype 3 isolates but it does highlight the need to include representatives of other genotypes when analysing interactions between viral and cellular proteins.

Chapter 1

Introduction

1.1. Historical background

1.1.1. Viral hepatitis

Liver disease is often associated with yellowing of the skin and the whites of the eyes, otherwise known as jaundice. This clinical presentation of hepatitis was described in European literature dating back to the third century A.D but the infectious nature of some forms of hepatitis was recognised only in the early 1900s (Specter, 1999). The hepatitis viruses are so-called because the primary disease they cause is inflammation of the liver but taxonomically they are unrelated. What follows is a brief description of those viruses, which together with hepatitis C virus (HCV), make up the group of "hepatitis viruses".

Hepatitis A virus (HAV) is transmitted primarily by the faecal-oral route and infections are often asymptomatic, especially in young children (reviewed by Cuthbert, 2001). HAV is endemic in developing countries where most individuals are infected early in life. In first world countries, higher levels of sanitation led to a reduction in the overall rate of infection but initial exposure occurs at an older age and infection is more likely to be symptomatic. HAV is a member of the *Picornaviridae* family and has been assigned to a new genus, Hepatovirus, of which it is the only member. The virus is a spherical, nonenveloped particle with a diameter of approximately 27 nm. It contains a 7.5 kb single-stranded, positive sense RNA genome that has a single open reading frame (ORF),

flanked by 5' and 3' noncoding regions (NCRs). The genome encodes a polyprotein that is cleaved to produce the structural and nonstructural proteins.

Hepatitis B virus (HBV) is transmitted primarily via sexual and blood to blood contact. As with HAV, the majority of infections are asymptomatic but unlike HAV, a proportion of HBV-infected individuals fail to clear the virus and become chronically infected. These individuals are at risk of developing cirrhosis and possibly hepatocellular carcinoma (HCC). HBV is a member of the family *Hepadnaviridae* which includes duck HBV (DHBV) and woodchuck HBV (WHBV), both of which have proved useful as models for HBV which only infects humans and chimpanzees. HBV has a spherical, enveloped virion of ~42 nm, with the envelope consisting of host-derived lipids and the glycoprotein surface antigen (HBsAg). The encapsidated genome is a circular, partially double-stranded DNA molecule of ~3.2 kb which encodes seven known viral products (Vyas & Yen, 1999).

Hepatitis D virus (HDV) is the only member of the genus Deltavirus within the *Deltaviridae* family. It is unique amongst animal viruses in that it relies on HBV to provide helper functions to establish an *in vivo* infection. Transmission is predominantly through parenteral routes. The HDV particle is 35-47 nm and consists of the viral RNA genome and the delta antigen (HDAg) enveloped by the surface antigen of HBV. The 1.7 kb single-stranded, circular RNA genome is self-complementary and folds into a rod structure and has both self-cleavage and self-ligation activities (Smedile & Verme, 1999).

Hepatitis E virus (HEV) is transmitted by the faecal-oral route and causes a self-limiting disease. It has caused epidemics of hepatitis in adults in the developing world and has a notably high mortality in pregnant women (Aggarwal & Krawczynski, 2000). The HEV genome was first cloned in 1990 and identified as a single-stranded, positive sense RNA molecule of approximately 7.5 kb. It contains three overlapping ORFs flanked by short 5' and 3' noncoding regions. The genome is enclosed in a nucleocapsid with icosahedral symmetry, measuring 32-34 nm in diameter. On the basis of its structural and

physiochemical properties, HEV has been provisionally classified in the family *Caliciviridae*, genus *Calicivirus.*

Hepatitis G virus (HGV or GBV-C), like HCV, belongs to the family *Flaviviridae* although there is only 25% amino acid sequence homology between these viruses (Bowden, 2001). HGV was cloned from an infected tamarin, which had been inoculated with serum from a patient with chronic hepatitis. The genome was identified as a single-stranded, positive sense RNA molecule of 9.4 kb with a similar gene organisation to that of HCV. HGV is transmitted by blood to blood contact but there is no conclusive evidence that it causes liver disease and the true site of replication has yet to be identified. It is increasingly being referred to by its original name, GBV-C.

1.1.2. Non-A, non-B hepatitis

The development of serological tests for HBV led to the introduction of blood donor screening for hepatitis B infection in the 1970s. This allowed for the evaluation of the role of this virus in transfusion-associated hepatitis (TAH). However, the removal of all HBV-positive blood donations from the supply did not eliminate cases of TAH. It became apparent that another blood-borne agent was responsible and these cases were defined as having non-A, non-B hepatitis (NANBH). Initially this was thought to be a mild disease but when long-term cases were analysed it was evident that chronic hepatitis often ensued with further progression to cirrhosis and liver failure.

Evidence of the transmissible nature of the NANB agent was shown when serum from patients with NANBH was administered intravenously to chimpanzees (Alter *et al.*, 1978). The chimpanzees developed hepatitis but importantly, were negative for HAV, HBV or other hepatotropic viruses such as cytomegalovirus (CMV) and Epstein Barr virus (EBV). Ultrastructural changes, not reported in humans, were observed in the hepatocytes of the infected chimpanzees. The appearance of either cytoplasmic membranous tubules or intracellular particles of 20-27 nm or both were reported. Feinstone *et al.* (1983)

demonstrated that the infectious sera were rendered non-infectious after treatment with chloroform, which suggested that the agent might be a virus with a lipid envelope.

1.1.3. Cloning of the NANB hepatitis infectious agent

Attempts at identifying the NANBH agent by conventional virological and immunological approaches failed, possibly due to insufficient concentrations of both viral antigen and viral antibody. The advent of recombinant DNA cloning technologies provided a new approach to identifying the nucleic acid of the agent. Choo *et al.* (1989) described the isolation of NANBH-specific nucleic acid by screening a bacterial cDNA library derived from chimpanzee plasma of high infectious titre. The plasma underwent ultracentrifugation to ensure isolation of a virus and both DNA and RNA were extracted from the resulting pellet. cDNA was synthesised using random primers and either DNA polymerase or reverse transcriptase and was cloned into the bacteriophage λgt11 and expressed in *Escherichia coli* (*E.coli*). The library was screened for clones expressing viral antigen, using serum from a patient with chronic NANBH as a presumed source of antibodies. After screening approximately 10^6 recombinants, two positive overlapping clones were identified. The cDNA of these clones failed to hybridise to human and chimpanzee DNA therefore they were not derived from the host genome. However, total RNA from infectious chimpanzee liver did hybridise to the clones and no hybridisation was observed with RNA from an uninfected liver or after ribonuclease treatment of total nucleic acid. The authors concluded that the clones were derived from an exogenous RNA molecule associated with NANB hepatitis infection. Further analysis showed that this was a positive sense single-stranded RNA molecule, approximately 10 kb in length containing one continuous translational open reading frame (ORF). The binding of oligo(dT)-cellulose indicated the presence of either a 3' terminal polyadenylated sequence or an A-rich tract elsewhere in the molecule. The NANB agent was named hepatitis C virus (HCV). Thereafter serological assays were developed (Kuo *et al.*, 1989) and showed that HCV represented the predominant cause of transfusion associated hepatitis.

1.1.4. Virus morphology

Before the identification of HCV, filtration experiments had shown that passage through an 80 nm filter did not have an effect of the infectious nature of the inoculum (Bradley *et al.*, 1985). Further filtration experiments narrowed down the predicted diameter of the HCV virion to 30-60 nm (He *et al.*, 1987). Subsequent to the identification of HCV, studies on the detailed structure of the virion have been hampered by the lack of a cell culture system and the low titres of infectious virus present in serum. More recently, electron microscopy, using immunogold staining of the putative envelope proteins, has identified spherical virus-like particles in human plasma. These particles were 55-65 nm in diameter and had spike-like projections of approximately 6 nm (Kaito *et al.*, 1994). Similarly, Shimizu *et al.* (1996) described the appearance of 50 nm virus-like particles in the liver of an infected chimpanzee. The buoyant density of HCV virions is heterogeneous, possibly due to association with immunoglobulins or low-density lipoproteins. In highly infectious sera, HCV was detected in fractions of low buoyant density (≤1.06 g/ml), equivalent to that of low density lipoproteins (LDL) (Hijikata *et al.*, 1993a). In contrast, virus complexed with antibody was detected in higher density fractions (~1.17 g/ml) and correlated with lower infectivity (Hijikata *et al.*, 1993a). Chloroform treatment, which removes the lipid envelope, also resulted in particles with a buoyant density of 1.17 g/ml, thought to represent the HCV nucleocapsids (Hijikata *et al.*, 1993a).

1.1.5. Classification of HCV

The morphological characteristics of the HCV virion and the identification of a single-stranded, positive sense RNA genome containing one ORF led to the suggestion that HCV was a member of the *Flaviviridae* family. This was confirmed following cloning and sequencing of the full-length genome, which showed regions of sequence identity with the pestiviruses and to a lesser extent with the flaviviruses (Choo *et al.*, 1991). This also suggested a similar pattern of gene organisation with the structural genes in the N-terminal

portion and the nonstructural genes in the remainder (Figure 1.1). Sequence comparison with other members of the *Flaviviridae* family also identified motifs that are characteristic of a serine protease, RNA helicase and RNA-dependent RNA polymerase (Miller & Purcell, 1990, Choo *et al.*, 1991). Despite these regions of sequence homology, there is little overall amino acid sequence identity between HCV and the pestiviruses and therefore a third genus of the *Flaviviridae* family was proposed for classification of HCV (Robertson *et al.*, 1998). This genus was named the Hepaciviruses and currently HCV is the only member.

1.2. The HCV genome and its replication

1.2.1. 5' Untranslated region (5' UTR)

The 5' UTR is the most conserved region of the HCV genome (Bukh *et al.*, 1992). This 341 nt region is much longer than the 5' UTRs of flaviviruses, which have an average length of 100 nt (Heinz, 1992) but is similar to that of pestiviruses such as bovine viral diarrhoea virus (BVDV) and classical swine fever virus (CSFV) (Brown *et al.*, 1992). The 5' UTRs of these viruses all contain several AUG triplets upstream of the initiation codon and they are all predicted to form extensive secondary structure. The HCV 5' UTR has four major structural domains, designated I to IV (Figure 1.2). Domains II and III are complex stem-loop structures, the latter containing an RNA pseudoknot. Domain IV contains the AUG initiation codon. This high degree of secondary structure and the multiple upstream AUG sequences is similar to the equivalent region in picornaviruses. Picornaviral genomes act as uncapped messenger RNAs and are translated by a mechanism which allows ribosomes to bind internally on the genome and initiate translation at a specific AUG sequence (Jackson *et al.*, 1990). These internal ribosomal entry sites (IRES) are dependent on the highly ordered RNA secondary structure and allow translation in a cap-independent manner. An IRES element was identified in the HCV 5' UTR by Tsukiyama-Kohara *et al.* (1992). Deletion analysis has mapped the

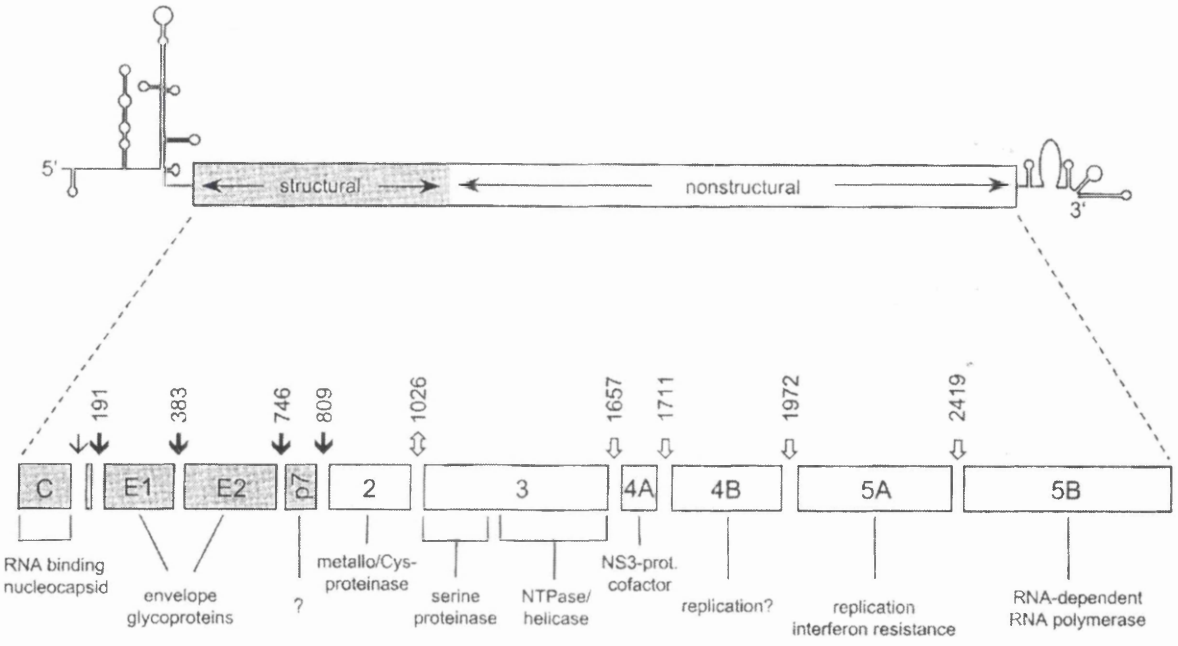


Figure 1.1. A schematic representation of the HCV genome and the individual protein products of the polyprotein. Cleavage sites for (⬇) host cell signalases, (⇳) the NS2-3 proteinase, (⇓) the NS3/4A proteinase complex and (↓) an unknown cellular enzyme are marked. Numbers above the arrows refer to the P1 positions of the corresponding cleavage sites (numbering according to genotype 1). The functions of the proteins are indicated, where known. (Taken from Bartenschlager & Lohmann, 2001).

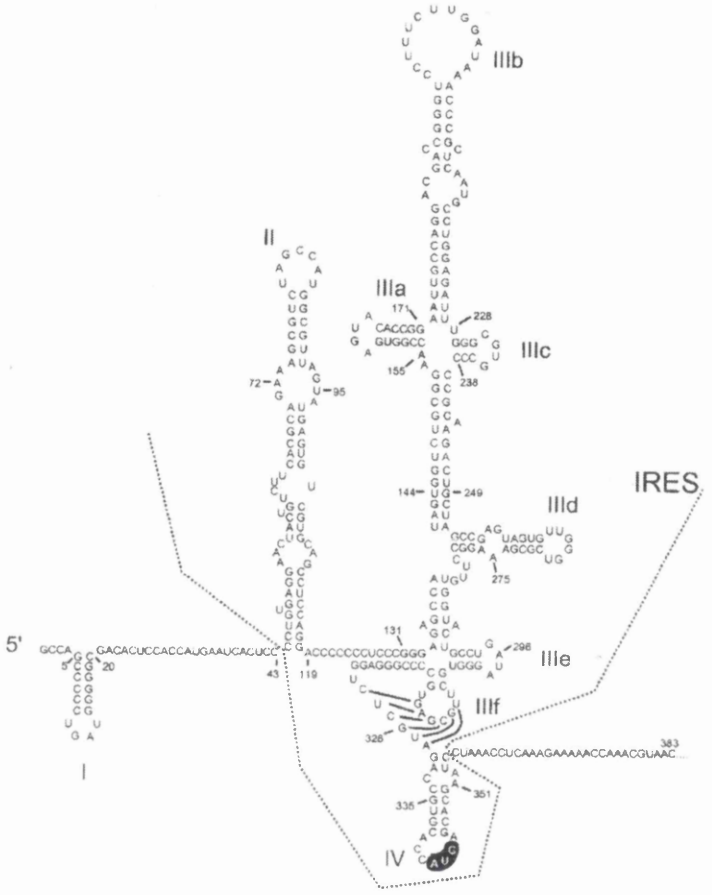


Figure 1.2. Secondary RNA structure of the HCV 5'UTR and the start of the downstream open reading frame. The region comprising the IRES extends from domain II to domain IV. The AUG start codon is highlighted. (Taken from Honda *et al.*, 1999).

HCV IRES activity to between nucleotides 40 and 370 (Wang *et al.*, 1993, Reynolds *et al.*, 1995), therefore the 3' boundary of the IRES lies in the coding sequence of the core protein, a feature unique to HCV. It has since been suggested that core may modulate the initiation of translation and Wang *et al.* (2000) have shown that the full-length core sequence suppresses translation and that this is a function of the core coding sequence not the core protein. The IRES has also been shown to bind cellular proteins such as PTB (Ali & Siddiqui, 1995) and La autoantigen (Ali & Siddiqui, 1997).

While the IRESs of HCV and the pestiviruses are structurally similar, their sequence differs by approximately 50%. However, most of these differences are concentrated in the stem-structures and base-pairing is maintained (Hellen & Pestova, 1999). There is strict conservation of sequence in the terminal loop (especially in domain II) and the internal bulges (Honda *et al.*, 1999, Hellen & Pestova, 1999), which suggests that these regions may be involved in a direct interaction with conserved elements of the host translational apparatus. Pestova *et al.* (1998) have demonstrated specific interactions between these IRES structures and both the 40S ribosomal subunit and the eukaryotic initiation factor 3 (eIF3). eIF3 is essential for formation of the 80S ribosome and subsequent polypeptide synthesis. Two subunits of eIF2 (eIF2Bγ and eIF2γ) have been shown to be cofactors of HCV IRES-mediated translation (Kruger *et al.*, 2000) but in contrast to other IRES elements, additional eIFs, such as eIF4A, eIF4B or eIF4F, are not required (Pestova *et al.*, 1998). Therefore the requirements for translation initiation in HCV are relatively simple and are so far unique in eukaryotes, being more similar to the situation in prokaryotes, with the IRES being comparable to the Shine-Dalgarno sequence. The HCV IRES functions by recruiting translation components to form an initiation complex and the highly ordered secondary structure orientates this complex such that the mRNA-binding cleft of the 40S subunit is positioned precisely at the initiation codon.

1.2.2. 3' UTR

The 3' UTR of HCV has 3 distinct regions – a short variable sequence, a poly(U) tract of variable length and a highly conserved sequence of 98 nt termed the 3'X region (Figure

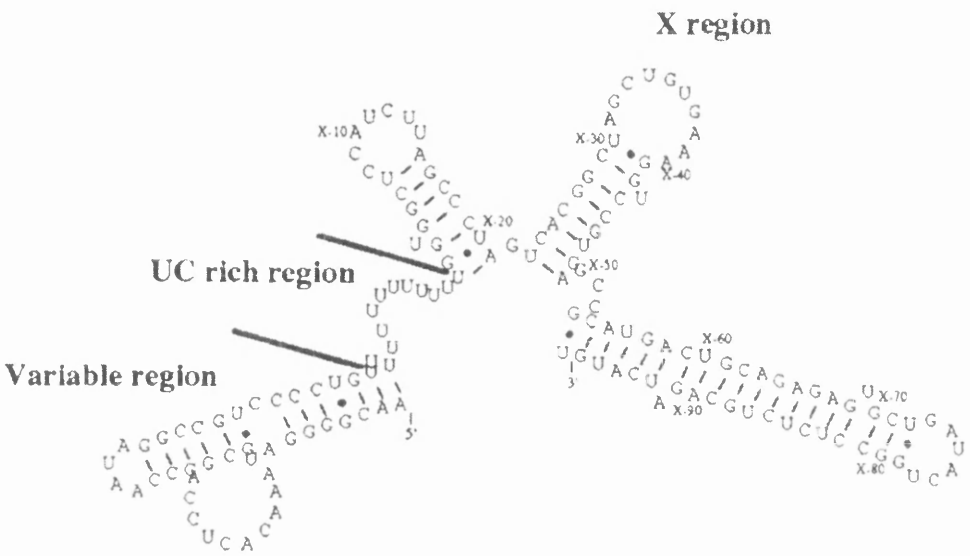


Figure 1.3. Predicted model of the secondary structure of the HCV 3'UTR RNA. (Taken from Ito & Lai, 1997).

1.3). The 3'UTR is predicted to fold into a stem-loop structure. The poly (U) region and the 3'X region but not the variable region have been shown to be critical for *in vivo* infectivity of HCV in chimpanzees (Yanagi *et al.*, 1999b, Kolykhalov *et al.*, 2000). Little is known about HCV replication (see section 1.2.9) but by analogy with other single stranded positive sense RNA viruses, transcriptional initiation by the RNA polymerase (see section 1.2.8) is predicted to occur at the 3' ends of both negative and positive RNA strands. The poliovirus helicase has been shown to bind the 3' terminal sequences of the negative RNA strand, possibly because it is required for unwinding the double-stranded structure formed after initiation of copying of the positive strand (Banerjee *et al.*, 1997). In contrast, the HCV helicase, NS3 (see section 1.2.5) has been shown to interact with the 3' UTR of both positive and negative RNA strands (Banjeree & Dasgupta, 2001). Binding to the negative strand was specific for the stem loop structure at the extreme 3' terminus, whereas binding to the positive strand required the entire 3' UTR. Several cellular proteins have also been shown to interact with the 3' UTR. Polypyrimidine tract-binding protein (PTB) binds to both the poly(U) tract and the 3'X region (Tsuchihara *et al.*, 1997, Luo, 1999). Ito *et al.* (1998) have demonstrated that the 3'X region is able to enhance translation from the IRES and that this enhancement is reduced (but not abolished) when the PTB binding site is mutated. As mentioned in section 1.2.1, the 5' UTR also binds PTB, suggesting a possible interaction of the 5' and 3' UTRs mediated by PTB. Other proteins that bind the 3'UTR include glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Petrik *et al.*, 1999), La autoantigen (Spangberg *et al.*, 1999), heterogeneous nuclear ribonucleoprotein C (hnRNP C) (Gontarek *et al.*, 1999) and human ribosomal proteins, L22, L3, S3 and mL3 (Wood *et al.*, 2001). The functional relevance of these interactions remains unknown.

1.2.3. Processing of the HCV polyprotein

The HCV genome encodes a polyprotein that must be proteolytically cleaved to produce the 10 viral proteins (Figure 1.1). The structural proteins (core, E1, E2 and p7) are in the N-terminal region of the polyprotein. Processing of the structural proteins appears to be catalysed by host signal peptidases in the endoplasmic reticulum (ER) lumen, as cleavage

in a cell-free assay is dependent on the addition of microsomal membranes (Hijikata *et al.*, 1991a). The cleavage of core and E1 from the polyprotein probably occurs cotranslationally, however, E2 is generated from an E2-p7-NS2 precursor. This is subsequently cleaved into NS2 and E2-p7. Final processing of E2 is inefficient and two stable proteins are produced, E2 and E2-p7 (Lin *et al.*, 1994). Cleavage in the non-structural protein region is mediated by two viral protease activities located in the NS2-NS3 region. The NS2/NS3 junction is cleaved in *cis* by a zinc-dependent metalloprotease that encompasses NS2 and the N-terminal portion of NS3 (Hijikata *et al.*, 1993b). The remaining non-structural proteins are cleaved by the action of the serine protease domain of the NS3 protein. Cleavage of NS3/NS4A junction occurs in *cis* and this is followed by *trans* cleavage of the NS5A/5B, 4A/4B and 4B/5A junctions (Bartenschlager *et al.*, 1995). A brief description of the non-structural proteins and their functions follows in this section and the structural proteins will be discussed in detail in section 1.4.

1.2.4. NS2

The mature NS2 is a transmembrane protein with its C-terminal region in the ER lumen and its N-terminal region in the cytoplasm (Santolini *et al.*, 1995). As mentioned in section 1.2.3, together with the N-terminal third of NS3, NS2 has proteolytic activity that is responsible for *cis* cleavage at the NS2/NS3 junction. Although this portion of NS3 contains the serine protease domain, it is not required for the 2/3 cleavage (Hijikata *et al.*, 1993b). The NS2-3 protease has been proposed to be a metalloprotease based on the observation that activity is inhibited by metal chelators such as EDTA and stimulated by exogenous zinc (Hijikata *et al.*, 1993b), however there is no direct structural evidence to support this hypothesis. However, elucidation of the crystal structure of the NS3 serine protease identified a zinc binding site that is co-ordinated by 3 cysteine residues (Love *et al.*, 1996), and the role of zinc ion appeared to be entirely structural. Mutagenesis of these cysteine residues significantly decreased serine protease activity and also had a measurable affect on the NS2-3 protease activity (Hijikata *et al.*, 1993b), so it is possible that the two proteases share the same zinc binding site. Site-directed mutagenesis also identified His-952 and Cys-993 in NS2 as being important for enzymatic activity (Hijikata

et al., 1993b), which led to the suggestion that the NS2-3 enzyme is a cysteine protease. Mutagenesis of the residues surrounding the cleavage site revealed that cleavage is inhibited only if the mutations alter the conformation of the region, suggesting that correct folding of the cleavage site and the protease is more important for activity than any particular amino acid side chain (Reed *et al.*, 1995). At this time, the true nature of the protease remains unclear. NS2 has also been shown to be involved in regulation of NS5A phosphorylation (Liu *et al.*, 1999).

1.2.5. NS3

The NS3 protein extends from amino acids 1027 to 1658 and has a molecular weight of approximately 70 kDa (Grakoui *et al.*, 1993a). It is a multifunctional enzyme. In addition to its role as part of the NS2-3 protease described above, NS3 also has a serine-type protease function in the N-terminal domain and nucleoside triphosphate (NTPase)/helicase activity in the remaining portion of the molecule (Hijikata *et al.*, 1993b, Suzich *et al.*, 1993, Kim *et al.*, 1995). These enzymatic activities are also functions of the flavivirus NS3 protein and the pestivirus p125 protein, which therefore may be functional homologues of HCV NS3 (Ryan *et al.*, 1998). NS3 may also play a role in the modulation of protein kinase A-mediated signal transduction (Borowski *et al.*, 1996) and the N-terminal half of NS3 has been reported to transform NIH 3T3 fibroblasts (Sakamuro *et al.*, 1995).

1.2.5.1. NS3 protease

The NS3 serine protease is responsible for *cis* cleavage at the NS3-4A junction and *trans* cleavage at the 4A-4B, 4B-5A and 5A-5B junctions. The minimal NS3 domain required for protease activity has been mapped to the N-terminal 180 residues (Bartenschlager *et al.*, 1994) and a catalytic triad has been identified within this region based on sequence comparison with proteases of the trypsin superfamily. The residues involved are His-1083, Asp-1107 and Ser-1165 and mutation of these was shown to abolish cleavage at the NS3-4A, NS4A-4B, NS4B-5A and NS5A-5B junctions but not at the NS2-3 junction

(Bartenschlager *et al.*, 1993, Grakoui *et al.*, 1993b). Processing at the NS3-4A junction is rapid and probably cotranslational. Thereafter, the order of cleavage is NS5A-5B followed by NS4A-4B and NS4B-5A (Bartenschlager *et al.*, 1994).

The NS4A protein functions as an essential co-factor for the NS3 serine protease and is required for cleavage at the NS3-4A, NS4A-4B and NS4B-5A sites and enhances cleavage at the NS5A-5B site (Failla *et al.*, 1994). NS4A forms a detergent-stable, non-covalent complex with the NS3 protease domain and this interaction requires the 22 N-terminal residues of NS3 and a 12-residue sequence in the centre of NS4A (Lin *et al.*, 1995, Satoh *et al.*, 1995). NS4A was thought to enhance protease activity by inducing conformational changes, which stabilise NS3 (Steinkuhler *et al.*, 1996). The resolution of the X-ray crystal structures of NS3 revealed that, in the absence of NS4A, the N-terminal region of NS3 is loosely structured and extends away from the protein (Love *et al.*, 1996). In the presence of NS4A, the N-terminal domain forms an 8-stranded β-barrel, with one strand contributed by NS4A (Kim *et al.*, 1996). It was also shown that NS4 binding affects the spatial arrangement of the catalytic triad, as in the absence of NS4A, the His-1083 and Asp-1107 residues were positioned too far away from each other and the Ser-1165 residue to form a typical catalytic triad (Love *et al.*, 1996). This explains the observed increase in protease activity in the presence of NS4A.

As mentioned in section 1.2.4., analysis of the crystal structure of the NS3 protease indicated the presence of a zinc binding site co-ordinated by three cysteine residues and (via a water molecule) a histidine residue (Love *et al.*, 1996). Due to the close proximity of the zinc ion to the catalytic serine residue, it is predicted to have a structural rather than a catalytic role. This is supported by the report that zinc can be substituted by cadmium or cobalt ions with no affect on the protease activity (Stempniak *et al.*, 1997). Also, in other enzymes where zinc is known to have a structural role, it is ligated by cysteine thiolates, as seems to be the case with NS3 (Berg & Shi, 1996).

1.2.5.2. NS3 NTPase/helicase

Sequence motifs within the C-terminal two thirds of NS3 suggest that this region of the protein has both NTPase and RNA helicase activity. Suzich *et al.* (1993) demonstrated that HCV NS3 does possess NTPase activity *in vitro* and it has the same requirements for magnesium concentration and pH as do the equivalent flavivirus and pestivirus enzymes. NTPase activity is enhanced by the presence of polynucleotides, particularly poly(U). The NS3 protein also contains a motif (DECH) which suggests it is a member of the DEXH subfamily of proteins. These proteins belong to the DEAD box family of RNA helicases and Kim *et al.* (1995) demonstrated that the carboxy terminal fragment of NS3 possesses RNA helicase activity. This activity requires ATP and either magnesium or manganese ions. The helicase domain is able to unwind dsRNA, RNA/DNA heteroduplexes and dsDNA in a 3' to 5' direction, but this is dependent on a 3' overhang region (Tai *et al.*, 1996). The minimal requirement for both NTPase and helicase activities has been mapped to the C-terminal 465 amino acids of NS3 (Kim *et al.*, 1997a) and mutations in either the ATPase or the helicase motifs affect both functions significantly (Kim *et al.*, 1997b). The helicase activity has also been shown to be required for viral replication in chimpanzees (Kolykhalov *et al.*, 2000).

The N-terminal protease domain and the C-terminal NTPase/helicase domain have been analysed independently and shown to be sufficient for their respective activities. There is no evidence of cleavage *in vivo* to separate these two domains and in other flavivirus proteins, the protease and helicase domains are also found on the same polypeptide. To investigate any potential cross-talk between these two domains, the protease, NTPase and helicase activities have also been analysed in terms of the full-length NS3 (Gallinari *et al.*, 1998). Once complexed with NS4A, the isolated domains and the full-length protein did not shown any significant differences in any of the three enzymatic activities. It was also shown that complete inhibition of the protease activity had no affect on the helicase activity. However, it has been shown that the protease domain is also involved in RNA-binding and that this leads to inhibition of protease activity, while ensuring optimal

binding of RNA to the helicase domain (Gallinari *et al.*, 1998). By this mechanism, there is potentially interplay between these two enzymatic domains.

1.2.6. NS4A and NS4B

Aside from its role as a cofactor for the NS3 serine protease (as discussed in section 1.2.5.1.), not much else is known about the 8 kDa NS4A protein. NS4A has an N-terminal hydrophobic domain, which has been shown to mediate localisation of NS3 to the ER membrane (Wolk *et al.*, 2000). NS4A has also been reported to form a stable complex with the NS4B-5A polyprotein, which may be required for NS3-mediated cleavage at the NS4B-5A junction (Lin *et al.*, 1997). In addition, NS4A has been identified as a modulator of NS5A hyperphosphorylation and this is mediated by the central region of NS4A, which is also involved in the interaction with NS3 (Koch & Bartenschlager, 1999). A recent study has also identified an interaction between NS4A and NS2, although the functional role of this complex is unknown (Flajolet *et al.*, 2000).

NS4B is the least well characterised of the HCV proteins. This hydrophobic 27 kDa protein localises to the ER membrane and displays properties of a cytoplasmically orientated integral membrane protein (Hugle *et al.*, 2001). No function as yet been ascribed to the NS4B proteins of related pestiviruses, however the Kunjin flavivirus NS4B protein has been shown to be essential for viral replication (Khromykh *et al.*, 2000). If this were true for HCV NS4B, it would support the suggestion that NS4B associates with the other non-structural proteins to form a replication complex. This was based on the observation that NS4A interacts with the NS4B-5A cleavage substrate (Lin *et al.*, 1997) and that, like NS3 and NS4A, NS4B is also involved in the modulation of NS5A hyperphosphorylation (Koch & Bartenschlager, 1999).

1.2.7. NS5A

The mature NS5A protein is a 56 kDa (p56) phosphoprotein but it also exists in a hyperphosphorylated form with a molecular weight of 58 kDa (p58) (Kaneko *et al.*, 1994). Phosphorylation occurs mainly at serine residues and is mediated by a cellular kinase (Reed *et al.*, 1997). Basal phosphorylation to produce p56 is independent of NS4A, however, hyperphosphorylation is dependent on the association of NS5A with NS4A (Asabe *et al.*, 1997). Furthermore, Koch & Bartenschlager (1999) demonstrated that hyperphosphorylation occurs only when NS5A is expressed as part of a continuous NS3-NS5A polyprotein and that NS3, NS4A and NS4B are all involved in modulation of NS5A hyperphosphorylation. The NS2 protein also plays a role in the regulation of NS5A phosphorylation (Liu *et al.*, 1999). Hyperphosphorylation of NS5A is a conserved feature among members of the *Flaviviridae* family but the role of the two different forms of the protein is unknown.

NS5A has been shown to form a multisubunit complex with NS3, NS4A and NS4B (Lin *et al.*, 1997) and NS5A has been co-precipitated with the remaining two nonstructural proteins, NS2 and NS5B (Hijikata *et al.*, 1993c). This suggests that the HCV nonstructural proteins form a replication complex as described for other single-stranded, positive sense RNA viruses (e.g. poliovirus). This complex is likely to be anchored on the intracellular membranes and NS5A is localised in the cytoplasmic membrane surrounding the nucleus (Tanji *et al.*, 1995). A potential nuclear localisation signal has been identified in the C-terminus of NS5A but there have been no reports of NS5A in the nucleus (Ide *et al.*, 1996). Apart from its putative role in virus replication, a region of NS5A (aa 2135-2331) has been shown to have transcription-activating activity (Kato *et al.*, 1997). Two acidic motifs within this region are essential for this activity. It is not known whether NS5A acts as a transcriptional activator *in vivo* but if so, it could have a role in the development of hepatocellular carcinoma (see section 1.7.2.2.) in HCV-infected patients.

NS5A has also been implicated as a mediator of viral resistance to interferon alpha (IFN) therapy. NS5A has been shown to bind to double-stranded RNA-dependent protein kinase (PKR) (Gale *et al.*, 1997), which is induced by IFN and is a key component of the antiviral response. A 40 amino acid stretch, called the interferon sensitivity-determining region (ISDR), was identified in the central region of NS5A as being predictive of outcome to IFN therapy (Enomoto *et al.*, 1995). However, this finding led to several conflicting reports and the proposed role of NS5A in IFN-resistance will be discussed in more detail in section 1.8.4.1.

1.2.8. NS5B

The 65 kDa NS5B protein has been shown to be an RNA-dependent RNA polymerase (RdRp) (Behrens *et al.*, 1996). It contains the classical GDD motif, which is conserved in all RNA polymerases, and is essential for polymerase activity. The crystal structure of NS5B has been determined and reveals that the catalytic domain has the typical right-handed "fingers-palm-thumb" structure, with the RNA-binding groove lying between the "fingers" and "thumb" (Bressanelli *et al.*, 1999). Further proof that the non-structural proteins form a replication complex is provided by the report that NS5B interacts with both NS3 and NS4A and this complex is localised to the endoplasmic reticulum (Ishido *et al.*, 1998). NS5B expressed alone is also associated with intracellular membranes (Hwang *et al.*, 1997), so it is likely that HCV replicates on intracellular membranes. A recombinant full-length form of NS5B has been shown to be capable of copying the full-length HCV RNA genome *in vitro*, without the need for addition factors (Oh *et al.*, 1999), although additional viral or cellular factors are probably necessary for regulation of RNA synthesis. The HCV RdRp was shown to use the 3'X tail of both the plus and minus RNA strands as templates for RNA synthesis (Oh *et al.*, 1999). This specificity for the 3'X tail is thought to be due to the recognition of specific stem-loop structures in the 3'UTR (Cheng *et al.*, 1999). This is in accordance with the reports that both NS5B and the 3'X tail are essential for viral replication *in vivo* (Kolykhalov *et al.*, 2000).

1.2.9. Replication of HCV

Since the identification of HCV in 1989, the characterisation of the viral gene products has been relatively rapid. However, research into the viral life cycle has been severely hampered by the lack of suitable *in vivo* and *in vitro* systems for the study of HCV infection, replication and pathogenesis. As mentioned above, coprecipitation experiments have indicated that the nonstructural proteins probably form a replicase complex on the cytoplasmic side of the ER membrane (Ishido *et al.*, 1998, Lin *et al.*, 1997), which is analogous to what happens in related flaviviruses and pestiviruses. In its role as the RdRp, the NS5B protein is responsible for the synthesis of both plus- and minus-strand RNA, although the individual steps involved in this process are unknown. Numerous attempts have been made at propagating HCV in cell culture, through infection of either primary cells or established cell lines, or cultivation of primary cells from chronically infected patients (Lanford *et al.*, 1994, Ito *et al.*, 1996, Rumin *et al.*, 1999, reviewed by Bartenschalger & Lohmann, 2001). These systems suffer from both poor efficiency and poor reproducibility and none permitted detailed analysis of the replication cycle. Recent progress has been made in this field with the development of subgenomic HCV RNA replicons which are capable of replicating successfully in the human hepatoma cell line (Huh7) (Lohmann *et al.*, 1999). This system will be covered in more detail in section 1.2.9.2.

The search for a suitable animal model in which to study HCV replication has not been any more successful. The chimpanzee is the only known animal to support HCV replication but its use is limited by its scarcity, high maintenance costs and for ethical reasons. In the absence of a successful cell culture system, the chimpanzee has proved invaluable for the study of molecular clones of HCV, which contain all the necessary requirements for infectivity. The development of such clones will be discussed in the following section (1.2.9.1). A recent mouse model for HCV infection has been described, which relies on the transplantation of human hepatocytes into immunodeficient transgenic mice whose own hepatocytes degenerate spontaneously (Mercer *et al.*, 2001). The resultant chimeric mouse/human liver could be infected with serum from an HCV-infected

patient. The chimeric livers were able to maintain relatively high levels of HCV RNA (3 x 10^4 to 3 x 10^6 copies/ml) for a period of 15-17 weeks. Although promising, this mouse model requires specialised techniques, which are unlikely to be available to most HCV researchers.

1.2.9.1. Infectious clones

In order to investigate the role of any specific viral product or domain in HCV replication, it would be useful to initiate the infection with RNA transcribed from a cDNA clone. If patient material is inoculated, the presence of a quasispecies (see section1.3.3.) of individual variants means that there is uncertainty over exactly which variants result in expression of viral proteins. RNA transcribed from a cDNA clone is advantageous as the infecting viral genome is then well defined and can be easily manipulated. Initial attempts at producing a clone that was infectious when RNA was injected directly into the liver of a chimpanzee were unsuccessful (Major & Feinstone, 2000). These clones were constructed by RT-PCR amplification of HCV RNA from patient material and PCR errors could have potentially rendered the virus incapable of replication. Additionally, the quasispecies nature of HCV could have meant that the selected clones contained natural mutations that were lethal for replication. To address this, a consensus clone was constructed from a total of six clones and intrahepatic injection of RNA derived from this clone resulted in HCV RNA levels of approximately 1 x 10^6 copies/ml (Kolykhalov *et al.*, 1997). In addition, serum aminotransferase levels, liver histology and the appearance of HCV-specific antibodies indicated that the chimpanzee was infected. Using the same strategy, infectious clones of genotypes 1a, 1b and 2a (see section 1.3.1.) have been reported (Yanagi *et al.*, 1997, Yanagi *et al.*, 1998, Yanagi *et al.*, 1999a, Beard *et al.*, 1999, Lanford *et al.*, 2001). Mutagenesis of these infectious molecular clones has demonstrated that NS2, NS3, NS4A and NS5B activities are all required for replication *in vivo*, as is the 3' X tail and poly(U) region of the 3'UTR (Kolykhalov *et al.*, 2000, Yanagi *et al.*, 1999b). The hypervariable region (HVR1) of E2 (see section 1.3.2) is not required for replication *in vivo* (Forns *et al.*, 2000). These studies have provided crucial information about the

essential viral genes required for HCV replication but the expense of such studies means that they are not practical for detailed analysis of the viral genome.

1.2.9.2. Subgenomic replicon system

The most recent breakthrough in the field of HCV replication has come from the development of selectable subgenomic replicons which can replicate to high levels after transfection into the human hepatoma cell line Huh7 (Lohmann *et al.*, 1999) (Figure 1.4). Similarly to the construction of the infectious clones, a consensus genome (genotype 1b) was constructed from material cloned from liver tissue of a chronically infected patient. The structural genes (and NS2) were replaced with the gene encoding the selectable marker neomycin phosphotransferase (neo) which confers resistance to the antibiotic G418. Translation of this gene was directed by the HCV IRES and an encephalomyocarditis virus (EMCV) IRES directed translation of the HCV NS genes. These bicistronic RNA transcripts were transfected into Huh7 cells and those cells harbouring the replicon were selected with G418. Low numbers of resistant colonies were obtained but Northern blot analysis demonstrated that such colonies contained large amounts of HCV RNA. The HCV proteins were detected by immunoprecipitation and immunostaining showed them to be closely associated with the ER, consistent with this being the site of replication.

This replicon system was validated by Blight *et al.* (2000) and their study also identified a series of adaptive mutations which arose during propagation of the cells containing the replicon. These mutations enhanced RNA replication, which increased the G418 transduction efficiency. The adaptive mutations were clustered in a 30 amino acid region in the centre of NS5A and when they were transferred into the original replicon, there was increased cloning efficiency in Huh7 cells (Blight *et al.*, 2000). Interestingly, only two of these mutations have been identified independently (Bartenschlager & Lohmann, 2001). Similarly, Krieger *et al.* (2001) reported the identification of a cell culture adaptive replicon that contained two mutations in NS3 and one in NS5A. These adaptive mutations were confirmed by replacement of the neo gene with the firefly luciferase gene

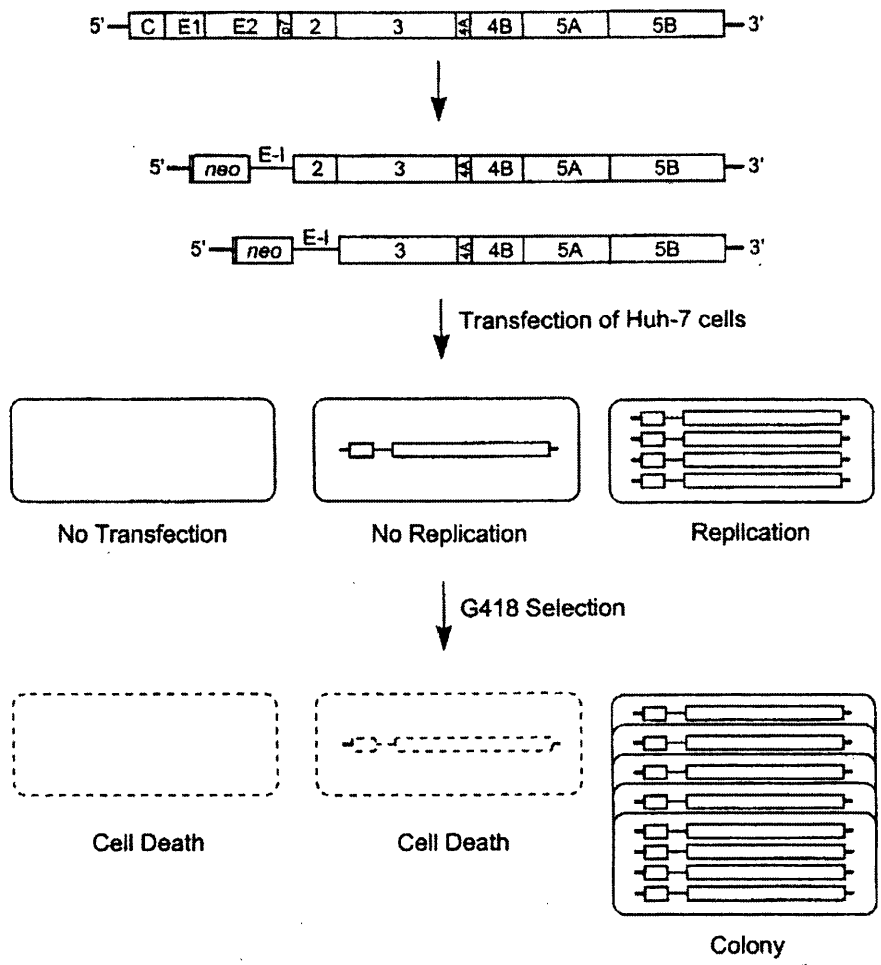


Figure 1.4. Schematic representation of the method used to establish HCV-replicon-containing cell lines. The HCV genome is shown at the top and the subgenomic RNAs are shown below. They are composed of the HCV 5'UTR plus a small fragment of the core-coding region, the *neo* gene, the encephalomyocarditis virus IRES (E-I), NS2-5B or NS3-5B and the 3'UTR. Upon transfection of Huh-7 cells, only those supporting replication of the HCV RNAs amplify the *neo* gene and develop resistance against the drug G418. Therefore only these cells will form colonies. (Taken from Bartenschlager & Lohmann, 2000).

in the replicon, which on transfection into Huh7 cells, showed increased levels of luciferase activity compared to the original replicon sequence (Krieger *et al.*, 2001). It is not known if similar mutations would be seen *in vivo* or whether they are simply due to adaptation to the Huh7 cell line. So far this cell line is the only one shown to support the HCV replicon system, possibly due to factors unique to these cells. Another enigma is that the H77 clone, known to be infectious in chimpanzees, could not be established as a subgenomic replicon (Blight *et al.*, 2000) and conversely, the isolate used to establish the replicon system has not been shown to cause infection in the chimpanzee. Although it is apparent that the structural proteins are not required for RNA replication, a full-length HCV replicon is required for analysis of the complete viral replication cycle. A recent report at an international meeting described the persistent replication of a selectable full length HCV genome in cell culture but there was no evidence of virion production, despite the presence of structural proteins (Pietschmann *et al.*, 2001).

1.3. Genetic variability of the HCV genome

1.3.1. Genotypes

Soon after reporting of the prototype HCV isolate, HCV-1 (Choo *et al.*, 1991), a number of sequences of other isolates were published and comparison revealed considerable diversity. Phylogenetic analysis of nucleotide sequences of complete genomes or subgenomic regions has led to the classification of six HCV genotypes (numbered 1-6) (Simmonds *et al.*, 1994) (Figure 1.5). Within each genotype are more closely related subtypes (a, b, c etc). Over the complete virus genome, genotypes differ from each other by more than 30%, subtypes by more than 20% and individual isolates of the same subtype, by less than 10% (Simmonds, 1995).

There are substantial regional differences in the distribution of HCV genotypes. Genotypes 1, 2 and 3 have a broad world-wide distribution, although their relative

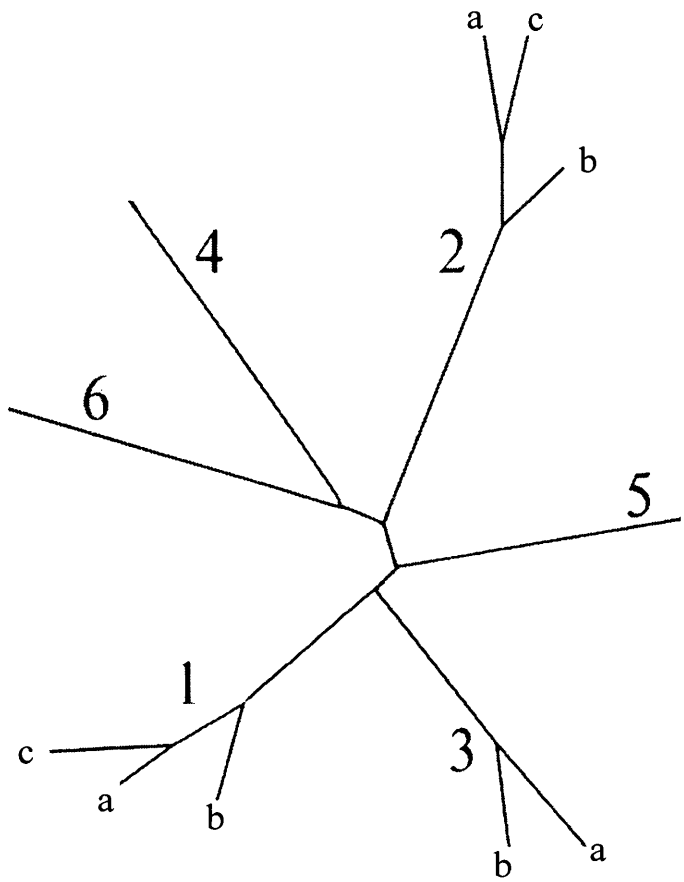


Figure 1.5. Phylogenetic analysis of nucleotide sequences in part of the HCV NS5A region. They are classified into six genotypes and subtypes a, b, etc. (Taken from Simmonds, 1995).

prevalence varies from one geographic area to another. In the United States and Western Europe, genotypes 1a and 1b are predominant (McOmish *et al.*, 1994). In Southern and Eastern Europe there is more frequent infection with genotype 1b (Simmonds, 1995) and in Japan, genotype 1b accounts for up to 73% of infections (Takada *et al.*, 1993). Genotypes 2a and 2b are commonly found throughout the United States, Europe and Japan. Genotype 3a is prevalent amongst intravenous (IV) drug users in the USA and Europe (Pawlotsky *et al.*, 1995) and is the most common genotype in Thailand (Kanistanon *et al.*, 1997). A wide variety of genotype 3 subtypes are found in Southern Asian countries (Tokita *et al.*, 1994, Simmonds, 1995). Genotype 4 is prevalent in Middle Eastern countries and in Central Africa (Simmonds, 1995, Xu *et al.*, 1994). Genotype 5 is confined to South Africa (Ohno *et al.*, 1994) and genotype 6 is restricted to Southeast Asia (Simmonds *et al.*, 1993).

The rate of nucleotide sequence change of HCV has been estimated as 10^{-3}-10^{-4} substitutions per site per year (Ogata *et al.*, 1991, Smith *et al.*, 1997a). Assuming this rate of change is maintained over long periods of time, it can be used to estimate the "age" of the individual genotypes and subtypes. Within subtype 1b, variants were estimated to have diverged 70-80 years ago and the absence of any genotype 1b specific country or region suggested that the initial spread was relatively rapid (Smith *et al.*, 1997a). In Western countries, genotype 3a variants were estimated to have diverged only 40 years ago (Simmonds, 2001) and this is consistent with the epidemiological link between genotype 3a and illicit intravenous drug use (IVDU) and needle sharing, which has dramatically increased since the 1960s. In contrast, the genotype 2a, 2b and 2c variants are predicted to have diverged 90-150 years ago (Simmonds, 2001). Documentation of sequence variability is not as good in non-Western countries but the wide range of subtypes in certain regions suggests that a particular genotype has been present in the human population for a long time and possibly that these represent the places of origin for a particular genotype. For example, genotype 2 shows considerable sequence diversity in Western African countries (Ruggieri *et al.*, 1996), as does genotype 4 in Central Africa (Xu *et al.*, 1994). The greatest diversity of genotypes 3 and 6 is found in Southern Asia (Simmonds, 2001).

Genotypes of HCV also differ with respect to disease manifestation and response to interferon therapy, as will be discussed later in this chapter (sections 1.7.3. and 1.8.2.).

1.3.2. Variability in the HCV genome

The HCV genome shows various levels of sequence diversity along its length. The 3'UTR and 5'UTR are highly conserved between genotypes, consistent with their roles in replication and translation. The core protein (see section 1.4.1) is the most conserved of the viral proteins with nucleotides in the core gene of isolates from all 6 genotypes having 79.4-99% identity (Bukh *et al.*, 1994). In contrast, the two envelope-encoding regions (E1 and E2) are highly divergent (Kato *et al.*, 1992). The E2 protein has been found to have an unusually high degree of amino acid variation in its N-terminus (Hijikata *et al.*, 1991b, Weiner *et al.*, 1991, Kato *et al.*, 1992). This hypervariable region (HVR) is the most variable region of the HCV genome. The HVR, mapped to aa 384-411, appeared to be a target for anti-HCV neutralising responses (Zibert *et al.*, 1995, Kato *et al.*, 1993) and it has been suggested that this region is subject to immune pressure which leads to the selection of escape mutants (Weiner *et al.*, 1992, van Doorn *et al.*, 1995).

1.3.3. Viral quasispecies

HCV variation is observed on four levels; genotypes, subtypes, quasispecies and the variants within a quasispecies. A quasispecies is defined as a population of viruses that share a common origin but which have distinct genomic sequences (Smith *et al.*, 1997b). HCV circulates in a single individual as a population of different but closely related genomes and therefore can be described as having a quasispecies distribution (Martell *et al.*, 1992). This quasispecies nature has been attributed to the RNA polymerase lacking proof-reading 3'-5' exonuclease activity and therefore each round of replication is accompanied by the incorporation of mutations. The mutation rate of the HCV genome has been estimated to be $\sim 10^{-3}$-10^{-4} nucleotide substitutions per site per year (Ogata *et al.*, 1991, Smith *et al.*, 1997a). However, as mentioned above, certain regions are more prone

to a substitution than others, which has been interpreted as meaning that these regions are under selective pressure. The flavivirus, HGV is closely related to HCV and its polymerase is presumed to have a similar error rate to that of HCV. However, compared to the 30% sequence diversity between HCV genotypes, HGV isolates from around the world differ by only 14% (Smith *et al.*, 1997c). Therefore the heterogeneous nature of HCV cannot be explained solely by the error prone RNA polymerase. This mixed virus population is thought to confer advantage to the virus in that it allows rapid selection of a variant that displays "better fitness" in response to a new environmental pressure. It has been suggested that the diversity of HCV quasispecies correlates with the progression of liver disease and contributes to the high level of chronicity seen in HCV-infected patients (Yuki *et al.*, 1997) (see section 1.7.2.2). This might be due to the selection of variants which are able to evade the immune system and therefore establish a persistent infection.

A difficulty in interpreting the degree of variation from nucleotide sequence data is that most data were obtained from amplification of the viral RNA by RT-PCR. Both the reverse transcriptase and *Taq* DNA polymerase enzymes have high error rates themselves and therefore these errors could be incorrectly interpreted as virus heterogeneity. This needs to be taken into consideration when analysing the sequence diversity of a virus within an infected individual. To avoid this problem, it has been suggested that, rather than analysing cloned sequences, it is preferable to sequence the PCR product directly (Smith *et al.*, 1997b). An error produced during PCR will be present only in a minority of products and will not be detected against the background majority.

1.4. The HCV structural proteins

1.4.1. Core

Core is the first peptide to be translated. It is cleaved from the polyprotein at residue 191 to produce a 23 kDa product (p23) but the presumed mature form of core is produced by

additional cleavage between residues 174 and 191 to give a 21 kDa species (p21) (Hussy *et al.*, 1996). Processing of p21 and p23 is dependent on the presence of microsomal membranes, indicating that cleavage is mediated by membrane-associated proteases (Santolini *et al.*, 1994). Core is highly conserved between the 6 genotypes and the hydrophobicity profile of core identifies 3 domains within the protein (Hope & McLauchlan, 2000) (Figure 1.6). Domain 1 (aa 1-122) contains a high proportion of basic residues and has two hydrophobic regions. Domain 2 (aa 123-174) is more hydrophobic than domain 1, while domain 3 (aa 175-191) is highly hydrophobic and acts as the signal sequence for E1, but is absent in p21 (Santolini *et al.*, 1994). Electron microscopy indicated that the majority of core was associated with either the cytoplasmic face of the ER or with granular structures, which were subsequently identified as lipid droplets (Moradpour *et al.*, 1996, Barba *et al.*, 1997). There is also evidence of nuclear localisation of a form of p21 core, which is conformationally different from the cytoplasmic form (Yasui *et al.*, 1998).

Core is presumed to form the viral capsid by its similar position in the genome and protein composition by comparison with the capsids of related flaviviruses and pestiviruses. However, very little is known about virus assembly as expression of the structural proteins in mammalian cells does not produce virus particles. Virus-like particles (VLPs) can be produced by infection of insect cells with recombinant baculoviruses (Baumert *et al.*, 1998) which express the HCV structural proteins. The core protein in the VLPs is the same size as that expressed in mammalian cells but because the VLPs are not released from infected cells, complete virus assembly cannot be analysed. In its role as the capsid protein, core is predicted to homo-oligomerise. Regions for oligomerisation have been identified in domain 1 (aa 36-91 and aa 82-102) (Matsumoto *et al.*, 1996) and another potential region has been described in domain 2 but has not been confirmed (Yan *et al.*, 1998). In positive-sense RNA viruses (e.g. Sindbis virus, rubella virus), the capsid protein has been shown to bind specifically to the plus strand viral RNA, and likewise, core interacts specifically with the positive-sense HCV genome (Shimoike *et al.*, 1999). VLPs produced in the baculovirus system have also been shown to encapsidate positive-sense HCV RNA (Baumert *et al.*, 1998). There was a suggestion that core protein could

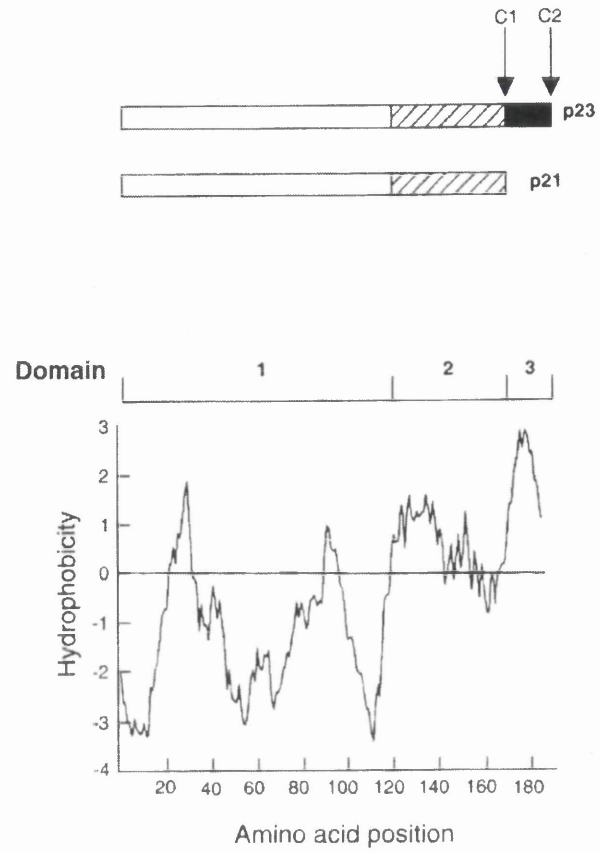


Figure 1.6. Hydrophobicity profile of the core polypeptide and the three proposed domains of core. p23 is the product of cleavage at C2 (position 191/192) and p21 is the product of cleavage at C1 (between residues 174 and 191). (Taken from McLauchlan, 2000).

modulate translation from the IRES (Shimoike *et al.*, 1999) but this effect has since been attributed to the core coding RNA sequence and not the protein itself (Wang *et al.*, 2000). One study has reported an interaction between core and E1 (Lo *et al.*, 1996) but no other interactions with HCV proteins have been reported, probably because core is highly insoluble and membrane-bound following extraction from cells.

Core has been shown to bind to several cellular proteins. These include heterogeneous nuclear ribonucleoprotein K (hnRNP K) (Hsieh *et al.*, 1998), lymphotoxin β receptor (LT-βR) (Chen *et al.*, 1997), tumour necrosis factor receptor 1 (TNF-R1) (Zhu *et al.*, 1998) and an RNA helicase from the DEAD box family of proteins (You *et al.*, 1999, Owsianka & Patel, 1999). The functional significance of these interactions has been difficult to ascertain in the absence of a cell culture system for HCV. Core expression has been demonstrated to affect apoptosis (Ruggieri *et al.*, 1997, Zhu *et al.*, 1998), possibly as a result of its interactions with LT-βR and TNF-R1, both of which are known to be involved in apoptosis. Core has also been implicated in cellular transformation and the development of hepatocellular carcinoma (Ray *et al.*, 1996, Moriya *et al.*, 1998), transcriptional regulation (reviewed in McLauchlan, 2000), immune presentation (Large *et al.*, 1999) and lipid metabolism (Moriya *et al.*, 1997). It would appear that core has multiple functions but these need to be confirmed in a model system that reproduces viral replication and disease pathogenesis as seen in humans.

1.4.2. p7

p7 is located between E2 and NS2 in the genome. Processing of the polyprotein at the E2/p7 and p7/NS2 junctions is mediated by a host signal peptidase to release a highly hydrophobic 7 kDa product (Lin *et al.*, 1994). However, cleavage at the E2/p7 junction is inefficient and leads to the production of two E2-specific species, E2 and E2-p7 (Mizushima *et al.*, 1994, Lin *et al.*, 1994). Baumert *et al.* (1998) demonstrated that p7 is not required for VLP formation in insect cells. In a related pestivirus, bovine viral diarrhoea virus (BVDV), p7 is required for generation of infectious virus, although the E2-p7 product is not (Harada *et al.*, 2000). The pestivirus p7 protein is known not to be a

major structural component of the virion (Elbers *et al.*, 1996) but its true function remains unknown. Similarly, the role of HCV p7 (and E2-p7) is unknown and awaits the development of a mammalian model system capable of generating virus particles.

1.4.3. E1 and E2

Sequence comparison with related flaviviruses and pestiviruses suggested that E1 and E2 were the viral envelope proteins (Choo *et al.*, 1991). On expression of the polyprotein, E1 and E2 were identified as glycosylated proteins of approximately 30 kDa and 70 kDa respectively (Grakoui *et al.*, 1993a, Selby *et al.*, 1993). Cleavage of E1 from the polyprotein is mediated by an ER-associated host signal peptidase between amino acids 191/192 and 383/384 (Hijikata *et al.*, 1991a). Cleavage of the C-terminus of E2 is not as efficient and various E2 species have been described. These include E2, E2-p7 and E2-p7-NS2, which are released after cleavage between amino acids 745/746, 810/811 and 1027/1028 respectively (numbering according to strain H77). The E2-p7-NS2 product is a short-lived precursor and is cleaved to release either E2 or E2-p7, which is stable and sometimes remains uncleaved (Lin *et al.*, 1994, Dubuisson *et al.*, 1994).

Both E1 and E2 are modified by *N*-linked glycosylation. This was determined by endoglycosidase digestion, which resulted in deglycosylated forms of E1 and E2 with molecular weights of approximately 21 kDa and 36 kDa respectively (Grakoui *et al.*, 1993a). From the amino acid sequence of genotype 1 (H77), E1 is predicted to have 5 glycosylation sites with 11 in E2 (Miyamura & Matsuura, 1993). In the other strains and genotypes, the number of predicted glycosylation sites differs by up to 3. It was recently shown that the predicted 5th glycosylation site at position 325 in genotype 1a E1 was not used (Meunier *et al.*, 1999). Mutagenesis revealed that this was due to the presence of a proline residue immediately following the recognition sequence.

The HCV glycoproteins have hydrophobic domains in the C-termini which act as membrane anchors and gives the proteins a type I membrane topology (Rice, 1996). Deletion of the hydrophobic domain results in secretion of the protein and this allows the

transmembrane domain (TMD) to be mapped to specific residues (Figure 1.7). The TMD of E2 maps to aa 718 to 746 (Cocquerel *et al.*, 1998) and is composed of two hydrophobic stretches connected by a short hydrophilic segment (Cocquerel *et al.*, 2000). The TMD of E1 has a similar organisation and was initially mapped to aa 311 to 383 (Michalak *et al.*, 1997). However, a study using chimeric proteins of the CD4 ectodomain fused to C-terminal hydrophobic sequences of E1, identified aa 353 as being the N-terminal limit of the E1 TMD (Cocquerel *et al.*, 1999). It was suggested that E1 contained a second membrane anchor between aa 262 and 290 (Matsuura *et al.*, 1994) but proteins truncated beyond these residues were still secreted, so it is unlikely that this region functions as a true membrane anchor. The TMDs of E1 and E2 are multifunctional. As well as anchoring the protein in the lipid membrane, their C-terminal halves act as ER signal sequences for E2 and p7 respectively and play a major role in subcellular localisation and assembly of the HCV envelope glycoprotein complex (see sections 1.4.4. and 1.4.5.).

1.4.4. HCV glycoprotein complexes

Expression of the HCV polyprotein and identification of the cleavage products revealed that E1 was coprecipitated by E2 antisera, implying that the E1 and E2 proteins interact to form a complex (Grakoui *et al.*, 1993a). These E1E2 complexes were also identified in several other studies using both viral and non-viral expression systems (Lanford *et al.*, 1993, Ralston *et al.*, 1993, Selby *et al.*, 1994, Dubuisson *et al.*, 1994, Deleersnyder *et al.*, 1997). Further characterisation of the E1E2 complex has relied on these expression systems because of the lack of a cell culture system that produces viral particles. The recent report of a full-length selectable replicon (Pietschmann *et al.*, 2001) paves the way for characterisation of the glycoprotein complex in the context of a replicating genome but unfortunately there is still no evidence of virus particle assembly in this system.

Initial characterisation of the E1E2 complex expressed using vaccinia virus recombinants provided contradictory reports, with Ralston *et al.* (1993) describing a non-covalently-linked complex, whereas Grakoui *et al.* (1993a) reported a complex linked by disulphide bonds. Subsequent studies using different HCV strains, expression vectors and cell lines

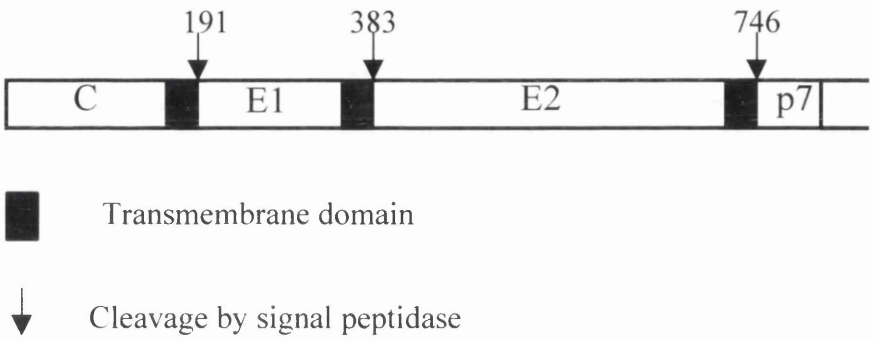


Figure 1.7. Illustration of the transmembrane domain position of the HCV structural proteins. The positions of the C-terminal residues of C, E1 and E2 are indicated at the top with numbering according to genotype 1, strain H77 (Ogata *et al.*, 1991).

showed that two forms of complex were detected in the same cell extract (Dubuisson *et al.*, 1994, Dubuisson & Rice, 1996). They were a heterodimer of E1 and E2 stabilised by non-covalent interactions and a high molecular weight, disulphide-linked aggregate. The non-covalently associated E1E2 heterodimer might be the functional subunit of the HCV virion envelope. The structure of the glycoproteins on the virion envelope has not been determined but the heterodimer is commonly termed the "native" complex. This is formed by proteins following a productive folding pathway (Deleersnyder *et al.*, 1997). Disulphide-linked aggregates are thought to be misfolded proteins that have followed a non-productive folding pathway (Dubuisson *et al.*, 1994) (Figure 1.8). It is possible to distinguish between these two complexes with the use of non-reducing gel electrophoresis conditions and conformation-dependent antibodies. In the absence of reducing agents such as β-mercaptoethanol (β-ME) or dithiothreitol (DTT), the disulphide-linked proteins are retarded at the top of the gel, whereas the non-covalently linked proteins migrate as monomers. Deleersnyder *et al.* (1997) were able to distinguish between the two forms of complex by using a conformation-sensitive E2 antibody (H2) that selectively recognised the native complex. Another such antibody (H53) has also been used for similar purposes (Cocquerel *et al.*, 1998). Other E2 or E1 antibodies used in studies on complex formation have recognised proteins involved in both aggregates and native complex. Interestingly, there are reports of human monoclonal antibodies that recognise both aggregates and native complex (Cardoso *et al.*, 1998, Habersetzer *et al.*, 1998). It has been suggested that aggregate formation is an artefact of high-level expression of the glycoproteins in cell culture but the isolation of these human antibodies suggests that both forms of complex are present in a natural infection.

1.4.4.1. Native E1E2 complexes

Kinetic studies have demonstrated that the formation of native complex is slow (Deleersnyder *et al.*, 1997). Dubuisson & Rice (1996) determined the rate of intramolecular disulphide bond formation in E1 and E2 by analysing the appearance of denatured, oxidised forms of the proteins, which adopt a more compact structure due to accumulation of intramolecular disulphide bonds and therefore have greater mobility on

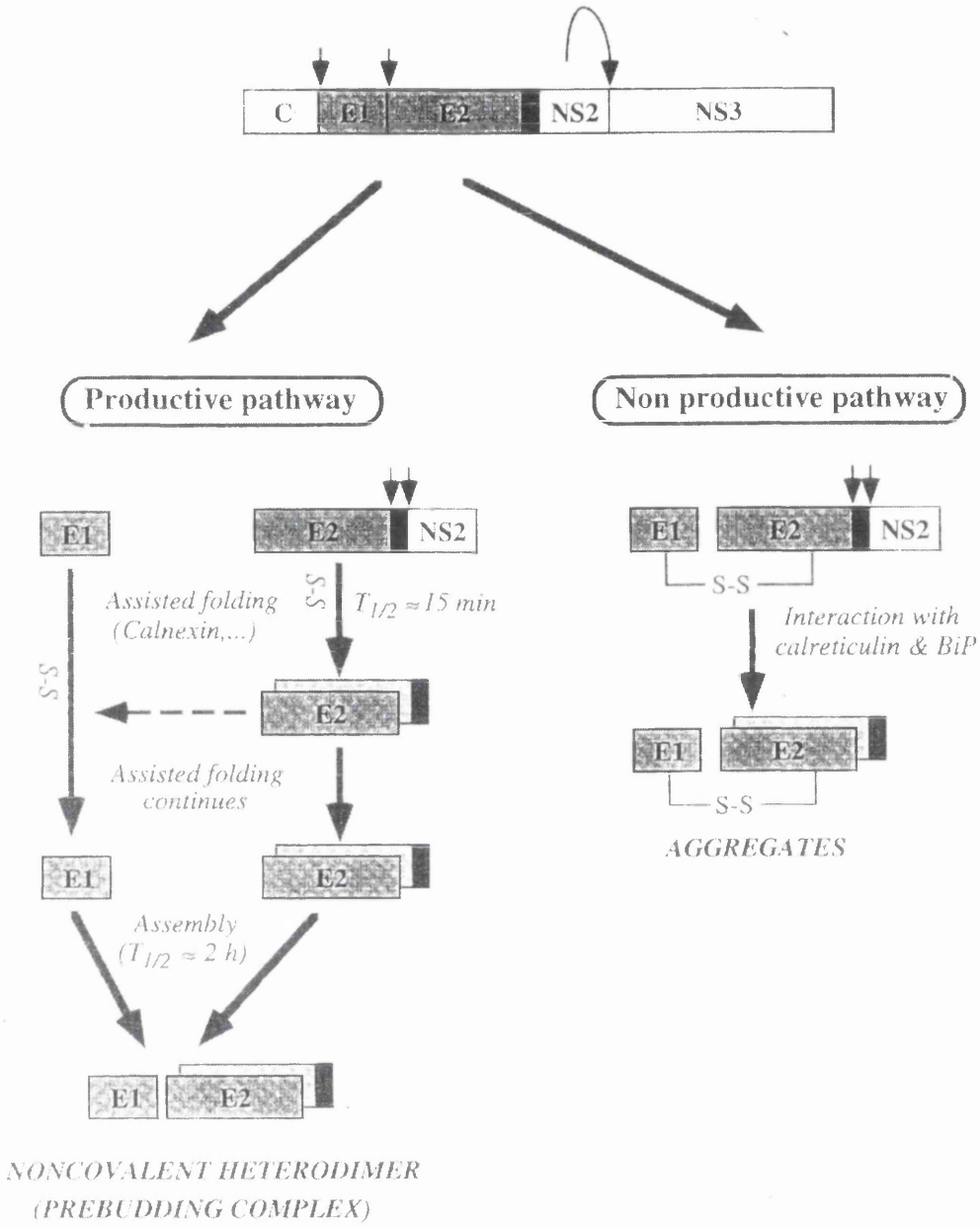


Figure 1.8. A model for HCV glycoprotein assembly. After cleavage from the polyprotein, the glycoproteins can follow two different pathways. In the non productive pathway, E1 and E2 form heterogeneous aggregates stabilised by disulphide bonds. These aggregates can interact with the molecular chaperones BiP and calreticulin. In the productive pathway, E1 and E2 fold slowly and interact to form a noncovalent heterodimer. During their folding, HCV glycoproteins interact with calnexin, and E2 plays a chaperone-like role for E1. Disulphide bonds are indicated by S-S. (Taken from Dubuisson, 2000).

the gel. Intramolecular disulphide bond formation for E1 was considerably slower than for E2, suggesting that this may be the rate limiting step for formation of native E1E2 complexes. The conformation-sensitive antibody that specifically recognised E2 associated with native complex, recognised both E2 and the unprocessed E2-p7 species (Deleersnyder *et al.*, 1997), although it is not known if both E2 and E2-p7 are found in mature virions. Certainly for the BK strain of HCV (genotype 1b), cleavage at the E2-p7 junction is known to be efficient (Dubuisson *et al.*, 1994) and therefore, for this strain at least, it is unlikely E2-p7 would be part of the native complex.

1.4.4.2. Aggregated E1E2 complexes

HCV glycoprotein aggregates are characterised by intermolecular disulphide bonds and a molecular mass greater than 158 kDa (Seong *et al.*, 1998). It has also been shown that E1 and E2 form aggregates when expressed individually, therefore glycoprotein aggregates may contain intermolecular disulphide bonds between E1/E2, E1/E1 and E2/E2 (Michalak *et al.*, 1997). It appears that this is further complicated by the presence of non-covalent interactions as well as covalent interactions in the aggregate (Patel *et al.*, 1999a). As mentioned earlier, it has been suggested that aggregate formation is an artefact of glycoprotein over-expression in cell culture. However, the appearance of such aggregates in cell lines where protein expression is tightly regulated, suggests that this is not the case (Duvet *et al.*, 1998). Therefore aggregate formation may be an intrinsic property of HCV glycoproteins. It has been suggested that this may be a tactic employed by the virus to down-regulate particle formation and hence minimise exposure of viral antigens to the immune system (Choukhi *et al.*, 1999). Fournillier *et al.* (1999) have also shown that the aggregate can induce antibodies that are able to neutralise binding of E2 to susceptible cells and furthermore, the native complex was unable to induce a significant antibody response. In conclusion, the glycoprotein aggregate may well play a functional role in the virus life cycle.

1.4.4.3. Requirements for E1E2 complex formation

Deletion of the C-terminal transmembrane domain of E2 has been shown to abolish the formation of the E1E2 complex (Selby *et al.*, 1994). Cocquerel *et al.* (1998) demonstrated that a chimeric protein consisting of the E2 ectodomain fused to the CD4 membrane anchor was no longer able to interact with E1. Conversely, it was shown that a chimeric protein containing the TMD of E2 fused to the ectodomain of Herpes simplex virus (HSV) 1 gD was sufficient for interaction with E1 (Patel *et al.*, 2001). The E2 TMD was also essential for correct folding of E1 and native complex formation. There is also a report that N-terminal regions in E2 and E1 are important for the E1/E2 interaction (Yi *et al.*, 1997). The major E1-interacting site on E2 was mapped to a discontinuous region within aa 415-500 and the highly conserved "WHY" tripeptide (aa 489-491) was crucial for this interaction. These data indicate that, in addition to an interaction via their TMDs, E1 and E2 may interact via regions in their ectodomains as well.

There is no apparent difference in the folding of E2 when expressed in the presence and absence of E1 but in the absence of E2, E1 does not fold correctly (Michalak *et al.*, 1997). Using chimeric proteins, Patel *et al.* (2001) showed that it is the C-terminal domain of E2 (containing the TMD) which is essential for ensuring correct folding of E1. Therefore E2 has a "chaperone-like" role which is predominantly mediated by its TMD but may also involve regions in the ectodomain. A recent report also indicated that core was required for the correct folding of E1 (Merola *et al.*, 2001). Dubuisson *et al.* (2000) have demonstrated that glycosylation of E1 is improved by the presence of E2 and membrane insertion of E2 has been shown to be stabilised by the presence of E1 (Cocquerel *et al.*, 2001). Therefore the maturation of the glycoproteins is tightly coupled. In addition, E1 and E2 have been shown to interact with molecular chaperones such as calnexin, calreticulin and BiP (Dubuisson & Rice, 1996, Choukhi *et al.*, 1998), which promote folding of glycoproteins in the ER (see section 1.4.6).

1.4.5. Subcellular localisation of HCV glycoproteins

The HCV glycoprotein complexes are retained in the ER, as demonstrated by their intracellular distribution, lack of complex glycans and the absence of cell-surface expression (Dubuisson *et al.*, 1994). This was confirmed by Cocquerel *et al.* (1998) and the E2 protein expressed in the absence of E1 showed a similar intracellular distribution to the E1E2 complex and was not resistant to digestion with Endo H. This indicated that E2 had not moved into the medial and trans-Golgi apparatus. Therefore an ER retention signal should be present in the ER. Replacement of the TMD of E2 with the membrane anchor sequence of CD4 (a protein normally expressed on the cell surface), resulted in expression of the E2 ectodomain on the cell surface and conversely, a chimeric protein containing the ectodomain of CD4 fused to the TMD of E2 was retained in the ER (Cocquerel *et al.*, 1998). These data indicated that an ER retention signal is located in the C-terminal 29 amino acids of E2. Similarly, the use of chimeric proteins has identified an ER retention signal in the C-terminal 31 amino acids of E1 (Cocquerel *et al.*, 1999, Flint & McKeating, 1999). ER retention was not due to misfolding of E1 in the absence of E2, as the chimeric proteins did not form aggregates and were recognised by a conformation-sensitive antibody (Flint & McKeating, 1999). It was concluded that the TMDs of E1 and E2 provided signals for ER retention of the glycoprotein complex. The highly conserved, charged residues between the two hydrophobic domains in the TMDs of E1 and E2 were shown to be crucial to ER retention (Cocquerel *et al.*, 2000).

ER localisation could be due to retention in the ER or retrieval from the *cis* Golgi. To determine if the HCV glycoproteins can cycle between these two compartments, Duvet *et al.* (1998) examined the sensitivity of the glycoproteins to the glycosidase Endo D. If glycoproteins enter the *cis* Golgi, their oligosaccharides chains would be exposed to Golgi α-mannosidase I, which processes high mannose ($Man_9GlcNAc_2$) sugars to $Man_5GlcNAc_2$ (Kornfeld & Kornfeld, 1985). Molecules containing $Man_5GlcNAc_2$ should be sensitive to Endo D but the HCV glycoproteins remained Endo D-resistant, indicating that they did not cycle between the ER and cis Golgi departments. However, a recent study has provided contradictory evidence that the HCV structural proteins were located in the

intermediate compartment and *cis* Golgi complex as well as the ER (Martire *et al.*, 2001). In this study the HCV structural proteins were expressed from a stably transfected clone and were under the control of an inducible promoter. The authors proposed that this system prevented protein over-expression, which could lead to accumulation of misfolded proteins in the ER. Using confocal immunofluorescence microscopy, cell fractionation and immunoelectron microscopy, a small proportion of the proteins were localised to the intermediate compartment and cis Golgi region, while the majority were retained in the ER. The accumulation of complex *N*-linked glycans was not examined in this study, so these results cannot be directly compared to those of Duvet *et al.* (1998). Confirmation of the exact localisation on the HCV glycoproteins awaits the development of a model system in which to study virus particle assembly.

1.4.6. Association of HCV glycoproteins with folding proteins

The ER contains molecular chaperones, which associate specifically with incompletely folded proteins and increase the efficiency by which they acquire their correct 3-dimensional structure. They therefore play a role in quality control in that they prevent exit of misfolded proteins from the ER (see section 1.5.). The ER chaperone proteins include the immunoglobulin heavy-chain binding protein (BiP/GRP78), glucose regulated protein (GRP94), calreticulin and calnexin. An interaction with calnexin and/or calreticulin has been shown for several viral glycoproteins such as vesicular stomatitis virus G protein (Hammond & Helenius, 1994), Semliki Forest virus p62 and E1 and influenza haemagglutinin (Molinari & Helenius, 2000). The preference for glycoproteins is due to their lectin-like affinity for monoglucosylated *N*-linked glycans and binding and release of substrate glycoproteins depends on trimming and reglucosylation of the *N*-linked glycans (Herbert *et al.*, 1995).

Calnexin, calreticulin and BiP have all been shown to interact with E1 and E2, whereas no interaction has been detected with GRP94 (Dubuisson & Rice, 1996, Choukhi *et al.*, 1998, Liberman *et al.*, 1999). Calnexin was shown to interact preferentially with monomeric HCV glycoproteins and native complexes, while calreticulin and BiP associated

preferentially with aggregates (Choukhi *et al.*, 1998). This indicated that calnexin was involved in the productive folding pathway of HCV glycoprotein complex formation, whereas calreticulin and BiP were involved in the non-productive pathway. The association of the HCV glycoproteins with calnexin and calreticulin was faster than with BiP but there was no appreciable difference between the kinetics of association with calnexin and calreticulin (Choukhi *et al.* 1998). This suggested that rather than interacting sequentially with the glycoproteins, calnexin and calreticulin recognise proteins that are in different states of folding. Over-expression of calnexin, calreticulin or BiP did not improve the formation of native E1E2 complexes (Choukhi *et al.*, 1998), which suggested that a proper balance of chaperone activities is required for optimal folding.

The ER also contains other folding enzymes such as protein disulphide isomerase (PDI) and peptidyl proline isomerase (PPI). PDI catalyses the formation of disulphide bonds and PPI is involved in proline isomerisation (Gething & Sambrook, 1992). The conservation of cysteine and proline residues in the HCV glycoproteins suggests that these residues are important for folding, so the two afore mentioned enzymes are likely to be involved in folding of E1 and E2, but as yet this has not been documented.

1.5. Folding of glycoproteins in the ER

The lumen of the ER is a critical site for protein maturation. Newly synthesised membrane and secretory proteins are transported into the ER in a largely unfolded state. Specialised proteins within the ER facilitate optimal folding and assembly of the new proteins, which confers stability and determines their biological activity.

1.5.1. Targeting proteins to the ER

Proteins that are destined for secretion or insertion into cellular membranes are targeted to the ER during translation by a signal peptide. This signal sequence is characterised by its hydrophobic nature and is recognised by the signal recognition particle (SRP) as it emerges from the ribosome (Siegel, 1995). The ribosome/nascent polypeptide chain/SRP complex docks at the ER membrane via the SRP receptor. Once bound, SRP is displaced and translation continues while the nascent polypeptide chain is simultaneously transferred into the translocon. The translocon is a multisubunit complex in the ER membrane. It forms a membrane channel and is responsible for the transfer of polypeptide chains across the lipid bilayer (Figure 1.9). On the lumenal side of the ER membrane, signal peptidase cleaves the signal peptide from the polypeptide chain.

1.5.2. Glycosylation

During translocation of the polypeptide chain, oligosaccharyl transferase (which is associated with the translocation machinery) adds *N*-linked glycan to the growing polypeptide chain as 14-residue oligosaccharides ($Glc_3Man_9GlcNAc_2$) (Parodi, 1999). The oligosaccharides are added to the side chains of asparagine residues in the consensus sequence N-X-S/T, where X is any amino acid except proline. Transfer of the oligosaccharide occurs co-translationally and can occur as soon as 12-14 aa residues of the polypeptide chain have been transferred into the ER lumen, if a suitable acceptor site is present. Immediately following addition of $Glc_3Man_9GlcNAc_2$, three glucose residues are removed sequentially by glucosidase I (for the first glucose residue) and glucosidase II (for the remaining two residues) to yield $Man_9GlcNAc_2$ (Parodi, 1999). During this process an intermediate glycan ($Glc_1Man_{7-9}GlcNAc_2$) is formed and this is recognised by calnexin and calreticulin, which bind and monitor folding of the protein (see section 1.5.3). Once correctly folded, the chaperones are released. The high mannose oligosaccharide moieties are further trimmed by ER mannosidase, yielding Man_8 species. Transport to the Golgi compartments will result in further modification of the glycans.

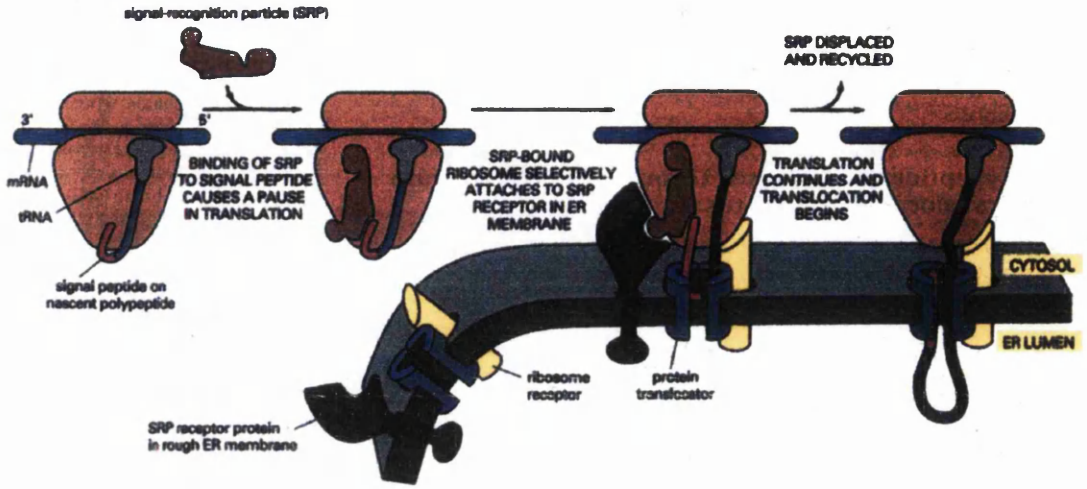


Figure 1.9. ER signal peptides and SRP direct ribosomes to the ER membrane. The SRP binds to the exposed ER signal peptide and to the ribosome, thereby inducing a pause in translation. The SRP receptor in the ER membrane binds the SRP ribosome complex. In a poorly understood reaction that involves multiple 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).

1.5.3. Protein folding in the ER

Folding of proteins in the ER begins before translation is complete, although the time taken to reach the final conformational state varies between proteins. The ER contains specialised proteins which assist the newly synthesised protein in reaching its final conformational state. As mentioned earlier, folding enzymes such as PDI and PPI are responsible for catalysing disulphide bond formation and isomerisation of proline peptide bonds, respectively. These two folding steps may be rate-limiting. As well as forming native disulphide bonds, PDI is also involved in unscrambling wrong cysteine bridges into native ones (Braakman & van Anken, 2000).

The ER also contains a group of proteins termed molecular chaperones, which bind to unfolded and misfolded proteins to assist them in attaining their native state and prevent their exit from the ER. Two major classes of molecular chaperones have been identified. The first class includes members of the heat shock protein 70 class, which in the ER is BiP (grp78). BiP has an affinity for hydrophobic stretches of amino acids, which are usually exposed on unfolded or misfolded proteins. The second class consists of the lectin chaperones, calnexin and calreticulin. Both these chaperones bind to monoglucosylated forms of *N*-linked glycans present on the unfolded protein (Helenius *et al.*, 1997). These are formed either during maturation of the oligosaccharide precursors as described in section 1.5.2. or as a result of reglucosylation of fully-trimmed oligosaccharides by the enzyme UDP-glucose:glycoprotein glucosyltransferase (GT) (Parodi, 1999). This enzyme is responsible for quality control, as reglucosylation results in renewed binding to calnexin and calreticulin. Therefore a cycle of glucose removal and addition is coupled to chaperone binding and release, during which the protein folds and is checked for structural quality (Figure 1.10). Calnexin and calreticulin can apparently bind to the same glycan, as demonstrated by folding of the influenza A virus HA molecule (Herbert *et al.*, 1997) but they sometimes show specificity for particular glycans, which seems to depend on the location of the glycan in the newly synthesised protein. It has been shown that the presence of *N*-linked glycans within 50 residues of the N-terminus promotes association with calnexin/calreticulin rather than BiP (Molinari & Helenius, 2000). The association

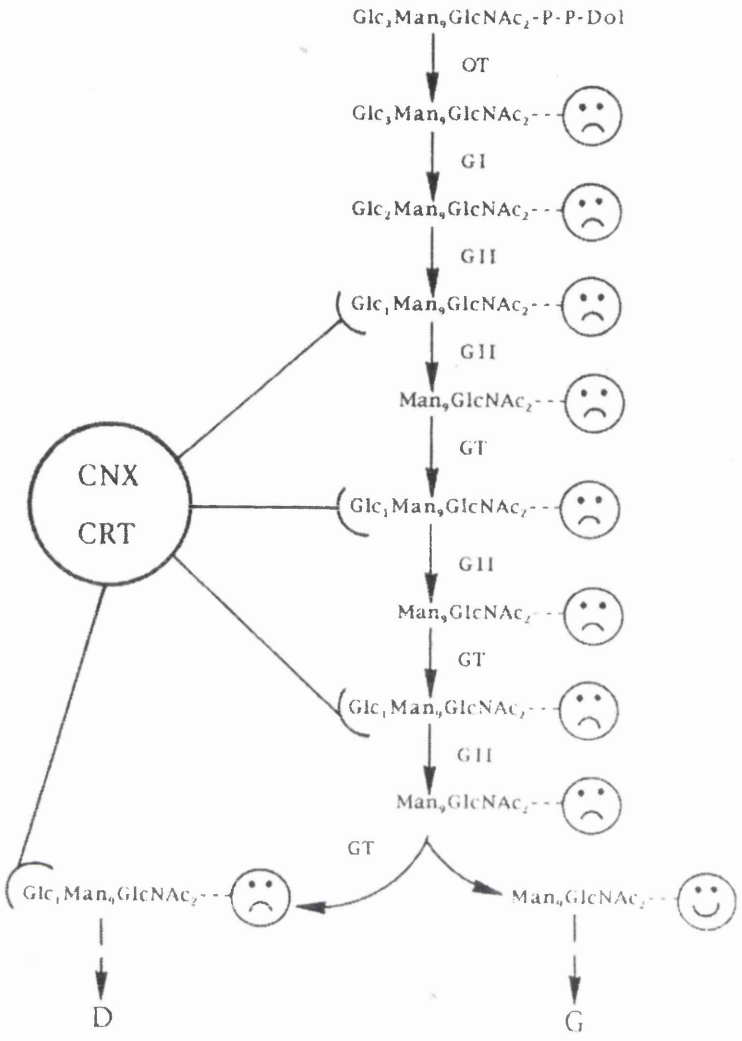


Figure 1.10. Schematic representation of the model proposed for the quality control of glycoprotein folding in the ER. Sad faces represent misfolded glycoproteins whereas happy ones are properly folded species. CNX: calnexin; CRT: calreticulin; OT: oligosaccharyltransferase; GI: glucosidase I; GII: glucosidase II; GT: UDP-Glc:glycoprotein glucosyltransferase; D: degradation; G: golgi. (Taken from Parodi, 1999).

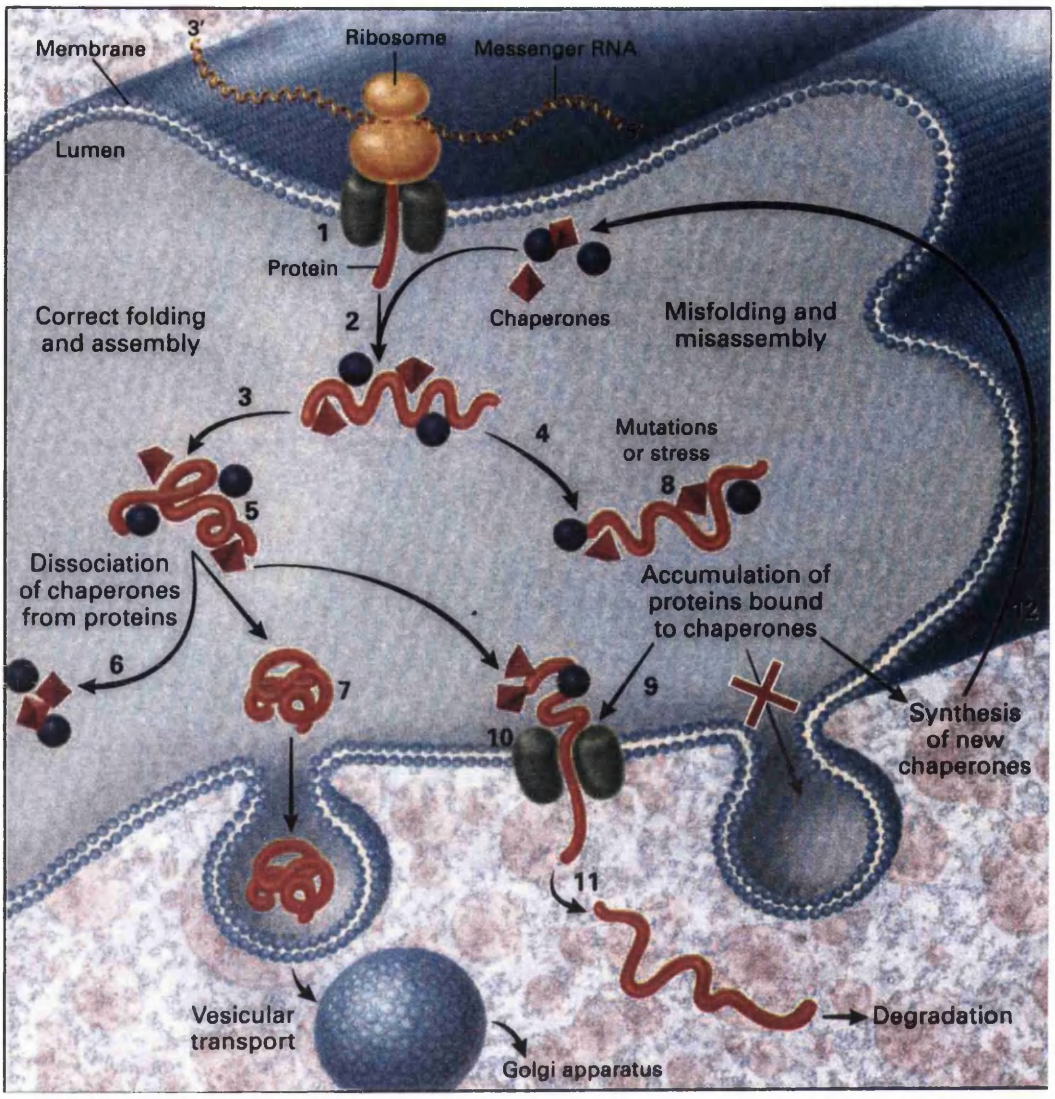


Figure 1.11. Role of molecular chaperones in the endoplasmic reticulum (ER) in protein folding. During synthesis,secretory and membrane proteins are cotranslationally translocated into the lumen of the ER through an aqueous gated channel (1). They then bind to molecular chaperones (2) and begin the folding process, which is facilitated by the chaperones and by folding and processing enzymes (3). After completion of the folding and modifications (5), the proteins (7) dissociate from the chaperones (6) and are transported to the Golgi apparatus by vesicular transport. When folding or assembly of proteins is not completed successfully (4), the misfolded proteins, bound to the chaperones, are retained in the ER (8). This retention also signals the synthesis of new chaperones (12). The misfolded proteins may then be transported back out of the ER to the cytoplasm by retrograde translocation (9 & 10). They are then degraded in the cytoplasm (11).
(Taken from Kuznetsov & Nigam, 1998).

of molecular chaperones with their substrates is normally short-lived but if the protein does not fold properly, this association may become stable and prolonged, thereby retaining the misfolded protein in the ER (Figure 1.11).

1.5.4. ER stress response

The accumulation of misfolded proteins in the ER would drastically affect the functioning of the cell, because approximately one third of all cellular proteins attain their proper tertiary structure in the ER. Eukaryotic cells have devised various mechanisms for relieving the ER stress induced by the accumulation of incorrectly folded proteins. These include transcriptional induction, translational attenuation (together known as the unfolded protein response (UPR)) (Patil & Walter, 2001) and protein degradation.

1.5.4.1. Transcriptional activation

When misfolded protein accumulate in the ER, a signal is sent across the ER membrane into the nuclear and cytoplasmic compartments. This results in transcriptional upregulation of a set of target genes (Kaufman, 1999). These genes include those that encode most of the chaperones and enzymes involved in protein folding in the ER. In this respect it is not unexpected to find that expression of HCV E2 leads to activation of the grp78 (encodes BiP) and grp94 (encodes GRP94) promoters (Liberman *et al.*, 1999), presumably due to the accumulation of E2 aggregates in the ER. The number of known targets of transcriptional upregulation has increased greatly and includes proteins working at every stage of the secretory pathway.

The mechanism of the UPR was initially characterised in the budding yeast, *Saccharomyces cerevisiae*. IRE1 is a central component of the UPR and functions by sensing the accumulation of unfolded protein accumulation in the ER lumen and transmitting signals across the ER membrane. IRE1 is a type I transmembrane protein with a stress-sensing N-terminal domain in the ER lumen and a protein kinase C-terminal domain in the cytoplasm. In unstressed cells, the lumenal domain forms a stable complex

with BiP. BiP is released on accumulation of misfolded proteins in the ER lumen and this converts IRE1 to an oligomeric active state (Bertolotti *et al.*, 2000). This is accompanied by trans-autophosphorylation via its kinase domain, which in turn stimulates the endonuclease activity present in its most C-terminal domain. The IRE1 endonuclease splices out an intron from the *HAC1*^{u} (uninduced) precursor mRNA, which encodes the basic leucine zipper transcription factor Hac1p (Sidrauski & Walter, 1997). The 5' and 3' halves of the mRNA are rejoined by tRNA ligase Rlg1p (Sidrauski *et al.*, 1996), generating a new mRNA, *HAC1*^{i}, which encodes a functional Hac1p. Hac1p enters the nucleus and activates target gene transcription by binding to an upstream sequence called the unfolded protein response element (UPRE) (Cox & Walter, 1996) (Figure 1.12). Mammalian homologues of IRE1 include human IRE1α and mouse IRE1β. IRE1α is expressed in all cell types, whereas IRE1β expression is mainly confined to cells of the gut. Another mammalian protein has recently been identified, which shares homology with IRE1, IRE1α and IRE1β in its lumenal domain but its cytoplasmic domain lacks endonuclease activity and shares homology with eIF2α kinases (Figure 1.13a). This human protein is called PERK or PEK (Shi *et al.*, 1998, Harding *et al.*, 1999).

1.5.4.2. Translational attenuation

A notable feature of the UPR is the rapid repression of protein synthesis. This phenomenon is mediated by PERK (Harding *et al.*, 1999). Its lumenal domain has been shown to be interchangeable with that of IRE1 (Bertolotti *et al.*, 2000), so it is also responsible for sensing ER stress via an interaction with BiP. However, its cytoplasmic domain shows ~40% homology to PKR, an eIF2α kinase. When active, PERK oligomerises and trans-autophosphorylates like IRE1 but its active kinase domain phosphorylates the translation initiation factor, eIF2α (Figure 1.13b). This causes a down-regulation in overall protein translation by preventing the association of mRNA with the ribosomal 60S and 40S subunits. Harding *et al.* (2000) showed that cells in which both copies of the PERK gene had been deleted (PERK$^{-/-}$), have lost the ability to attenuate translation in response to accumulation of unfolded proteins and hence experience greater ER stress, as shown by increased levels of IRE1α phosphorylation.

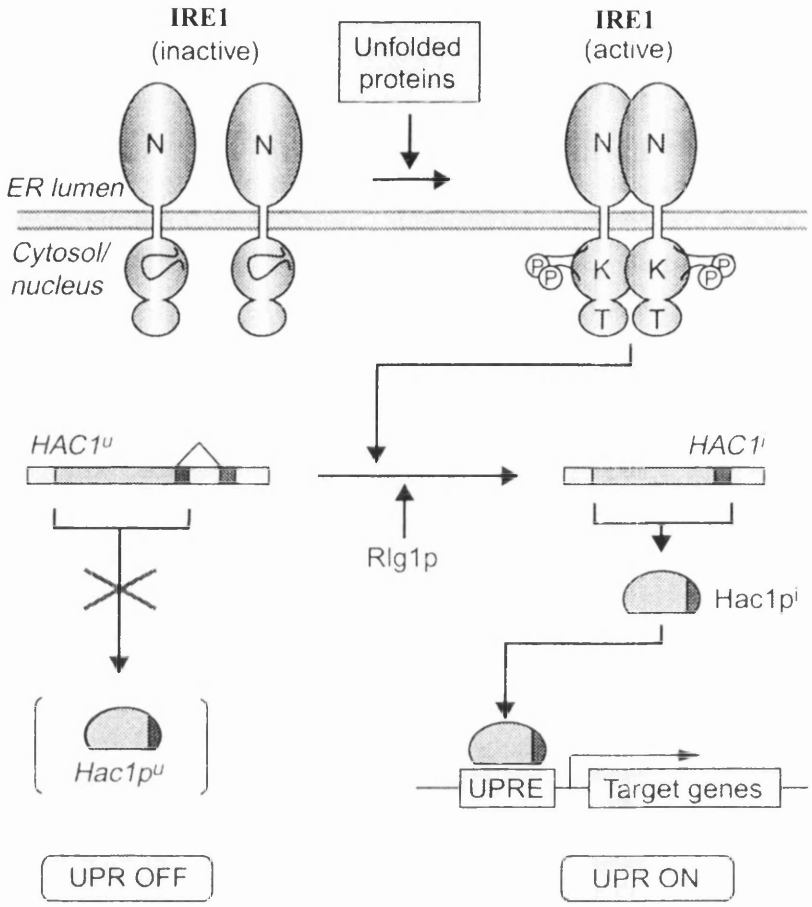


Figure 1.12. Illustration of the unfolded protein response. IRE1 senses the accumulation of unfolded protein in the ER lumen via its N domain. IRE1 then oligomerises, trans-autophosphorylates via its kinase domain (K) and activates the endonuclease in the tail domain (T). The endonuclease cuts *HAC1* mRNA at two sites, removing an intron. The two exons are rejoined by the ligase Rlg1p and *HAC1*^{i} mRNA is translated into Hac1p^{i}, a transcriptional activator that upregulates expression of UPR target genes after binding to the UPRE in the promoters of genes encoding ER-chaperones and other proteins. (Taken from Patil & Walter, 2001).

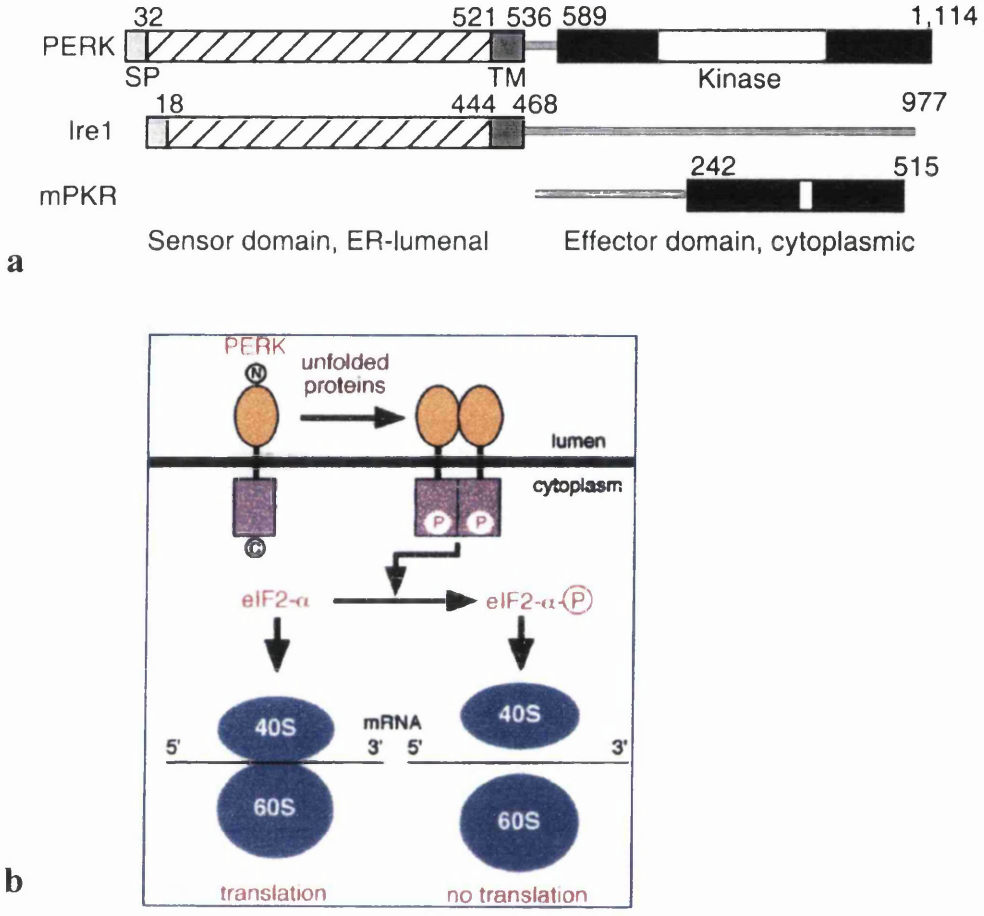


Figure 1.13. **a)** The domain structure of PERK, Ire1 and PKR. SP: signal peptide, TM: transmembrane domain. The lumenal domain is cross-hatched and the predicted two lobes of the kinase domain, conserved between PERK and PKR, are shown in black. (Taken from Harding *et al.*, 1999).

b) Proposed mechanism for ER stress-induced translational attenuation. The N-terminal domain of PERK senses accumulation of unfolded proteins in the ER lumen. It oligomerises and auto-phosphorylates, which activates the kinase domain and results in phosphorylation of eIF2α at serine 51, leading to inhibition of translation initiation. (Taken from Mori, 2000).

Recently another human homologue of IRE1 has been identified, which is also responsible for translational repression. This protein, hIRE1β, which shows 76% identity with the previously identified mouse IREβ, functions as a repressor of protein synthesis through cleavage of 28S ribosomal RNA during ER stress (Iwawaki *et al.*, 2001). Therefore there appear to be two distinct pathways for attenuation of protein translation in response to ER stress. It is not yet known how the ER chaperones escape from this general inhibition of translation.

1.5.4.3. Protein degradation

After prolonged association with the molecular chaperones, particularly BiP, the misfolded proteins are targeted for degradation. It has been shown that these misfolded proteins can be extruded from the ER into the cytoplasm via the translocon, where they undergo ubiquitin- and proteasome-dependent degradation (Figure 1.11). This process is termed ER-associated degradation (ERAD) (Plemper & Wolf, 1999). It has been shown that ERAD and the UPR are intimately co-ordinated, as the degree of activation of the UPR correlates with the severity of the ERAD defect (Patil & Walter, 2001).

Certain viral proteins, such as human cytomegalovirus (HCMV) US2 and US11, and Human immunodeficiency virus (HIV) 1 Vpu can exploit this retrograde translocation to mediate export and degradation of ER-associated MHC class I and CD4 molecules, respectively (Plemper & Wolf, 1999). Accumulation of misfolded proteins in the ER is also responsible for the pathology of several diseases. This is due either to the protein not reaching its final destination, as in cystic fibrosis, or the accumulation of protein in the ER causing cellular damage, as in hepatocellular damage in patients with alpha$_1$-antitrypsin deficiency (Kuznetsov & Nigam, 1998).

1.6. Putative HCV receptors

1.6.1. Cellular tropism of HCV

The cellular tropism of a virus is important when considering candidates for the virus receptor. The principle site of HCV replication was thought to be the liver and therefore it seemed likely that the cellular receptor would be a hepatocyte-specific molecule. However there are also reports of both positive and negative strand HCV RNA being present in polymorphonuclear leukocytes, monocytes/macrophages and B-lymphocytes but not T-lymphocytes (Lerat *et al.*, 1998). Replication in B-cells is an intriguing link with a lymphoproliferative disorder (mixed cryoglobulinaemia) which is associated with HCV infection (see section 1.7.2.3). There are also reports of high prevalence of HCV infection amongst patients with B-cell lymphoma (Dammacco *et al.*, 1998), although this has not been confirmed. These data suggest that HCV may affect B-cell proliferation. Studies on HCV quasispecies have shown the presence of distinct variants between the liver and PBMC fractions (Okuda *et al.*, 1999), which further suggests the possibility of extrahepatic sites of replication.

1.6.2. E2 mediates binding of HCV

Binding of HCV virions to a specific receptor is most likely to be mediated by the outermost proteins on the virion, in this case the envelope proteins E1 and E2. The lack of a model system in which virus particles are produced has meant that it has not been possible to study events involved in virus entry, such as attachment and penetration. Lagging *et al.* (1998) described the production of pseudotypic vesicular stomatitis virus (VSV) particles, which contained chimeric envelope proteins consisting of the HCV E1 and E2 ectodomains fused to the TMD and cytoplasmic tail of VSV G protein. The pseudotype virus was able to infect mammalian cells (HepG2 and BHK) and entry was inhibited by sera from chimpanzees which had been immunised with E1 and E2. This suggested that E1 and E2 could mediate infection of cells. It has subsequently been

shown that cells expressing these same chimeric E1 and E2 proteins can induce cell fusion in a low pH-dependent manner, which suggests entry via an endosomal pathway (Takikawa *et al.*, 2000).

The development of an assay for the binding of E2 to human cells was the first step towards the identification of a cellular receptor. This assay was based on the ability of antibodies or sera to prevent binding of E2 to the human MOLT-4 T-cell line, which was termed neutralisation of binding (NOB) (Rosa *et al.*, 1996). E2 expressed in mammalian cells but not yeast, showed strong binding, while E2 expressed in insect cells showed weak binding. As well as binding MOLT-4 T cells, E2 also bound to hepatocarcinoma cell lines and human B-cells but not mouse cell lines or mouse hepatocytes. Sera from chimpanzees vaccinated with E1/E2 and protected from challenge were shown to inhibit binding of E2 to cells, suggesting that this could be a surrogate test for the presence of neutralising antibodies. Ishii *et al.* (1998) demonstrated the presence of antibodies capable of blocking E2-binding to cells in the sera of patients with chronic hepatitis C. Furthermore, the appearance of antibodies with high NOB titres coincided with viral clearance in those patients who resolved the infection naturally. Overall, these results strongly indicated that E2 is responsible for mediating binding of HCV to cells.

1.6.3. CD81

To identify the E2-binding molecule on human cells, Pileri *et al.* (1998) screened a cDNA library prepared from the MOLT-4 T-cell line and identified a clone which conferred E2-binding to mouse fibroblast cells on transient transfection. This clone encoded human CD81.

CD81 is a 26-kDa surface protein composed of four transmembrane domains and two extracellular loops (reviewed by Levy *et al.*, 1998) (Figure 1.14). It shares this topology with the tetraspanin family of proteins, which are expressed in evolutionarily diverse organisms. CD81 is expressed on most human tissues with the exception of red blood cells and platelets, although most studies have concentrated on expression in lymphoid

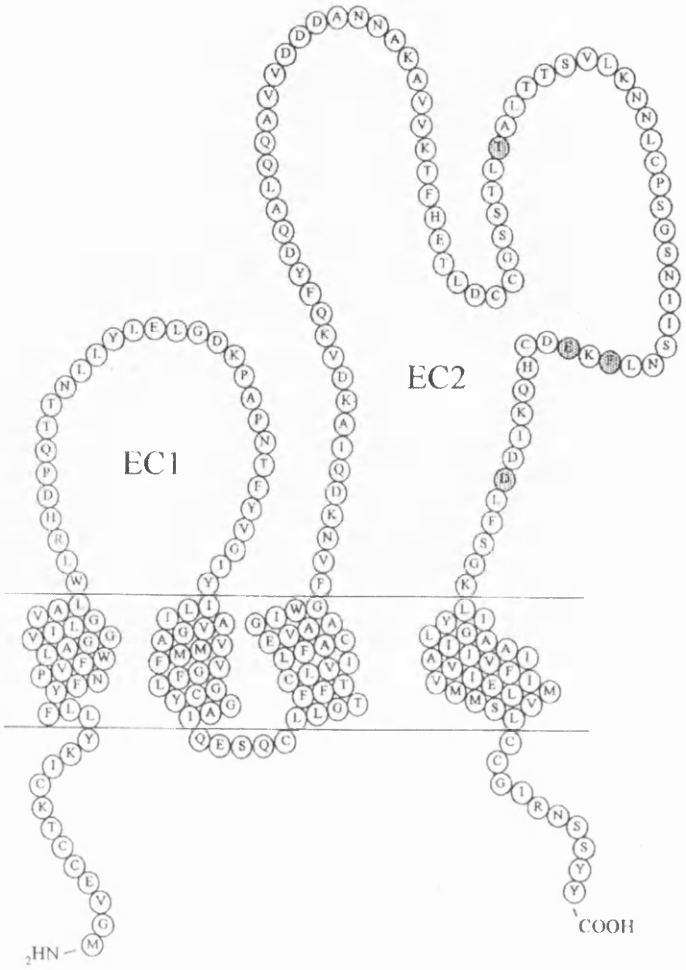


Figure 1.14. The predicted secondary structure of CD81. EC: extracellular domain. Shaded aa indicate the four positions in EC2 where the African Green Monkey sequence differs from the human. (Taken from Flint & McKeating, 2000).

cells. On B-cells CD81 is present in a complex with CD19, CD21 and Leu-13 (Fearon & Carter, 1995). This complex amplifies the signalling cascade generated through the B-cell receptor. Binding of anti-CD81 antibody induces anti-proliferative effects on B-cells and triggers aggregation (Levy *et al.*, 1998). It is not known if these effects mimic those induced by the natural ligand of CD81 or if they are specific to the antibodies. CD81 is also found on the surface of T cells in association with CD4, CD8 and CD82 (Imai *et al.*, 1995) and may be involved in the maturation of T cells. It is likely that CD81 is part of a protein complex on the surfaces of non-lymphoid cells as well, but, as yet, these have not been characterised.

Binding of E2 was mapped to the large extracellular loop (LEL) of CD81 (Pileri *et al.*, 1998). This is the most variable part of the molecule but significantly, it is highly conserved in humans and chimpanzees, the only species permissive for HCV infection. In accordance with this, E2 does not bind to murine, rat, hamster or African green monkey cells (Flint *et al.*, 1999a). However, E2 does bind to rat cells transfected with human CD81, suggesting that additional human-cell-specific factors are not required for the E2 interaction with CD81 (Flint *et al.*, 1999a). African green money CD81 differs from human CD81 at only 4 residues. Mutation of these residues revealed that amino acid 186 in human CD81 was crucial for interaction with E2 (Higginbottom *et al.*, 2000). Additionally, the four cysteine residues in the LEL form two disulphide bridges, the integrity of which was necessary for the binding of E2 (Petracca *et al.*, 2000). The species specificity of the E2/CD81 interaction appeared to correlate with susceptibility to HCV infection but the finding that E2 could bind to CD81 of tamarins (Allander *et al.*, 2000a, Meola *et al.*, 2000), which are not permissive for HCV infection, puts this into doubt.

In an assay more relevant to the *in vivo* situation, recombinant human CD81 LEL attached to polystyrene beads has been shown to precipitate HCV virions from infectious chimpanzee plasma (Pileri *et al.*, 1998, Hadlock *et al.*, 2000). Pre-incubation of the plasma with sera from protected E1E2-vaccinated chimpanzees inhibited binding of virions to CD81, whereas no inhibition was seen with sera from unprotected, vaccinated

chimpanzees. This indicated that antibodies capable of neutralising infection *in vivo* are also capable of inhibiting the interaction of HCV with CD81. Several studies have used antibodies with known epitopes to map the region of E2 involved in the CD81 interaction (Flint *et al.*, 1999a, Yagnik *et al.*, 2000, Patel *et al.*, 2000, Owsianka *et al.*, 2001). There is no absolute agreement from these studies but it is clear that the CD81-binding region on E2 is discontinuous and therefore conformational in nature. This was confirmed when it was shown that the aggregated form of E2 did not bind CD81 and interestingly, intracellular forms of E2 were bound with greater affinity than secreted forms (Flint *et al.*, 2000). The secreted forms of E2 were found to have acquired more complex glycans during passage through the secretory pathways of the cell, which suggests that glycosylation might affect the interaction with CD81. As yet, the type of glycans present on the glycoproteins in the HCV virion is not known. Despite the strong interaction between E2 and CD81, cell-surface expressed E2 does not induce fusion with CD81-positive cells (Flint *et al.*, 1999b), which would be necessary for infection. It therefore seems likely that if CD81 is involved in viral attachment, other factors (cellular and/or viral) are required for virus entry.

1.6.4. LDL receptor

The LDL receptor has also been proposed as a candidate for HCV receptor because HCV particles have been shown to associate with low-density lipoproteins (LDLs) (Prince *et al.*, 1996). Indeed it has been shown that HCV is able to bind to COS-7 cells which express the LDL receptor (Monazahian *et al.*, 1999). Moreover, Agnello *et al.* (1999) have reported that endocytosis of HCV and two other flaviviruses, hepatitis G virus and BVDV, is mediated by LDL receptors. Endocytosis of HCV correlated with LDL receptor activity and was completely inhibited in the presence of anti-LDL receptor antibody. However, small amounts of intracellular HCV were detected in LDL-deficient fibroblasts inoculated with HCV and could not be inhibited by anti-LDL receptor antibody, so viruses can enter these cells by some other means.

In a recent study, Wunschmann *et al.* (2000) compared the characteristics of the LDL receptor and CD81 in terms of binding both HCV and E2. In agreement with previous reports, HCV binding and entry correlated with LDL receptor expression and was inhibited by the addition of LDL but not soluble CD81. In contrast, E2 binding was independent of LDL receptor expression and was inhibited by human soluble CD81 but not mouse soluble CD81 or LDL. There was no evidence of an interaction between E2 and LDL or E2 and LDL receptor, so the mechanism by which HCV interacts with LDL or the LDL receptor remains unclear. The authors concluded that while E2 binds CD81, virus particles utilise the LDL receptor for binding and entry.

1.7. Natural history of HCV infection

1.7.1. Epidemiology and transmission

It is estimated that there are more than 170 million persons infected with HCV world-wide (World Health Organization, 1997). In developed countries the prevalence rate of HCV amongst the general population is less than 3% but higher prevalence rates are seen in parts of Africa and Eastern Europe (Naoumov, 1999). Egypt has a particularly high prevalence rate of 22%, which has been shown to be due to reuse of needles during a national campaign to treat schistosomiasis infections during the 1970s (Frank *et al.*, 2000).

HCV is generally transmitted by the parenteral route, although cases of non-parenteral transmission have also been documented.

Parenteral Transmission of HCV

There is more than an 80% chance of acquiring HCV infection when receiving blood from an HCV-positive blood donor (Thomas, 2000). However, the risk of transfusion-transmission of HCV has been substantially reduced since the introduction of blood donor screening for anti-HCV antibodies in 1989. Certain groups are particularly at risk for parenteral transmission of HCV. Prior to blood donor screening, haemophiliac patients

and patients with immunoglobulin deficiency were at risk from contaminated batches of clotting factor and immunoglobulin, respectively (Heintges & Wands, 1997, Yap *et al.*, 1994). Health care workers are at risk through occupational exposure to blood and needle-stick injuries. Procedures such as tattooing, piercing and acupuncture are also associated with potential risk of HCV transmission. Illicit use of drugs, especially by injection, now accounts for the majority of HCV infections in developed countries. At least 60% of new HCV infections in the United States are due to intravenous drug use (IVDU) (Alter, 1999).

Non-parenteral Transmission of HCV

Cases of hepatitis C infection have been reported where there is no apparent evidence of parenteral exposure. Sexual transmission of HCV is infrequent, although co-infection with HIV seems to increase the risk of HCV sexual transmission (Lissen *et al.*, 1993). Perinatal transmission of HCV is uncommon and most likely occurs at the time of delivery. A correlation has been shown between the risk of vertical transmission and the HCV RNA titre of the mother (Ohto *et al.*, 1994).

1.7.2. Clinical features of HCV infection

The natural history of HCV infection has been difficult to assess because of the usually silent onset of the acute phase. Also, the interval between infection and the development of cirrhosis can exceed 30 years, which means there are few prospective studies.

1.7.2.1. Acute hepatitis C

The majority of patients with acute infection are symptom free but when symptoms are present, they develop an average of 7 weeks after exposure (Alter *et al.*, 1992). Symptoms usually consist of jaundice, malaise and nausea. The acute phase of HCV infection is characterised by elevated levels of serum alanine aminotransferase (ALT) released from damaged hepatocytes. HCV RNA becomes detectable by RT-PCR in

serum 1-3 weeks after exposure (Farci *et al.*, 1991) and anti-HCV antibodies appear within 3-20 weeks (Hoofnagle, 1997).

1.7.2.2. Chronic hepatitis C

Following acute infection it is estimated that only about 15-20% of patients will clear the virus. In the remaining 80%, HCV infection becomes chronic as demonstrated by the persistence of HCV RNA in serum. Serum ALT levels are abnormal or intermittently elevated in roughly two thirds of chronically-infected patients but remain normal in the other third (Theodore & Fried, 2000). Disease progression during chronic HCV infection is typically slow and less than 20% of patients develop cirrhosis of the liver within 20 years post infection (Marcellin, 1999). The rate of progression of fibrosis towards cirrhosis can be predicted by the severity of hepatic fibrosis, as diagnosed by liver biopsy and certain factors such as alcohol intake, co-infection with HIV or HBV and older age can accelerate the development of fibrosis (Lauer & Walker, 2001). HCV-related end-stage liver disease is now a leading cause of liver transplantation (Marcellin, 1999) and cirrhosis may lead to hepatocellular carcinoma (HCC) in 0.4-2.5% of patients (Colombo, 1999).

1.7.2.3. Extrahepatic HCV-associated diseases

A variety of extra-hepatic manifestations have been associated with HCV infection. Most of these are immune complex-related diseases and include mixed cryoglobulinaemia (aggregation of immunoglobulins in blood vessels) and membranoproliferative glomerulonephritis (deposition of immune complexes in the capillaries of the glomeruli) (Manns & Rambusch, 1999). Porphyria cutanea tarda (alteration of iron metabolism) has also been associated with HCV as 60-80% of patients in one study had anti-HCV antibodies (Gumber & Chopra, 1995).

1.7.3. Genotypes and natural history of HCV

The issue of whether pathogenicity of HCV infection varies according to genotype is controversial, as there are multiple other host and viral factors which could influence the disease. There have been reports that genotype 1b was associated with a greater frequency of cirrhosis and the development of HCC than other genotypes (Dusheiko *et al.*, 1994, Bruno *et al.*, 1997), while other studies have failed to show any relationship between disease severity and genotype (Benvegnu *et al.*, 1997, Brechot, 1997). It has also been shown that patients infected with genotype 1b are older than those infected with other genotypes (Zein *et al.*, 1996) and therefore more severe disease may be a reflection of a longer duration of infection rather than a more pathogenic form of HCV (Zein, 2000). A recent report has described an association between genotype 3 and steatosis, a condition in which lipids accumulate in the liver (Rubbia-Brandt *et al.*, 2001). The potential role of other risk factors such as obesity, alcohol abuse and IVDU was taken into account and the link was still found to be significant, although further studies await confirmation of this correlation.

1.8. Treatment and prevention of HCV

1.8.1. Interferon as a therapeutic agent for chronic hepatitis C

Before the identification of HCV in 1989, interferon (IFN) was shown to have beneficial effects against NANB hepatitis (Hoofnagle *et al.*, 1986) and had been used successfully for the treatment of malignancies (Kaposi's sarcoma, leukaemia) and viral infections due to HBV and human papillomavirus. Interferon alpha (IFNα) became the recommended therapy for chronic hepatitis C and a 6 month course of treatment led to normalisation of ALT levels, loss of detectable virus in blood and reduction of inflammation in liver biopsies in some patients. However, IFNα monotherapy showed a sustained response in a minority of patients and even when the duration of treatment was extended to 12

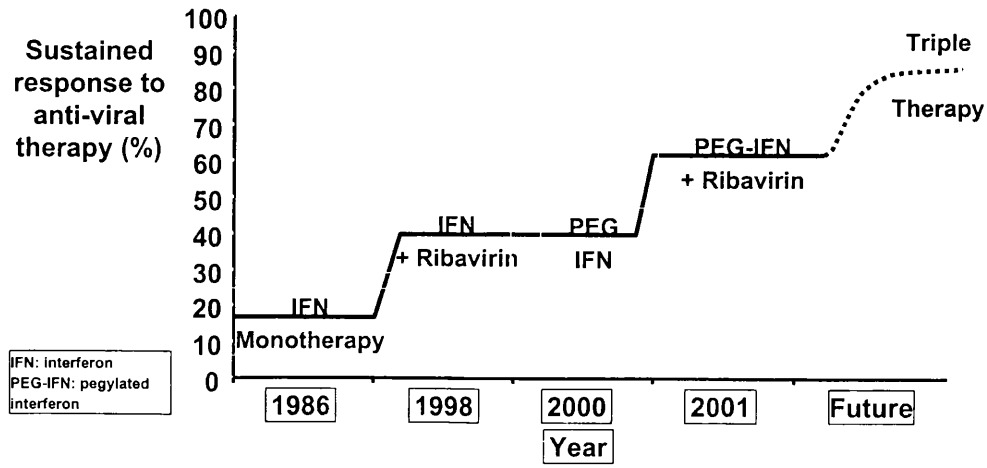


Figure 1.15. Illustration of the improving efficacy of interferon (IFN) therapy for chronic hepatitis C.

months, only 13-25% of patients showed a sustained response (Poynard *et al.*, 1996). More recently, the nucleoside analogue, ribavirin, has been used in combination with IFNα and this has increased the efficacy of treatment. When used alone, ribavirin leads to a reduction in serum ALT levels but has no effect on HCV viraemia (Di Bisceglie *et al.*, 1995), but when used in combination with IFNα, the response rates increase to 38-43% for 48 weeks of treatment (Poynard *et al.*, 1998, McHutchinson *et al.*, 1998). Combination therapy is also beneficial for patients who have relapsed after initial treatment with IFNα alone (Schalm *et al.*, 1997). IFNα and ribavirin are currently the only licensed treatments for HCV in the UK (Booth *et al.*, 2001). The efficacy of a pegylated form of IFN has recently been investigated. This involves attaching a large inactive molecule, polyethylene glycol, to the protein, which increases the half-life and allows larger doses to be given infrequently. Initial trials have shown that a combination of pegylated interferon and ribavirin leads to sustained response in approximately 60% of patients (Trepo *et al.*, 2000) (Figure 1.15).

1.8.2. Predicting response to interferon treatment

The high cost of IFN treatment and the associated side effects mean that the identification of markers that are predictive of treatment outcome would be of great importance. Both host and viral factors have been identified as being predictive of response to IFN treatment but it is often difficult to assess the effects of these factors independently and none is sufficiently accurate to exclude a patient from IFN treatment. Host factors that have been reported as predictive of response are sex and age, with younger women having a more favourable outcome of treatment than older men (Booth *et al.*, 2001). The better response in younger patients (<40 years) may be related to a shorter duration of infection and less severe liver disease. In fact, the absence of cirrhosis has been associated with an increased response to IFN therapy (Tsubota *et al.*, 1994, Jouet *et al.*, 1994).

Viral factors thought to be important in determining treatment response include genotype, level of viraemia and level of viral heterogeneity. Multivariant analyses have shown that genotype is an independent factor associated with treatment outcome, with genotypes 2

and 3 being predictive of a better response than genotype 1 (Davis & Lau, 1997, Trepo, 2000). In a trial using IFNα and ribavirin, 69% of non-genotype 1 infected patients showed sustained response after 24 weeks of treatment compared to 16% of patients infected with genotype 1 (McHutchinson *et al.*, 1998). Similar figures were reported by Poynard *et al.* (1998). If treatment was extended to 48 weeks, genotype 1 patients showed an increased response rate of around 30%, whereas the response rate of non-genotype 1 patients remained the same. Therefore it is recommended that patients infected with non HCV-1 (mostly genotypes 2 and 3) should be treated for only 6 months (Booth *et al.*, 2001). An inverse correlation between pre-treatment HCV RNA titre and response to IFN has been reported (Lau *et al.*, 1993, Trepo, 2000). In the trials reported by McHutchinson *et al.* (1998) and Poynard *et al.* (1998), patients with $<2x10^6$ copies/ml of viral RNA showed a 44% response rate, compared to 27% for those with $>2x10^6$ copies/ml. The degree of sequence heterogeneity has also been reported as predictive of IFN responsiveness (Kanazawa *et al.*, 1994). It has been suggested that this is because the greater the diversity of variants, the greater the chance that IFN-resistant species will be present or escape mutants will be generated that will survive IFN treatment.

1.8.3. Antiviral action of interferon

Interferons comprise a family of cytokines secreted in eukaryotic cells in response to various stimuli, particularly viral infection. Upon binding to their specific cell surface receptors on target cells, IFNs trigger a cascade of intracellular reactions that culminate in the induction of a large number of genes. The protein products of these genes mediate the multiple effects of IFN, which include induction of an antiviral state in the target cells, immune stimulation and regulation of cell growth (reviewed by Goodbourn *et al.*, 2000), all of which serve to block virus replication and target infected cells for apoptosis to limit virus spread. There are two types of IFN. Type I IFNs include IFNα and IFNβ and are produced by most cell types, in direct response to virus infection. Type II IFN consists of IFNγ, which is produced by activated T lymphocytes and natural killer cells in response to their recognition of virus-infected cells (Goodbourn *et al.*, 2000).

1.8.3.1. dsRNA-dependent protein kinase R (PKR)

Type I IFNs induce the synthesis of PKR, a serine/threonine kinase. PKR has an N-terminal regulatory domain that contains a dsRNA-binding site and a C-terminal catalytic domain. PKR is normally inactive and is activated on binding dsRNA. This induces a conformational change that triggers the catalytic activities of PKR. The active form of PKR is a dimer, with two PKR molecules binding one molecule of dsRNA. Once bound, PKR autophosphorylates and this renders the kinase domain active. Activated PKR phosphorylates the α subunit of the eukaryotic translation initiation factor eIF2α, which results in inhibition of protein synthesis. Viruses that produce dsRNA during the course of their replication cycles have the potential to activate PKR, which in turn leads to a block in viral replication at the level of protein synthesis. Several viruses have evolved mechanisms to block the effects of PKR (Gale & Katze, 1998). Influenza virus activates a cellular inhibitor of PKR, which binds PKR and inhibits kinase activity (Polyak *et al.*, 1996). Influenza virus also encodes a protein, NS1, which blocks the dsRNA-mediated activation of PKR, possibly by sequestering dsRNA (Lu *et al.*, 1995). The vaccinia virus E3L protein acts via a similar mechanism (Chang *et al.*, 1992). HIV, adenovirus and HSV also produce PKR inhibitors (Gale & Katz, 1998).

1.8.3.2. 2'-5' oligoadenylate synthetase / RNAse L pathway

Type I IFNs induce the synthesis of 2'-5' oligoadenylate (2-5(A)) synthetases. 2-5 (A) synthetases are activated by dsRNA and catalyse the synthesis of small oligomers, with a unique 2'-5' linkage, from precursor ATP. These products bind to endoribonuclease L (RNAse L) and induce its activation via dimerisation. Activated RNAse L cleaves ssRNA, including mRNA and thereby inhibits protein synthesis. EMCV, HSV and HIV have all evolved strategies to counteract this pathway (Goodbourn *et al.*, 2000).

1.8.3.3. Alternative antiviral pathways

The Mx proteins are also induced by IFNα. The human protein, MxA, inhibits the growth of several RNA viruses including members of the *Orthomyxoviridae*, *Paramyxoviridae*, *Rhabdoviridae* and *Bunyaviridae* (Pavlovic *et al.*, 1990, Zhao *et al.*, 1996, Frese *et al.*, 1996). Additional antiviral pathways also exist, as demonstrated by the generation of mice that are triply deficient in RNAse L, PKR and Mx1 (Zhou *et al.*, 1999). IFNα can also control viral infections indirectly, by activating the cellular immune system to destroy virus infected cells (Goodbourn *et al.*, 2000).

1.8.4. Mechanisms for HCV evasion of the interferon response

Some HCV-infected individuals do not respond to treatment and for patients with HCV genotype 1, over 50 % are resistant to IFNα therapy. This strongly suggests that HCV has evolved a mechanism or mechanisms to overcome the antiviral effects of IFNα. Two HCV proteins, NS5A and E2, have been implicated in mediating resistance to IFNα.

1.8.4.1. NS5A and ISDR

Enomoto *et al.* (1995) compared the full-length sequences of IFN-responsive and IFN-sensitive viruses from Japanese HCV genotype 1b-infected patients and identified a region within NS5A that corresponded with IFN responsiveness. This region was designated the interferon-sensitivity-determining region (ISDR) and patients who responded to IFN therapy were found to carry viruses that contained 4 or more amino acid residues that differed from the reference sequence HCV-J in this region. Many studies from North America and Europe subsequently reported conflicting data on the relationship between ISDR sequence and clinical resistance to IFN (Zeuzem *et al.*, 1997, Khorsi *et al.*, 1997, Hofgartner *et al.*, 1997). These apparently discrepant results were difficult to interpret due to viral strain differences and differences in study design or treatment regimen. A recent statistical analysis of all the published ISDR sequences in the database has

nonetheless demonstrated a positive association between ISDR sequence and IFN resistance (Witherell & Beineke, 2001).

A mechanism for IFN resistance was first proposed by Gale *et al.* (1997) when it was demonstrated that NS5A is capable of interacting with PKR and inhibiting its phosphorylation of eIF2α. Moreover, PKR and NS5A interact directly via the ISDR domain and mutating the ISDR sequence to match that seen in responders in the Enomoto study, blocks the interaction (Gale *et al.*, 1998). However, subsequent studies cast doubt on the role of the ISDR. Using a virus-rescue system, NS5A could rescue interferon-sensitive viruses from the antiviral effects of IFN in cell culture and the same effects were seen when using NS5A-deletion mutants lacking the ISDR domain (Polyak *et al.*, 1999). More recently, Podevin *et al.* (2001) have used NS5A derived from responder and non-responder patients and shown that NS5A can rescue IFN-sensitive viruses irrespective of the number of mutations in the ISDR or the phenotype of HCV from which the NS5A protein is derived. Therefore, while NS5A has been shown to block the antiviral effects of IFN, sequence differences in the ISDR may not be responsible for this activity. NS5A has recently been shown to induce the expression of interleukin-8 (IL-8), a cytokine that inhibits the antiviral action of IFN. Increasing levels of NS5A correlated directly with the increasing IL-8 concentrations and NS5A was shown to activate the promoter of the IL-8 gene (Polyak *et al.*, 2001a). IL-8 alone was able to mimic the effects of NS5A on rescuing IFN-sensitive viruses in cell culture. Additionally, HCV-infected patients who did not respond to IFN treatment were found to have raised levels of IL-8 (Polyak *et al.*, 2001b). This suggests that as well as conferring IFN-resistance via an interaction with PKR, NS5A may also be involved in an alternative pathway, mediated by IL-8. This may explain why the ISDR sequence is not always predictive of sensitivity/resistance to IFN.

1.8.4.2. E2 and PePHD

Taylor *et al.* (1999) identified a 12 amino acid sequence in the E2 protein that has homology to several putative autophosphorylation sites within the N-terminal domain of PKR, as well as its natural substrate, eIF2α. This sequence was termed the PKR- eIF2α

phosphorylation homology domain (PePHD). The PePHDs of the more IFN-resistant genotypes (1a and 1b) show greater homology to PKR and eIF2α than those of the less IFN-resistant genotypes (2 and 3). Further investigation demonstrated that genotype 1a E2 interacted with PKR and the interaction was dependent on the PePHD. E2 was also shown to inhibit PKR activity in yeast and HeLa cells and the PePHD was required for this inhibition. The exact mechanism by which E2 inhibits PKR is not known but it is not thought to be due to E2 acting as a pseudosubstrate for PKR autophosphorylation sites (Taylor *et al.*, 2001). Furthermore, PKR activity was not inhibited when the PePHD was mutated to resemble that of genotype 2 and 3 (Taylor *et al.*, 1999). These results indicated that the PePHD might be responsible for the inherent resistance of genotype 1 and sensitivity of genotypes 2 and 3. However, several clinical studies have shown that the PePHD is a highly conserved region and when mutations do occur, they are not predictive of response to IFN treatment for genotype 1 (Chayama *et al.*, 2000, Berg *et al.*, 2000, Polyak *et al.*, 2000) or resistance to IFN treatment for genotype 3 (Sarrazin *et al.*, 2000a, Puig-Basagoiti *et al.*, 2001). Therefore the role of E2 in mediating IFN resistance remains to be established.

1.8.5. Prevention of HCV infection

There is currently no vaccine available for the prevention of HCV infection. Difficulties in HCV vaccine development include the heterogeneous nature of HCV, not only with respect to genotypes/subtypes but also because of its quasispecies nature, which leads to rapid changes within an infected individual, possibly due to immune selection. There is also no evidence of lasting immunity in individuals who resolve the infection, leaving the potential for reinfection with the same or different HCV isolate. The two main areas of vaccine development have included recombinant subunit proteins and nucleic acids as candidate vaccines. Vaccination of chimpanzees with purified E1 and E2 proteins resulted in protection from challenge with homologous virus in 5/7 vaccinees (Choo *et al.*, 1994). However, the same chimpanzees were not protected from challenge with a heterologous HCV strain, implying the absence of cross-neutralising antibodies (Houghton, 2000). The potential use of DNA vaccines has also been explored. DNA

vaccines against other viruses (e.g. influenza, rabies, HSV1) successfully induce humoral and cellular immune responses (Ulmer *et al.*, 1993, Xiang *et al.*, 1994, Manickan *et al.*, 1995) and therefore potentially may result in both prevention of infection and immunotherapy of chronic infections. Immunisation of mice with naked plasmid DNA encoding HCV core can generate both humoral and lymphoproliferative immune responses (Lagging *et al.*, 1995, Inchauspe *et al.*, 1997). Plasmids encoding HCV E2 have also been used successfully to induce anti-E2 antibodies in immunised mice and macaques (Nakano *et al.*, 1997, Forns *et al.*, 1999). As with most areas of HCV research, the development of an HCV vaccine is hampered by the lack of a susceptible, small animal model.

With no vaccine available, the most effective means of reducing the incidence of HCV infection is the screening of all blood donations and educating susceptible individuals, particularly IVDUs and hospital staff, on safe practices.

1.9. Aims and objectives

The aim of this project was to amplify the HCV structural genes from the serum of a patient infected with HCV genotype 3, to express the genotype 3 glycoproteins and to compare their characteristics with those of genotype 1.

The following objectives were met:

1. To amplify the genotype 3 structural genes by RT-PCR and to obtain a majority sequence to compare with published genotype 3 and 1 sequences.
2. To express the genotype 3 E1 and E2 proteins and to investigate the reactivity of genotype 1 anti-E1 and anti-E2 antibodies against the genotype 3 proteins.
3. To compare the glycoproteins of genotypes 1 and 3 with respect to subcellular localisation, complex formation and glycosylation status.
4. To determine if genotype 3 E2 binds to human CD81, as reported for genotype 1 E2 (Pileri *et al.*, 1998) and to investigate the ability of HCV genotypes 1 and 3 virions in human sera to bind CD81.
5. To determine whether or not the PePHD within the E2 protein is predictive of response to interferon treatment in patients infected with HCV genotype 3. To investigate the proposed interaction of genotype 1 E2 with PKR (Taylor *et al.*, 1999).

Chapter 2

Materials and Methods

2.1. Materials

2.1.1. Bacterial Strains

***Escherichia coli (E. coli)* strain**	**Phenotype**
DH5α	F'/*end*A1 *hsd*R17 (r_k^-m+) *sup*E44 *thi*1 *rec*A1 *gyr*A (Nal^r) *rel*A1 Δ(*lac*ZYA-*arg*F) U169 (ϕ80*dlac*(*lac*Z)M15)
Top 10F' (Invitrogen)	F'{*lac*I^q, Tn*10*(Tet^R)} *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15Δ*lac*X74 *deo*R *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L(Str^R) *end*A1 *nup*G
Epicurian Coli XL1-Blue (Stratagene)	*rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac* [F' *pro*AB *lac*I^{q}ZΔM15 Tn*10* (Tet^r)]^c

2.1.2. Vectors

pSFV1 expression vector (Liljestrom & Garoff, 1991)	Life Technologies
pcDNA 3.1 (+)	Invitrogen
pGEM1	Promega

2.1.3. Synthetic oligonucleotides

Oligonucleotides were initially synthesised "in house" by Alex Orr and Dave McNab and later ordered from MWG-Biotech (Germany).

2.1.4. Kits and enzymes for DNA/protein modification

Recovery Quick Flow Mini kit, Recovery Quick Flow Midi kit, DNA purification kit	Hybaid
Restriction Enzymes, CIP, T4 DNA ligase, NTPs SP6 RNA polymerase, $m^7G(5')G$ RNA cap analogue, Endoglycosidase H, PNGase F	Roche
QIAmp RNA extraction kit, SuperFect transfection reagent, Omniscript reverse transcriptase	Qiagen
Superscript reverse transcriptase, RNase H	Life Technologies
RNasin	Promega
QuikChange Site-directed Mutagenesis kit	Stratagene
Advantage 2 PCR kit, Advantage TaqPlus PCR kit	Clontech

2.1.5. Mammalian cell lines and culture media

Proteins were expressed in baby hamster kidney (BHK) (clone C13) cells and human hepatocellular carcinoma (Huh7) cells.

Glasgow minimal Eagles Medium (GMEM), Dulbecco's modified Eagles Medium (DMEM), new born calf serum (NBCS), foetal calf serum (FBS),

penicillin/streptomycin, L-glutamine, non-essential amino acids and S.O.C medium were supplied by Life Technologies.

PBS, PBS(A), versene, trypsin, L-broth, 2YT broth, tryptose phosphate (TP) broth, and components of Eagles medium used during radiolabelling were produced "in-house" by the media department.

2.1.6. Human sera

Human sera used in this study were stored in the Regional Virus Lab at Gartnavel General Hospital, Glasgow. Informed consent for the use of the sera was obtained. Sera were separated within 4 hours of sampling by Carol-Anne Smith. Sera were stored in small aliquots at -20°C or in liquid nitrogen.

2.1.7. Radiochemicals

^{35}S-L-Methionine	10μCi/μl	Amersham

2.1.8. Antibodies

Monoclonal and polyclonal antibodies raised against HCV E1 and E2 (see Table 4.2.), anti-GST polyclonal	A. Patel
Conformational anti-E2 MAbs H30, H50, H53, H60 H61.	J. Dubuisson
Anti-calnexin polyclonal antibody	J. McLauchlan
Anti-calnexin monoclonal antibody	StressGen Biotechnologies Corp.
Anti-PERK polyclonal antibody (N-18),	Santa Cruz,
Anti-PKR monoclonal antibody (B-10)	Biotechnology.
Anti-PKR monoclonal antibody (71/10)	A. Hovenessian

Anti-mouse FITC raised in rabbit	Dako Ltd.
Anti-rabbit HRP	Diagnostics Scotland
Anti-mouse FITC raised in goat, anti-goat Cy3 raised in rabbit, anti-mouse Cy5 raised in goat, anti-rabbit TRITC raised in goat, anti-mouse HRP, protein-A-HRP, rabbit anti-mouse IgG.	Sigma

2.1.9. Chemicals

Unless otherwise stated, all chemicals were obtained from Sigma Chemical Co., BDH Chemicals Ltd or Prolabo.

ECL, Hybond nitrocellulose membrane	Amersham
30% acrylamide, ammonium persulphate	Bio-Rad
En^{3}Hance	DuPont
Dried skimmed milk	Marvel
Viraferon (interferon alpha-2b)	SP(Brinny) Co., Ireland
TMB substrate	Zymed Laboratories

2.1.10. cDNA clones and proteins

The pSFV1 clones of core, E1, E2 and E2 alone from strains H77 and Glasgow were provided by J. Patel. The pcDNA clone encoding E1 and E2 of strain H77 was provided by A. Patel. The extracellular loop 2 of human CD81, expressed as a GST fusion protein, was provided by A. Patel.

2.1.11. Solutions

Agarose gel loading buffer	0.1M EDTA, 50% sucrose, 1μg/ml bromophenol blue.
Boiling mix	29% SGB, 6% (w/v) SDS, 2M β-ME (excluded for non-reducing conditions), 29% glycerol, 1μg/ml bromophenol blue.
Cell Lysis solution	200mM NaOH, 1% SDS
Citrate phosphate buffer	0.1M citric acid / 0.2M Na_2HPO_4 (pH 5.5).
Eagles A	1.5mM $CaCl_2.2H_2O$, 1mM $MgSO_4.7H_2O$, 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 meth)	7:1 (v/v) Eagles A: Eagles B (1/5 normal concentration of methionine), 2% (v/v) acetic acid in dH_2O.
Elution solution	1250mM NaCl, 100mM Tris-HCl (pH 8.5).
Equilibration solution	600mM NaCl, 100mM sodium acetate / acetic acid (pH 5.0), 0.15% Triton X-100.
Fix solution	50% methanol, 7% acetic acid in dH_2O.
Immunoprecipitation (IP) buffer	50mM Tris-HCl (pH 7.6), 200mM NaCl, 1mM EDTA, 1% Triton X-100, 1mM PMSF, 20mM NEM.
L-broth	10g NaCl, 10g Bactopeptone and 5g yeast extract per litre.
L-broth agar	L-broth plus 1.5% (w/v) agar.
Lysis buffer	20mM Tris-HCl (pH 7.5), 1mM EDTA, 20mM iodoacetamide, 100mM NaCl, 0.5% Triton X-100.
Neutralisation solution	3.2M potassium acetate / acetic acid (pH 5.5).

PBS	PBS(A) plus 6.8mM $CaCl_2.2H_2O$ and 4mM $MgCl_2.6H_2O$.
PBS(A)	170mM NaCl, 3.4mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, 25mM Tris-HCl (pH 7.2).
PBST	PBS(A) plus 0.05% (v/v) Tween 20.
Phosphate buffer	0.2M Na_2HPO_4 / 0.2M $NaH_2PO_4.2H_2O$ (pH 7.5).
Resolving gel buffer	1.5M Tris-HCl (pH 8.9), 0.4% SDS.
Resuspension solution	50mM Tris-HCl (pH 8.0), 10mM EDTA, 100µg/ml RNase.
Running gel buffer	40mM Tris, 185mM glycine, 0.1% SDS.
Stacking gel buffer	0.5M Tris-HCl (pH 6.8), 0.4% SDS.
TBE (1X)	90mM Tris-HCl (pH 8.0), 90mM Boric Acid, 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 solution	800mM NaCl, 100mM sodium acetate / acetic acid (pH 5.0).

2.1.12. Other materials and apparatus

X-Omat UV film was purchased from Kodak Ltd., as was developer and fixer for the KONICA. All tissue culture plasticware was supplied by Nunc. The immulon ELISA plates were purchased from Dynex.

2.2. Manipulation of DNA and RNA

2.2.1. Purification of synthetic oligonucleotides

At the start of the project, oligonucleotides were synthesised "in-house" and were supplied with the DNA molecule covalently linked to a support matrix on a column. Oligonucleotides were removed from this column by pushing 1.5ml of ammonia through the column between two 5ml syringes at a rate of 0.5ml every 20 minutes. The solution was mixed by pushing it back and forth through the column several times and was then transferred to a screw cap Eppendorf tube. To remove the amide protecting groups, the oligonucleotide was deprotected by incubating the solution at 55°C for 5 hours and then cooled to room temperature. The oligonucleotide was dried under vacuum overnight to remove the ammonia and the pellet resuspended in 50µl deionised water. Oligonucleotides were stored at -20°C until required. In the latter stages of the project oligonucleotides were ordered from MWG-Biotech (Germany) and were supplied as lyophilised pellets.

2.2.2. Small scale preparation of plasmid DNA (minipreps)

The extraction of plasmid DNA from culture was performed using the Hybaid Recovery Quick Flow Mini Kit according to the manufacturers instructions. Briefly, single colonies of transformed bacteria were inoculated into 5ml of L-broth containing 100µg/ml ampicillin and incubated with shaking overnight at 37°C. A 1.5ml aliquot of culture was centrifuged at 13000rpm for 30sec, the supernatant removed and the cell pellet resuspended in 50µl Pre-Lysis solution. This was followed by the addition of 100µl of Cell Lysis solution, which was mixed thoroughly until the solution was clear and viscous. 75µl of Neutralising solution was added before centrifugation for 3 min at 13000rpm. The supernatant was transferred to a Hybaid spin filter, 250µl of Binding buffer was added and the mixture centrifuged for 1 min before the addition of 350µl of Wash solution and a further 1min centrifugation. The pellet was dried by a final 1min centrifugation and the DNA eluted by the addition of 50µl dH_2O and a 30sec centrifugation. The plasmid DNA was stored at -20°C.

2.2.3. Large scale preparation of plasmid DNA (midipreps)

Large quantities of plasmid DNA were isolated from cells using the Hybaid Recovery Quick Flow Midi Kit as specified by the manufacturers instructions. 100ml of L-broth containing 100μg/ml ampicillin was inoculated with bacteria and incubated with shaking overnight at 37°C. Cells were pelleted by centrifugation at 3000rpm for 20 min at room temperature. The supernatant was removed completely and the cell pellet resuspended in 4ml of cold Resuspension solution (containing RNase). 4ml of Cell Lysis solution was added and mixed by inversion before leaving at room temperature for no longer than 5 min. 4ml of Neutralisation solution was added, immediately mixed by inversion and the cell debris removed by centrifugation at 12000rpm for 30min at 20°C. The Hybaid columns were equilibrated with 10ml of Equilibration solution prior to the addition of the supernatant which was allowed to flow through the column by gravity. Two 10ml aliquots of Wash solution were applied to the column and the DNA was eluted by the addition of 5ml of Elution solution. The DNA solution was mixed with 3.5ml of isopropanol and the DNA precipitated by centrifugation at 12000rpm for 1 hour at 4°C. The DNA pellet was washed with 70% ethanol and centrifuged at 12000rpm for 30 min at 4°C. Finally, the DNA was air-dried and resuspended in 100μl of dH_2O and stored at -20°C.

2.2.4. Quantification of plasmid DNA and oligonucleotides

The Heλiosα version 4.55 spectrophotometer was used to determine DNA concentrations by measuring the absorbance at 260nm, assuming 1 A_{260} equals 50μg/ml for double stranded plasmid DNA and 1 A_{260} equals 20μg/ml for single-stranded oligonucleotides. Purity of nucleic acids was determined by readings at 260nm and 280nm (A_{260}/A_{280}), where a ratio of 1.8 indicated that the preparation was relatively free of protein contaminants.

2.2.5. Restriction enzyme digestion of DNA

Restriction enzyme digestions and conditions as outlined by Sambrook *et al.* (1989) were adhered to, or adapted as required. All reactions were incubated at 37°C (or the temperature specified by the supplier) for 2-3 hours in a total volume of 20µl which contained 10U of restriction enzyme per 1µg of DNA in the buffer supplied for the enzyme. Digestion reactions were terminated by addition of ¼ volume of agarose gel loading buffer and resulting DNA fragments were analysed by agarose gel electrophoresis as described in 2.2.10.1.

2.2.6. Dephosphorylation of linearised plasmid DNA

Removal of the 5' terminal phosphates of digested vectors was carried out using calf intestinal phosphatase (CIP). In a total volume of 50µl, the DNA was added to 5µl of 10x CIP buffer (Roche) and 1 unit of CIP per 5µg of DNA. The reaction was incubated at 37°C for 30 min and then another unit of CIP per 5µg of DNA was added and incubation continued for a further 30 min. The linearised, dephosphorylated vector was purified as described in 2.2.10.3 in preparation for ligation.

2.2.7. Extraction of RNA from human sera

RNA was extracted from human sera using the Qiagen QIAmp kit, as specified by manufacterers instructions. For lysis of virus particles, a 140µl aliquot of serum was combined with 560µl of AVL buffer containing carrier RNA and mixed by vortexing, then centrifugation for 1 min at 8000rpm. The mixture was incubated at room temperature for 10 min before the addition of 560µl of ethanol (100%) and further vortexing. A 630µl aliquot of the solution was added to a QIAmp spin column and centrifuged for 1 min at 8000 rpm. The filtrate was discarded and the remaining 630µl applied to the column and the centrifugation step repeated. The filter was washed with 500µl of buffer AW1, centrifuged at 1 min at 8000rpm and the filtrate discarded. This wash step was repeated with 500 µl of buffer AW2 and centrifugation at 14000rpm for 3 min. The RNA was eluted by the addition of 60µl buffer AVE,

incubation at room temperature for 1 min and centrifugation at 8000rpm for 1 min. The viral RNA was stored at -20°C.

2.2.8. Reverse transcription of RNA

2.2.8.1. Superscript reverse transcriptase

Reverse transcription was performed using SUPERSCRIPT reverse transcriptase from Life Technologies. In a total volume of 12.5µl, 40 pmoles of primer (MS2) was added to 3µl of viral RNA. Primer annealing occurred by heating to 70°C for 10 min and cooling to 4°C. This was added to a reaction mix containing 50mM Tris-HCl (pH 8.3), 7.5mM KCl, 3mM $MgCl_2$, 20 units of RNAsin, 10mM DTT, 1mM each of dGTP, dATP, dTTP and dCTP, and 200U of SUPERSCRIPT reverse transcriptase to make a total reaction volume of 20µl. This was heated to 42°C for 50 min and 70°C for 15 min before cooling to 4°C. To remove RNA complementary to the cDNA, 2 units of Ribonuclease H (RNAse H) was added and the reaction incubated at 37°C for 20 min. cDNA was stored at 4°C.

2.2.8.2. Omniscript reverse transcriptase

For the purposes of quantification, Omniscript reverse transcriptase (Qiagen) was used for the reverse transcription reaction. Briefly, 3 µl of RNA was added to 9.7 µl water and incubated at 65°C for 5 min then cooled to 4°C. To this was added a mix containing 2 µl of 10X RT buffer (Qiagen), 2 µl dNTP mix (5mM each dNTP), 40 pmoles RT primer, 12U RNasin and 4U Omniscript reverse transcriptase. This mix was incubated at 37°C for 60 min and then 93°C for 5 min before cooling to 4°C

2.2.9. Polymerase chain reaction (PCR) amplification of cDNA

Amplification of HCV DNA was performed using "nested" PCR to increase the specificity and amount of DNA amplified. Two sets of primers were used, one set

being internal to the other. The products from the first round of amplification were used as template for a second round of amplification using the internal primers.

2.2.9.1 Amplification of core, E1 and E2

Both first and second round PCR reactions were performed using the Clontech Advantage 2 PCR kit which contained a polymerase mix which included AdvanTaq DNA polymerase, a small amount of a proofreading polymerase and TaqStart antibody to provide automatic hot-start PCR. For the first round PCR reaction, 4µl of cDNA was added to a 16µl reaction mixture containing 40mM Tricine-KOH (pH 9.2), 15mM KOAc, 3.5mM $Mg(OAc)_2$, 3.75µg/ml BSA, 0.005% Tween20, 0.005% NP-40, 0.2mM each of dATP, dCTP, dGTP and dTTP, 16 pmoles of each primer and 0.4µl of the polymerase mix. PCR reactions were carried out under the following reaction conditions: a denaturation step of 94°C for 4 min and then thermal cycling of 94°C for 90 sec (strand separation), 65°C for 90 sec (primer annealing) and 70°C for 3 min (strand elongation) for 25 cycles before cooling to 10°C. For the second round PCR reaction, 1µl of the first round product was added to a 19µl reaction mixture containing the same components as for the first round reaction except, this time using the internal set of primers (16 pmoles of each). As for the first round, the reaction was heated to 94°C for 4 min and then underwent thermal cycling of 94°C for 90 sec, 60-65°C (dependent on primer) for 90 sec and 70°C for 3 min for 25 cycles. Finally, the reaction was heated to 70°C for 7 min for a final extension step and cooled to 10°C. For purification of the PCR product, the second round PCR reaction volume was increased to 50µl. All PCR reactions were carried out in a Biometra TRIO-Thermoblock.

2.2.9.2 Amplification of the PePHD region of E2

Extraction of RNA and reverse transcription were performed as described in 2.2.7 and 2.2.8.1. The PCR reactions were performed using the AdvanTaq Plus PCR kit from CLONTECH. This kit used a polymerase mix which contained AdvanTaq DNA polymerase and TaqStart antibody for automatic hot start PCR. For the first round PCR reaction, 4µl of cDNA was added to a 16µl reaction mixture containing 40mM

Tricine-KOH (pH 8.0), 16mM KCl, 3.5mM $MgCl_2$, 3.75µg/ml BSA, 0.2mM each of dATP, dCTP, dGTP and dTTP, 16 pmoles of each primer and 0.4µl of the polymerase mix. Following a 4 min denaturation step at 94°C, the reactions underwent thermal cycling of 94°C for 90 sec, 65°C for 90 sec and 68°C for 3 min for 30 cycles. For the second round PCR reaction, 1µl of the first round product was added to a 19µl reaction mixture containing the same components as for the first round reaction except, this time using the internal set of primers (16 pmoles of each). The second round cycling conditions were the same as for the first round except only 25 cycles were used and in addition, there was a final extension at 68°C for 7 min. For purification of the PCR product, the second round PCR reaction volume was increased to 50µl. All PCR reactions were carried out in a Biometra TRIO-Thermoblock.

2.2.9.3 Amplification of HCV cDNA for quantification purposes

For quantification purposes the cDNA was amplified by PCR using a Perkin Elmer Applied Biosystems 5700 sequence detection system and all PCR reagents and primers were obtained from Applied Biosystems. 1 µl of cDNA was added to a mix containing 12.5 µl of PCR mastermix (AmpliTaqGold DNA polymerase and dNTPs), 7.091 pmoles of EMC HCV F primer, 1.225 pmoles of EMC HCV R primer and 5.1 pmoles of EMCMGBP fluorescent probe. The cDNA was amplified with the following thermal cycling program: 50°C for 2 min, 95°C for 10 min and 40 cycles of 95°C for 15 sec and 60°C for 1 min. All reactions were quantified against a HCV positive serum of known titre to generate a standard curve.

2.2.10. Electrophoretic separation and isolation of DNA

2.2.10.1. Agarose gel electrophoresis

DNA fragments produced by PCR or restriction enzyme digestion were resolved by agarose gel electrophoresis. A 1% (w/v) agarose gel was prepared in 1x TBE buffer and contained a final concentration of 0.05µg/ml of ethidium bromide. Agarose gel loading buffer was added to each DNA sample prior to electrophoresis at 60-100V in

1x TBE buffer. An appropriately sized molecular weight marker (Roche) was included for comparison. Following electrophoresis, DNA was visualised under short-wave UV light or, for preparative gels, under long-wave UV light. Photography was carried out using the Bio-Rad Gel Doc 2000 Imaging system and accompanying software.

2.2.10.2. Purification of DNA from agarose gels

Agarose slices containing appropriate DNA fragments were excised from gels under long-wave UV transillumination. The DNA was recovered using the Hybaid Recovery DNA Purification Kit II, which uses a silica gel based system. The binding buffer containing the silica gel matrix was added to a spin column. The excised gel slice was added and incubated at 55°C until the gel had dissolved. The solution was applied to the column by centrifugation at 13000 rpm for 30 sec. DNA was washed with 500 µl Wash solution (containing ethanol) and the centrifugation step repeated. The pellet underwent further centrifugation for 1 min and the DNA was eluted by the addition of 10-20 µl DNAse-free water, gentle vortexing and a final centrifugation for 30 sec.

2.2.10.3. Purification of linearised DNA and digested PCR fragments

Linearised plasmid DNA and PCR fragments were purified from solution using the Hybaid Recovery DNA Purification Kit II. The restriction enzyme digestion was mixed with binding buffer (provided by the supplier) which contains a silica gel matrix and added to a spin column. The solution was applied to the column by centrifugation at 13000 rpm for 30 sec. Thereafter, the procedure was the same as described above in 2.2.10.2.

2.2.10.4. Phenol/Chloroform extraction

Proteins were removed from bulk restriction enzyme digestions (containing more than 5µg DNA) by the addition of an equal volume of a 25:24:1 solution of phenol:chloroform:isoamyl alcohol (Sigma). This mixture was mixed vigorously and

centrifuged for 5 min at 13000rpm. This resulted in the organic layer containing the protein, being separated from the aqueous layer containing the nucleic acid. The aqueous layer was transferred to an equal volume of a 24:1 solution of chloroform:isoamyl alcohol to remove any residual phenol and the centrifugation step repeated. Finally, the aqueous layer was removed and the nucleic acid concentrated by ethanol precipitation as described in 2.2.10.5.

2.2.10.5. Ethanol precipitation

Nucleic acids were precipitated by the addition of 2 ½ volumes of ethanol and 1/10 volume of 3M sodium acetate. After mixing, the solution was placed on dry ice for 15-30 min and the precipitated nucleic acids collected by centrifugation at 13000rpm for 10 min. The pellet was washed with 70% ethanol and centrifuged at 13000rpm for 3 min. The pellet was then air-dried and resuspended in 10-50µl of dH_2O.

2.2.11. Ligation reactions

Vector and "insert" DNA fragments, prepared by restriction enzyme digestion and purification, were ligated in a 1:3 molar ratio. In a total volume of 10µl, the vector and insert DNA fragments were combined with 1x ligase buffer and 10 units of T4 DNA ligase. Reactions were incubated at 25°C overnight before transformation of competent *E.coli* as described in 2.4.

2.2.12. *In vitro* transcription of linearised DNA

Recombinant pSFV1 constructs were linearised with *Spe*I and purified by phenol/chloroform extraction and ethanol precipitation. 1µg of this linearised template was added to a reaction mixture containing transcription buffer (40mM Tris-HCl (pH 8.0), 6mM $MgCl_2$, 2mM spermidine and 10mM DTT), NTP mix (1mM ATP, UTP and CTP, 0.5mM GTP), 1mM $m^7G(5')G$ RNA cap analogue, 1mM DTT, 50 units ribonuclease inhibitor (RNasin) and 40 units of SP6 RNA polymerase (as described by Liljestrom & Garoff, 1991). The reaction was incubated at 37°C for 2 hours and 2µl

of the reaction was analysed on a 1% agarose gel. The remaining RNA was used for electroporation into mammalian cells as described in 2.6.1.

2.2.13. Automated DNA sequencing

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

2.2.14. Site-directed mutagenesis

Site-directed mutagenesis was performed using the QuikChange Site-Directed Mutagenesis kit from Stratagene. A set of PCR primers, complementary to opposite strands of the plasmid and incorporating the mutation, were designed for each desired point mutation. All changes were made consecutively, apart from two, which were located close together and therefore done simultaneously.

The PCR reaction was carried out in a total volume of 50µl, including 10ng of template plasmid, 125ng of each primer, 1x buffer (10mM KCl, 10mM $(NH_4)_2SO_4$, 20mM Tris-HCl (pH 8.8), 2mM $MgSO_4$, 0.1% Triton X-100, 0.1mg/ml BSA), 0.2mM each of dATP, dCTP, dGTP and dTTP and 2.5 units of *PfuTurbo* DNA polymerase. The reactions were overlaid with mineral oil. PCR reactions were carried out as follows: denaturation at 95°C for 30 sec and then thermal cycling of 95°C for 30 sec, 55°C for 1 min and 68°C for 12 min for a total of 12 cycles before cooling to 4°C. PCR cycling conditions were optimised for each set of primers, usually by increasing the annealing temperature to 60°C or increasing the number of cycles to 18 or 24. Following the PCR, 10µl of the reaction was analysed on a 1% agarose gel to check that the PCR product was of the correct size. The remainder of the reaction was incubated with 10 units of *Dpn*I (to remove the parental plasmid) for 1 hour at 37°C. 1µl of this reaction was then used to transform Epicurian Coli XL1-Blue supercompetent cells. Briefly, the cells were thawed on ice, the DNA added and the cells placed on ice for 30 min. The cells were then incubated at 42°C for 45 sec and placed on ice for a further 2 min before the addition of 500µl S.O.C medium and

incubation at 37°C for 1 hour with shaking. 500µl of the culture was plated onto L-broth agar plates containing 100µg/ml ampicillin and incubated overnight at 37°C. Single colonies were inoculated into 5ml of L-broth (containing 100µg/ml ampicillin) and incubated over night at 37°C with shaking. Plasmid DNA was isolated from the culture as described in 2.2.2. To ensure that no nicked plasmid DNA was transferred to the next mutagenesis reaction, the plasmid DNA was separated on a 1% agarose gel and the supercoiled plasmid DNA was cut out of the gel and extracted as described in 2.2.10.2. This DNA was used as a template for the next PCR reaction with a new set of mutagenesis primers.

2.3. Preparation of electrocompetent *E.coli* (DH5α) cells

To prepare electrocompetent bacteria, a 5ml overnight culture of *E.coli* strain DH5α was inoculated into 2 x 400ml 2YT broth. The cells were grown for 3-4 hours at 37°C in a shaking incubator until the OD_{550} was approximately 0.6. The culture was chilled on ice for 15-30 min before centrifugation at 3000rpm for 10 min at 4°C. The cells were gently resuspended in 200ml ice-cold deionised water, pelleted as before and resuspended again in 100ml water. Following a final centrifugation step, the cells were resuspended in 2ml 10% glycerol (in deionised water), aliquoted into 40µl volumes, frozen on dry ice and stored at -70°C.

2.4. Transformation of competent *E.coli* cells

2.4.1. Electroporation of DH5α cells

A 40µl aliquot of electrocompetent *E.coli* DH5α was thawed on ice before the addition of 2µl of the ligation reaction. This mixture was transferred to a chilled electroporation cuvette (0.1cm gap, Bio-Rad) and electroporated by a single pulse at 1800V in a Hybaid Cell Shock apparatus. The cells were resuspended in 1ml of S.O.C. medium and incubated at 37°C for 1 hour to allow expression of the antibiotic resistance gene. One tenth of the culture was plated onto a L-broth agar plate containing 100µg/ml ampicillin and incubated at 37°C overnight.

2.4.2. Transformation of TOP10F cells

Occasionally, commercially available chemically competent TOP10F cells (Invitrogen) were used to obtain plasmid DNA. A 50µl aliquot of cells was thawed on ice before the addition of 2µl 0.5M β-mercapto ethanol and 2µl of the ligation reaction. This mixture was chilled on ice for 30 min and then incubated at 42°C for 45 seconds to "heat shock" the bacteria and then chilled on ice for another 2 min prior to the addition of 950µl S.O.C. medium. The cells were incubated at 37°C for 1 hour in a shaking incubator and then plated onto agar plates as described before.

2.5. Maintenance of mammalian cells

2.5.1. BHK cells

BHK cells were cultured in Glasgow minimal Eagles Medium (GMEM) containing 10% new born calf serum (NBCS), 4% tryptose phosphate (TP) broth and 100 units/ml of penicillin/streptomycin at 37°C with 5% CO_2 in 160cm^2 tissue culture flasks. Cells were harvested at confluency, with a yield of 2-4x10^7 cells per flask.

2.5.2. Huh7 cells

Huh7 cells were cultured in Dulbecco's modified Eagles Medium (DMEM) containing 10% foetal calf serum (FCS), 4mM L-glutamine, 1X non-essential amino acids and 100 units/ml of penicillin/streptomycin at 37°C with 5% CO_2 in 160cm^2 tissue culture flasks. Cells were harvested at confluency, with a yield of 2-4x10^7 cells per flask.

2.6. Transfection of mammalian cells

2.6.1. Electroporation of recombinant SFV RNA

Cells approaching confluency in a 160cm^2 flask were washed with versene at room temperature and removed by trypsin treatment. The detached cells were suspended in 10ml of cell medium and a single cell suspension was achieved by passing the cells

through a narrow hole 5ml pipette several times. The cells were pelleted by centrifugation at 600rpm for 5 min at room temperature, resuspended in 30 ml of PBS(A), pelleted again and resuspended in a final volume of PBS(A) to give a concentration of ~$1x10^7$ cells/ml. A 0.8ml aliquot of cells was placed in an electroporation cuvette (0.4cm gap, Bio-Rad) with the *in vitro* transcribed recombinant pSFV1 RNA (2.2.12). The RNA/cell suspension was mixed by inversion and electroporated using the Bio-Rad Gene Pulser II apparatus. For BHK cells, the electroporation conditions were two pulses (with inversion between pulses) at 1.2kV and 25μF at room temperature. For Huh7 cells the electroporation conditions were one pulse at 360V and 950μF at room temperature. The electroporated cells were transferred to 6ml of cell medium. Finally, 35mm tissue culture dishes were seeded with $2x10^6$ cells (in 3ml medium) and 13mm coverslips in 24-well plates were seeded with $1x10^5$ cells (in 1ml medium). Cells were incubated at 37°C.

2.6.2. Transfection of pcDNA constructs

SuperFect transfection reagent (Qiagen) was used for transfection of Huh7 cells with recombinant pcDNA constructs. On the day prior to transfection, 2 x 10^5 cells were seeded onto 35mm plates or 5 x 10^4 cells onto coverslips in a 24-well plate. On the day of transfection the cells had grown to approximately 60% confluency. A pre-determined amount of DNA (approximately 4μg per 35mm dish, depending on construct) was added to 100μl of cell medium lacking FCS and antibiotics. The mixture was mixed thoroughly, incubated for 2-5 min at room temperature and mixed again by centrifugation at 13000rpm for a few seconds. A pre-determined amount of SuperFect transfection reagent (approximately 8μl per 35mm dish) was added and mixed by pipetting. The mixture was incubated at room temperature for 5-10 min, during which time, the cells were washed once with PBS. 600μl of normal cell medium was added to each transfection reaction and the mixture added to the cell monolayer. The cells were incubated at 37°C for 2-3 hours and were then washed once with PBS before normal cell medium was added and incubation continued for 40-45 hours to allow for gene expression. When transfecting cells on 13mm coverslips, the relative volumes of transfection reagent and medium were reduced as specified by the manufacturers instructions.

2.7. Radiolabelling of cells with ^{35}S-methionine

Cells were radiolabelled with ^{35}S-methionine 6-8 hours post electroporation. The medium was replaced with Eagles medium containing 1/5 normal concentration of methionine [Eagles (low meth)]. The cells were then incubated for a further 30 min at 37°C before the medium was replaced with Eagles (low meth) containing 30μCi/ml ^{35}S-methionine. Cells were radiolabelled for 14 hours at 37°C before harvesting.

2.8. Treatment of Huh7 cells with interferon-α

Cells were treated with Viraferon (interferon alpha-2b) to induce protein kinase R (PKR) expression. The medium was removed and replaced with Huh7 cell medium containing 1000 IU/ml interferon-α. The cells were incubated for 18 hours at 37°C before harvesting.

2.9. Preparation of cell extracts

2.9.1. Western blot analysis

Cells were washed twice with PBS and lysed in 250μl of boiling mix for 10 min at room temperature. The cell extract was collected into an Eppendorf and heated to 100°C for 5 min. Cell extracts were stored at –20°C.

2.9.2. ELISA analysis

Cells were washed twice with PBS and lysed in 500μl of lysis buffer for 30 min at 4°C. The extracts were collected and the cell debris removed by centrifugation at 13000rpm for 5 min. The supernatant was stored at 4°C.

2.9.3. Immunoprecipitation analysis

Cells were washed twice in PBS containing 20mM N-ethylmaleimide (NEM), an alkylating agent which blocks free sulphydryl groups on cysteine residues. Cells were lysed in 500µl of immunoprecipitation buffer (also containing 20mM NEM) for 10 min at room temperature. The extracts were collected and the cell debris removed by centrifugation at 13000rpm for 5 min. The proteins of interest were isolated from the extracts according to the procedure in 2.10.

2.10. Immunoprecipitation

For immunoprecipitation, 1-5µl of the appropriate antibody was added to cell lysates (prepared as described in 2.9.3) and the mixture incubated at 4°C overnight. Protein-A sepharose (7.5mg per precipitation) was equilibrated overnight in IP buffer at 4°C. A bridging antibody, rabbit anti-mouse antibody (5µl per precipitation) was added and allowed to bind to the protein-A sepharose overnight. The equilibrated protein-A resin was washed three times with IP buffer and resuspended in 100Xµl (where X was the number of precipitations) to achieve a 50% slurry. 100µl of this protein-A resin was added to each precipitaion and protein-A-antibody binding was allowed to occur for 1-2 hours at 4°C on a rotator. The protein complexes were subsequently pelleted by brief centrifugation at 6000rpm and washed three times with 500µl of IP buffer. Bound proteins were eluted with 40µl of boiling mix.

2.11. Deglycosylation of proteins using Endo H and PNGase F

Proteins were immunoprecipitated as described above but were eluted and denatured with 0.5% SDS and 1% β-ME at 100°C for 15 min. The solution was cooled on ice and the sepharose beads pelleted by brief centrifugation. For endoglycosidase H treatment, citrate phosphate buffer (pH 5.5) was added to a final concentration of 50mM, followed by 0.01 units of endo H. For PNGase F treatment, phosphate buffer (pH 7.5) (final concentration of 50mM), NP-40 (1% final concentration) and EDTA (final concentration of 2mM) were added, followed by 0.125 units of PNGase F.

Reactions were incubated at 37°C for 2 hours. Finally, boiling mix was added and the deglycosylated proteins analysed by SDS-PAGE as described in 2.12.

2.12. Protein analysis by SDS-PAGE

Proteins were separated by electrophoresis through SDS polyacrylamide gels using the discontinuous buffer system (Laemmli, 1970) and the Bio-Rad Miniprotein II apparatus. The apparatus and glass plates (10 x 8cm) were assembled according to the manufacturers instructions using 1.5mm wide spacers. A 10% acrylamide resolving gel mix was used containing 3.6ml of 30% acrylamide solution (consisting of acrylamide and *N',N'*methylene bisacrylamide (ratio 37.5:1), 2.7ml of resolving gel buffer, 4.5ml of dH_2O, 70µl of 10% ammonium persulphate (APS) and 5µl of TEMED. This solution was immediately poured into the gel assembly to approximately 2cm from the top of the plates. The resolving gel was overlaid with butan-2-ol to leave a smooth interface after polymerisation. After polymerisation, the top surface of the resolving gel was washed with resolving gel buffer. The stacking gel was prepared containing 1ml of 30% acrylamide solution, 1.5ml of stacking gel buffer, 3.5ml dH_2O, 50µl 10% APS and 5µl TEMED. This was poured on top of the set resolving gel surface and a 1.5mm wide comb inserted to form loading wells for the protein samples. Following polymerisation, the comb was removed and the wells were washed with running buffer. The gel was placed in the electrophoresis apparatus and the reservoirs filled with running buffer.

Protein samples that were analysed under reducing electrophoretic conditions only, were denatured and reduced in boiling mix containing β-ME by heating to 95°C for 5 min before loading. Alternatively, for analysis of proteins under both reducing and non-reducing conditions, samples were denatured in boiling mix lacking β-ME and heated to 95°C for 5 min. The samples were then 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 was further promoted by another incubation step at 95°C for 5 min. If reduced and non-reduced were analysed on the same gel, NEM to a final concentration of 20mM was added to each sample prior to loading to prevent reduction of non-reduced samples by diffusion of the reducing agent.

Molecular weight markers (Amersham) were resolved alondside the samples to allow protein size determination. Gels were run at 100V until the bromophenol blue dye reached the bottom of the resolving gel. Thereafter, gels were removed from the apparatus and immersed in fix solution for 30 min, then in En^{3}Hance for 30 min and finally washed in water for 15 min. Gels were then dried under vacuum and exposed to Kodak X-Omat UV film for the detection of radiolabelled proteins. In the latter half of the project, the En^{3}Hance step was omitted and the gels were exposed to phosphorimager screens and later analysed using the Bio-Rad personal molecular imager FX and accompanying software. For non-radiolabelled proteins, the proteins of the gel were analysed by Western blot as described below.

2.13. Western blot analysis

2.13.1. Electroblotting to nitrocellulose membrane

Proteins were resolved by SDS-PAGE as described in 2.12 and were then transferred to nitrocellulose membrane in a Bio-Rad mini transblot apparatus, as described by Towbin *et al.* (1979). Briefly, a blotting sandwich was set up where the gel and nitocellulose membrane were placed, in contact with each other, between two pieces of Whatman 3mm paper which in turn were placed between two fibre pads. All materials were soaked in Towbin buffer prior to assembly of the sandwich and the sandwich was then transferred to the transblot apparatus and the reservoir filled with Towbin buffer. Electrotransfer was carried out at 100mA for 4 hours or 25mA overnight at 4°C.

2.13.2. Immunodetection

After transfer, membranes were blocked in 5% skimmed milk (Marvel) prepared in PBS(A) at 4°C overnight. Following three 10 min washes with PBS(A) containing 0.05% (v/v) Tween-20 (PBST), membranes were incubated with the primary antibody (diluted in 1% BSA in PBST) for 4 hours. After three washes with PBST, the secondary antibody (either anti-mouse IgG or protein-A conjugated to horseradish peroxidase), diluted in 1% BSA in PBST, was added and incubated for 1hour. The

membranes were washed for a final three times and the proteins detected using the Amersham enhanced chemiluminescence (ECL) system. The two supplied reagents were mixed in a 1:1 ratio and placed on the membranes for 2 min with agitation. The membranes were then placed between two sheets of mellanine and exposed to Kodak XS-1 film for 10-60 seconds. All the washing and antibody incubation steps were performed at room temperature with shaking.

2.14. Enzyme linked immunosorbant assay (ELISA)

2.14.1. GNA capture ELISA

Immulon 2HB 96 well plates (Dynex) were coated with 125ng of *Galanthus nivalis* agglutinin (GNA), diluted in 100µl of PBS(A), per well and left at room temperature overnight. The wells were washed with PBST three times before the addition of 200µl of block solution (2% skimmed milk (Marvel) in PBST) for 2 hours. Wells were again washed with PBST and then 100µl of two fold dilutions of cell extract expressing HCV E2 (prepared as described in 2.9.2) or a pre-determined dilution, was added and incubated for 1 ½ hours. The cell extracts were diluted in PBS(A). After, three more PBST washes, 100µl of an anti-E2 antibody (diluted in block solution) was added and incubated for 1 ½ hours. Another washing step was followed by the addition of 100µl of the secondary antibody (anti-mouse IgG or anti-rabbit IgG conjugated with horseradish peroxidase) diluted in block solution and incubated for 1 ½ hours. After a final wash step, 100µl of Tetramethyl benzidine (TMB) substrate was added to each well and incubated for 30 min with the plate covered in tin foil. To stop the reaction, 50µl of 0.5M H_2SO_4 was added to each well and the absorbance reading of each well at 450nm determined using the Anthos Labtec IIT2 plate reader. All incubations were done at room temperature.

2.14.2. CD81 capture ELISA

The human CD81 used for the ELISA was the extracellular loop 2 of CD81, expressed as a GST fusion protein. The CD81 capture ELISA was performed exactly as described above for the GNA capture ELISA, except that the plates were coated with

525ng of human CD81 per well (diluted in PBS(A)) instead of GNA. For the CD81 capture ELISA in reverse, wells were coated with anti-E2 antibody (ALP98 at 1/1000), diluted in PBS(A). After washing and blocking as before, HCV E2 cell extracts were added and incubated for 1 ½ hours. This was followed by washing and the addition of 525ng of human CD81 per well which was incubated for 1 ½ hours. After washing, CD81 was detected by the addition of a rabbit anti-GST antibody (diluted in block solution) which was incubated for 1 ½ hours. And after another wash step, anti-rabbit antibody conjugated with horseradish peroxidase (diluted in block solution) was added and incubated for a further 1 ½ hours. Following a final wash step, 100µl of TMB was added to each well and incubated for 30 min with the plate covered in tin foil. As, described above, the reaction was stopped by the addition of 0.5 M H_2SO_4 and the absorbance read at 450nm. All incubations were done at room temperature.

2.15. Immunofluorescence

Transfected cells on coverslips were fixed in ice-cold methanol and placed at -20°C for 20 min. Cells were rehydrated by incubation with PBS for 15 min and then permeabilised with PBST for 15 min. The cells were then incubated with 200µl of the appropriate primary antibody (diluted in PBST) for 1 ½ hours with agitation at room temperature. Following three 5 min washes with PBST, the cells were incubated for 1 ½ hours with 200µl of diluted secondary antibody, conjugated with either FITC, TRITC, Cy3 or Cy5. For double-labelling, cells were incubated with two primary antibodies and two secondary antibodies which were conjugated with different fluoresceins. Cells were then washed twice with PBST and twice with PBS before being rinsed in water, dried and mounted onto slides with 5µl of Citifluor mounting fluid (supplied by Chem. Lab., Canterbury). The coverslips were sealed with nail varnish and stored at 4°C in the dark.

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

2.16. CD81 binding of virus particles

Glutathione agarose beads were coated with the human CD81-GST fusion protein or GST alone by incubation overnight at room temperature. The beads were then blocked with 2% BSA in TEN buffer (1M Tris-Cl pH 8.0, 50mM EDTA, 2M NaCl) for 1 hour at room temperature on a rotator, followed by 5 washes with 3 ml TEN buffer. A slurry was made with the addition of 0.5 ml TEN buffer and 50 µl aliquots of beads were transferred to eppendorf tubes for each reaction. After a brief centrifugation at 6000 rpm and removal of the supernatant, 200 µl of serum (either neat or diluted in TEN buffer) was added to the beads. The tubes were sealed with parafilm and incubated for 2 hours at 37°C on a rotator. The beads and serum were removed to a column and the beads washed 5 times with 5 ml of TEN buffer. After washing, the beads were removed to Eppendorf tubes, centrifuged briefly and the supernatant removed. Any bound virus was removed by the addition of 560 µl AVL buffer (Qiagen QIAmp kit) and incubation at room temperature for 10 min. RNA was then extracted as described in 2.2.7, reverse transcribed using Omniscript reverse transcriptase as described in 2.2.8.2 and amplified and quantified as described in 2.2.9.3.

2.17. Computer software

Sequences were initially analysed and edited using the Sequence Navigator program. Alignments and final editing of DNA and protein sequences were carried out using the programs FRAGMENT ASSEMBLY SYSTEM, PILEUP and PRETTY in the Genetic Computer Group sequence analysis package version 10.2. The Sequence Detector Software from Perkin-Elmer was used for detecting PCR product produced in the Perkin Elmer Applied Biosystems 5700 sequence detection system. An AGFA StudioScan II si scanner and Adobe Photoshop 5 and Adobe Illustrator 7.0 programs were used to process the figures. EndNote 3.0 was used for creating a bibliography.

Chapter 3

Cloning of HCV genotype 3 core, E1 and E2 from the serum of an infected patient

3.1. Introduction

Nucleotide sequence analysis of HCV isolates from around the world has revealed substantial variation over the viral genome. Phylogenetic analysis of nucleotide sequences of the complete genome or subgenomic regions (core, E1 and NS5B) has led to the classification of six distinct HCV genotypes (1-6) which can be further divided into subtypes (a-h) (Simmonds *et al.*, 1993 & 1994). Over the complete genome, the degree of nucleotide sequence identity between major genotypes ranges from 66-69% and from 77-80% between subtypes (Simmonds, 1995). Genotype 3 was first identified by sequence comparison of subgenomic regions (Cha *et al.*, 1992, Chan *et al.*, 1992). The first full-length genome sequence for genotype 3 to be published was that of NZL1 (accession no: D17763) which was isolated from a New Zealand blood donor (Sakamoto *et al.*, 1994). When compared to the full-length sequence of other HCV isolates (genotypes 1a, 1b, 2a and 2b), this isolate showed 65.7%-68.9% nucleotide homology, indicating that this isolate was distinct from those previously reported and hence was classified as a genotype 3a isolate. The HPCHK6 (accession no: D28917) isolate is the only other published full-length sequence of a genotype 3a isolate (Yamada *et al.*, 1994) but two more unpublished sequences are available from the EMBL database, HCVCENS1 (accession no: X76918) and an unnamed isolate (accession no: AF046866). These four

are the only sequences covering the entire core, E1 and E2 genes from genotype 3a isolates.

Geographically, genotype 3 has a wide distribution although its prevalence varies from one geographic area to another. In Thailand it accounts for the majority of infections (Kanistanon *et al.*, 1997) and a wide range of genotype 3 subtypes (3a-3f) has been reported in Nepal (Tokita *et al.*, 1994). Subtypes 3c-3f have not been reported elsewhere, suggesting that they have evolved and stayed within Nepal, presumably due to the isolated position of the country. In Australia, genotypes 1 and 3 are the most commonly reported genotypes (Kaba *et al.*, 1998). Genotype 3 is found throughout Western Europe and North America (along with types 1a, 1b, 2a and 2b) and accounts for approximately 16% and 8% of infections respectively (Trepo & Pradat, 1999, Blatt *et al.*, 2000). In Scotland the prevalence of genotype 3 is 47% (Sheila Cameron, Consultant Scientist, Regional Virus Laboratory, Gartnavel General Hospital, Glasgow). This figure was taken from all HCV positive patients (up to June 2000) who have had samples sent to the Glasgow laboratory but these patients reside all over Scotland and therefore this figure represents approximately 4/5 of Scotland with only the Edinburgh area being excluded. This high prevalence may be explained by the fact that genotype 3 is particularly common amongst intravenous (IV) drug users (Pawlotsky *et al.*, 1995, Mihm *et al.*, 1997) and in Glasgow an estimated 70% of IV drug users are HCV positive (Glasgow General Health Board report on HCV, 2000). Collecting the epidemiological data reported in the studies above is labour-intensive and expensive and therefore large studies are rarely undertaken. Additionally, the population being sampled should be carefully considered. Often this population is voluntary blood donors who are healthy adults. This obviously does not reflect the general population and leads to an underestimate in HCV prevalence and may misrepresent the true genotype distribution.

Whether or not genotype influences the clinical course of disease is controversial, as it is likely that host and other viral factors also play a role. Several studies have suggested an association between infection with genotype 1b and the development of more severe liver disease including hepatocellular carcinoma (Dusheiko *et al.*, 1994, Bruno *et al.*, 1997).

However, it has also been shown that genotype 1b is significantly associated with older age and longer duration of infection and when these factors are taken into consideration there is no correlation between genotype and severity of disease (Martinot-Peignoux *et al.*, 1999, Mihm *et al.*, 1997). Some manifestations of the disease do appear to be related to genotype. Steatosis is a condition in which lipids accumulate in hepatocytes. It has been shown to be significantly associated with genotype 3 (Mihm *et al.*, 1997, Rubbia-Brandt *et al.*, 2000). However, these studies were criticised as genotype 3-infected patients often have a past history of IV drug use and hence could have been exposed to other blood-borne infections. Also, accompanying alcohol abuse is not uncommon. In answer to these objections, Rubbia-Brandt *et al.* (2001) conducted a thorough study investigating the relationship with genotype independently from other known risk factors such as obesity, alcohol abuse and IV drug use and found that there was still a correlation between steatosis and genotype 3.

Genotype is also an established predictor of response to interferon (IFN) therapy. Numerous studies have shown that genotypes 2 and 3 are associated with a better response to treatment than genotype 1. This is true whether IFN is used alone (Martinot-Peignoux *et al.*, 1995) or in combination with ribavirin (McHutchinson *et al.*, 1998) or if the pegylated form of IFN is used (Zeuzem *et al.*, 2000). The E2 and NS5A proteins have both been implicated in the response to IFN therapy and their involvement will be discussed in detail in Chapter 5. Suffice to say, the exact mechanisms behind the differing response to IFN and the development of steatosis remain unknown.

Although there are a large number of research groups around the world involved in analysis of the HCV proteins, by far the majority has concentrated on those of genotype 1 only. Comparison of the characteristics of these proteins with those of other genotypes would surely further the current understanding of the virus life cycle and virus/host interactions. A conserved region between genotypes suggests that these regions are essential to the virus life cycle whereas a unique region may point to a function that is unique to that particular genotype. In this study these ideas were applied to the two putative envelope proteins, E1 and E2. There is no infectious clone available for

genotype 3 so it was necessary to amplify the structural genes from the serum of a genotype 3 infected patient. I reasoned that by ensuring the PCR majority sequence is maintained in the clone, the expressed protein would function in the same way as that produced in the natural infection. Support for this is given by construction of infectious genotype 1 cDNA clones that were successful only when the cDNA sequence matched that of the PCR majority sequence (Kolykhalov *et al.*, 1997, Yanagi *et al.*, 1997). The encoded proteins could then be expressed and the similarities and differences between the type 1 and type 3 envelope proteins assessed.

3.2. Selection of patient serum

The criteria for selection of a suitable patient for this study were; infection with HCV genotype 3 and the availability of several aliquots of stored sera. The patient from whom serum was obtained was a 46-year-old former IV drug user. Informed consent was obtained at the times of blood sampling and approval for the study had been given by the West Glasgow Ethical Committee. The infecting HCV genotype was determined as 3a by restriction fragment length polymorphism (RFLP) and direct sequencing of the NS5A gene. A liver biopsy sample taken 12 weeks pre-treatment was given an Ishak score (Ishak *et al.*, 1995) of 3 for inflammation and 1 for fibrosis. Following treatment in 1994, the patient was classified as a non-responder after failing to clear HCV RNA (as determined by RT-PCR) after 12 weeks of treatment with interferon alpha-2a (Roferon; Roche Products Ltd) at a dose of six million units thrice weekly. Aliquots of sera were available pre-treatment, during treatment and post-treatment. This same patient was chosen as part of a study on NS5A and the ISDR amino acid sequence was found to be identical to that of the prototype 3a sequence, NZL1 while the complete NS5A sequence was shown to have 95% homology to NZL1 (McKechnie *et al.*, 2000). An immediate pre-treatment serum sample, from which the NS5A gene had been amplified in previous studies, was chosen for amplification of the HCV structural genes.

3.3. Amplification of the HCV structural genes by RT-PCR

Oligonucleotide primers were designed to amplify the HCV genome from the end of the 5'UTR to the beginning of NS2, thereby incorporating the core, E1, E2 and p7 genes. The sequences of these primers and the nucleotide positions at which they bind can be seen in Table 3.1. *BamH*I and *Bgl*II restriction endonuclease sites were incorporated into the 5' ends of the primers for second round amplification. This would facilitate future cloning. In addition, stop codons (one for each reading frame) were incorporated into the anti-sense primer for second round amplification. These stop codons would ensure that the correct size protein product was subsequently made. Following extraction from the infected serum, the viral RNA was reverse-transcribed using the MS2 primer. The resulting cDNA was then amplified, first with MS1 and MS2 and then with the internal primers, MS3R and MS4R (Figure 3.1). The PCR product was resolved by agarose gel electrophoresis (Figure 3.2). Initially an annealing temperature of 65°C was used for both rounds of amplification, which resulted in a product of the correct size (2.5 kb) (Figure 3.2A). Although there was no amplification of non-specific products, the overall amount of product was not optimal, so the annealing temperature was reduced in 2 or 3 degree increments for the second round of amplification (Figure 3.2B). Maximal amounts of product were seen with annealing temperatures of 58°C and 60°C (lanes 2 and 3 of fig 3.2B), but 60°C was chosen for all future amplifications, as there were fewer non-specific products at this temperature.

3.3.1. Sequence comparison

For sequencing purposes, the PCR product was purified and initially sequenced using the inner PCR primers. The nucleotide sequence obtained with these primers was then used to design more sequencing primers which were specific for the PCR product. Therefore it was possible to obtain sequence of overlapping fragments which spanned the PCR product twice, once in each direction. An alignment of the full sequence (Gla-3a) with four

Table 3.1. Oligonucleotide primers for nested RT-PCR of genotype 3 structural genes.

Primer	*Sequence 5' to 3'*	*Position*	*Use*
MS1 (s)	GCTTGCGAGTGCCCCGGGAG	301 - 320	PCR 1
MS2 (as)	TGAC/TAAGGTAAAGAAGCCG	2841 - 2859	PCR 1 & RT
MS3R (s)	GCGCGGATCCAGATCTGGGAGGT CTCGTAGACCGTG	316 - 335	PCR 2
MS4R (as)	GCGCGGATCCAGATCTTCATTTAT TTAAGAAGCCGAAGAGGACCAAG	2829 - 2848	PCR 2

Nucleotide numbering according to NZL1 (Sakamoto *et al.*, 1994).

s = sense, as = antisense

HCV sequence is shown in red.

*BamH*I restriction sites are shown in blue.

*Bgl*II restriction sites are shown in green.

Stop codons are underlined.

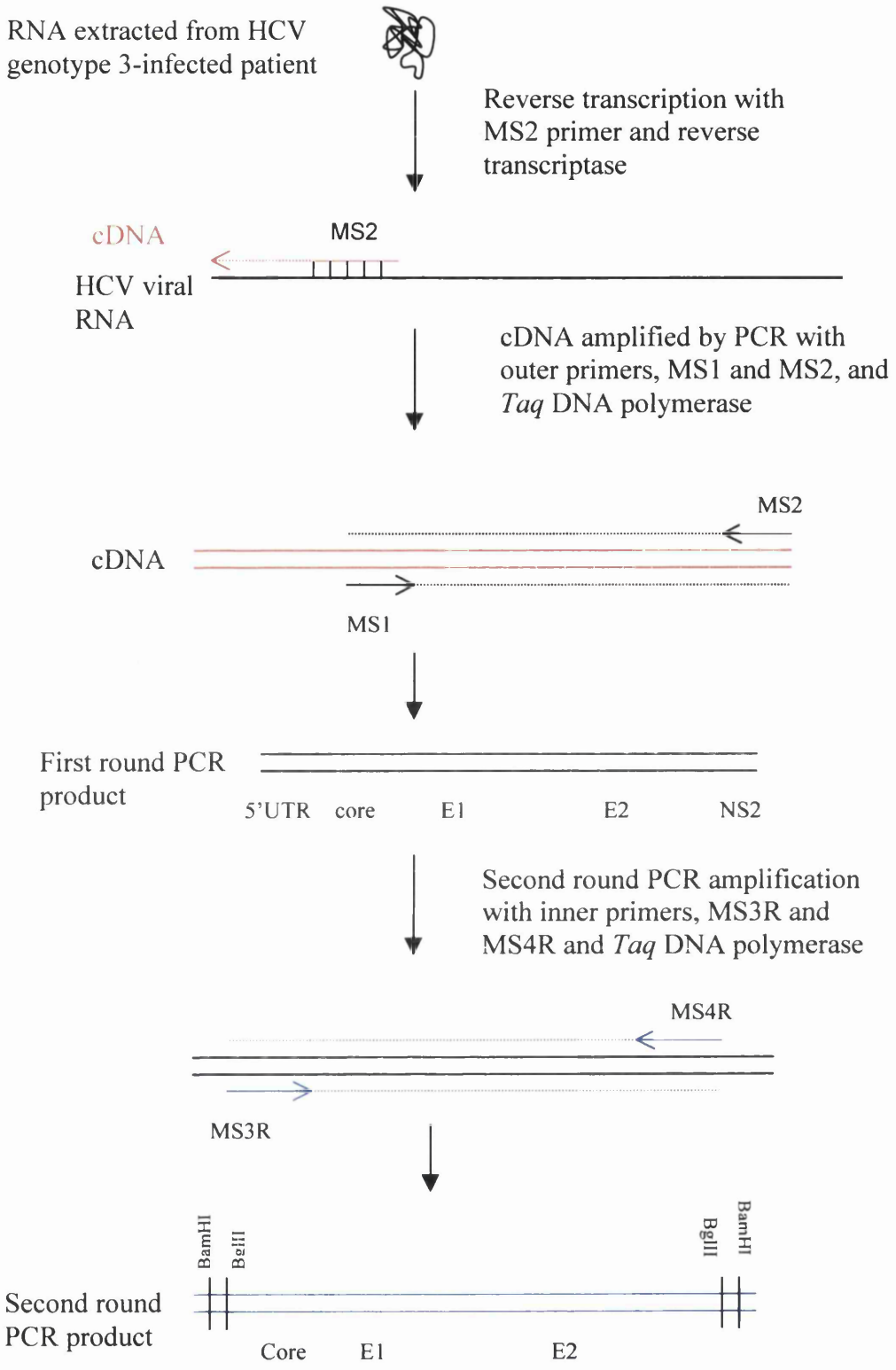


Figure 3.1. Illustration of the amplification of the core, E1 and E2 genes from the serum of a HCV genotype 3 infected patient by reverse transcription of viral RNA and amplification of cDNA by nested PCR.

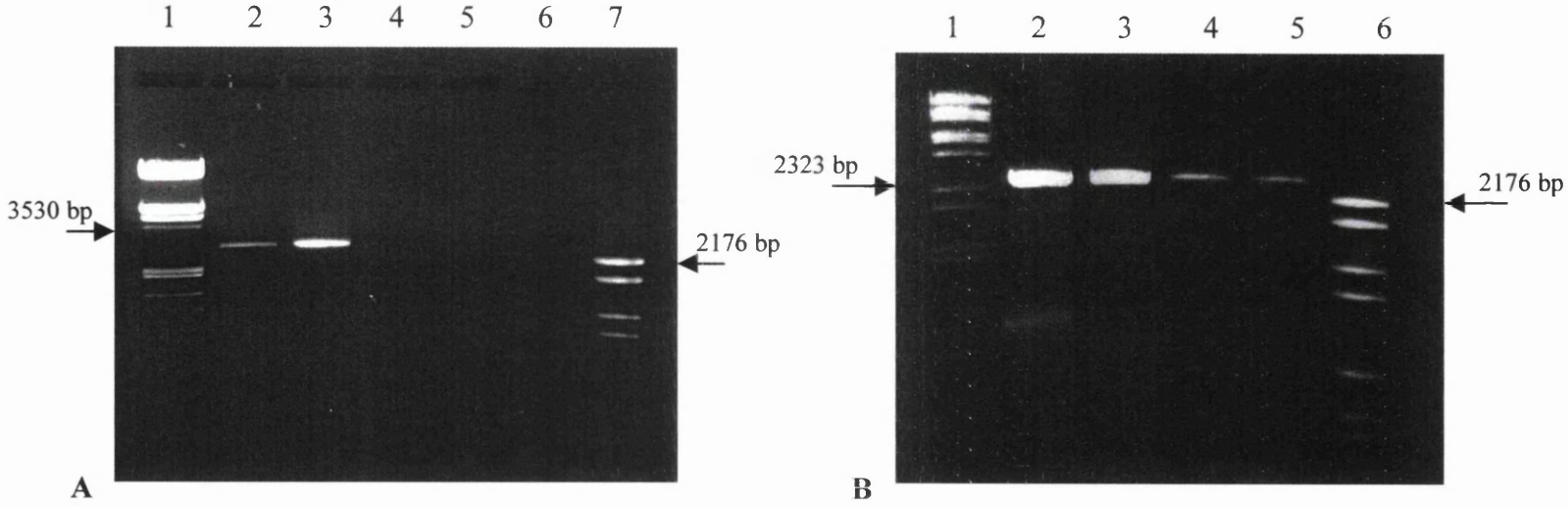


Figure 3.2. Agarose gel electrophoresis of second round PCR products of the type 3 HCV structural genes. A) shows the initial experiment where RNA was extracted from serum, reverse transcribed using the MS2 primer and underwent the first round of amplification with MS1 and MS2 and a second round of amplification with MS3R and MS4R. This was done in duplicate (lanes 2 and 3) and an annealing temperature of 65°C was used. Lanes 4, 5 and 6 show the "no template" controls for the RT, PCR-1 and PCR-2 stages respectively. Lanes 1 and 7 contain the molecular weight markers III and VI respectively (Roche). B) shows an annealing temperature optimisation experiment where annealing temperatures of 58, 60, 62 and 65°C (lanes 2-5) where used for the second round of amplification. Lane 1 contains the λ *BstE*II molecular weight marker and lane 6 contains the molecular weight marker VI (Roche).

genotype 3a reference isolates is shown in Figure 3.3. Overall, the five nucleotide sequences for core, E1 and E2/p7 showed 83.8% homology. Further analysis showed 90.4% homology for core, 85.6% for E1 and 80% for E2/p7.

In Figure 3.4, the derived amino acid sequence of Gla-3a was aligned with that of the four genotype 3 isolates. For comparison, the genotype 1a reference strain (H77) was also included. Analysis of the genotype 3 protein sequences revealed 93.2% homology in core, 91.6% homology in E1 and 80.9% homology in E2/p7. By comparison with genotype 1a (H77), Gla-3a showed 92.7% homology in core, 70.7% in E1 and 72.6% in E2/p7. The E1 protein of both genotypes 1 and 3 was 191 amino acids in length and there were 8 conserved cysteine residues at positions 207, 226, 229, 238, 272, 281, 304 and 306 (indicated with "*"). Proline residues were also conserved in both genotypes at positions 224, 228, 244, 295, 310, 328 and 341 of E1. Additional prolines were found at positions 208 and 254 in H77, position 241 in all four type 3 isolates and position 283 in HPCHK6.

The E2 protein of genotype 1a was 355 amino acids in length whereas E2 of genotype 3 was 361 amino acids. The insertions in genotype 3 occurred at positions 480, 571 and 579-582. There were 19 cysteine residues in E2 that were conserved between genotype 1 and 3 and one in the p7 protein (indicated with "*"). Genotype 1 had an additional cysteine at position 740. There were 21 proline residues in E2 and one in p7 that were conserved in both genotypes. All 5 genotype 3 isolates had an additional proline residue at position 461. There were another 8 proline residues that were not conserved within the genotypes, including two in the hypervariable region (HVR1). The HVR1 is highlighted in yellow and consists of the first 27 amino acids of E2. It was evident that the degree of variation in this region was substantially higher than elsewhere in E2, and was not related to genotype. This sequence was expected to differ between the different quasispecies present in the one isolate and therefore no interpretation could be made from the sequence of the HVR1 seen here, which presumably represented that of the majority sequence.

When comparing Gla-3a to the four previously published genotype 3 isolates, there were positions in both E1 and E2 where Gla-3a differed from the other four. There were 3

Figure 3.3. Alignment of the majority nucleotide sequence of genotype 3a core, E1 and E2 genes amplified by PCR (Gla-3a) and the equivalent sequence from four genotype 3a reference isolates, NZL1, HPCHK6, HCVCENS1 and AF046866. The core sequence is shown in black, the E1 sequence in red, the E2 sequence in blue and the p7 sequence in green. "-" indicates ambiguities between the sequences. Nucleotide numbering is according to that of NZL1 (Sakamoto *et al.*, 1994).

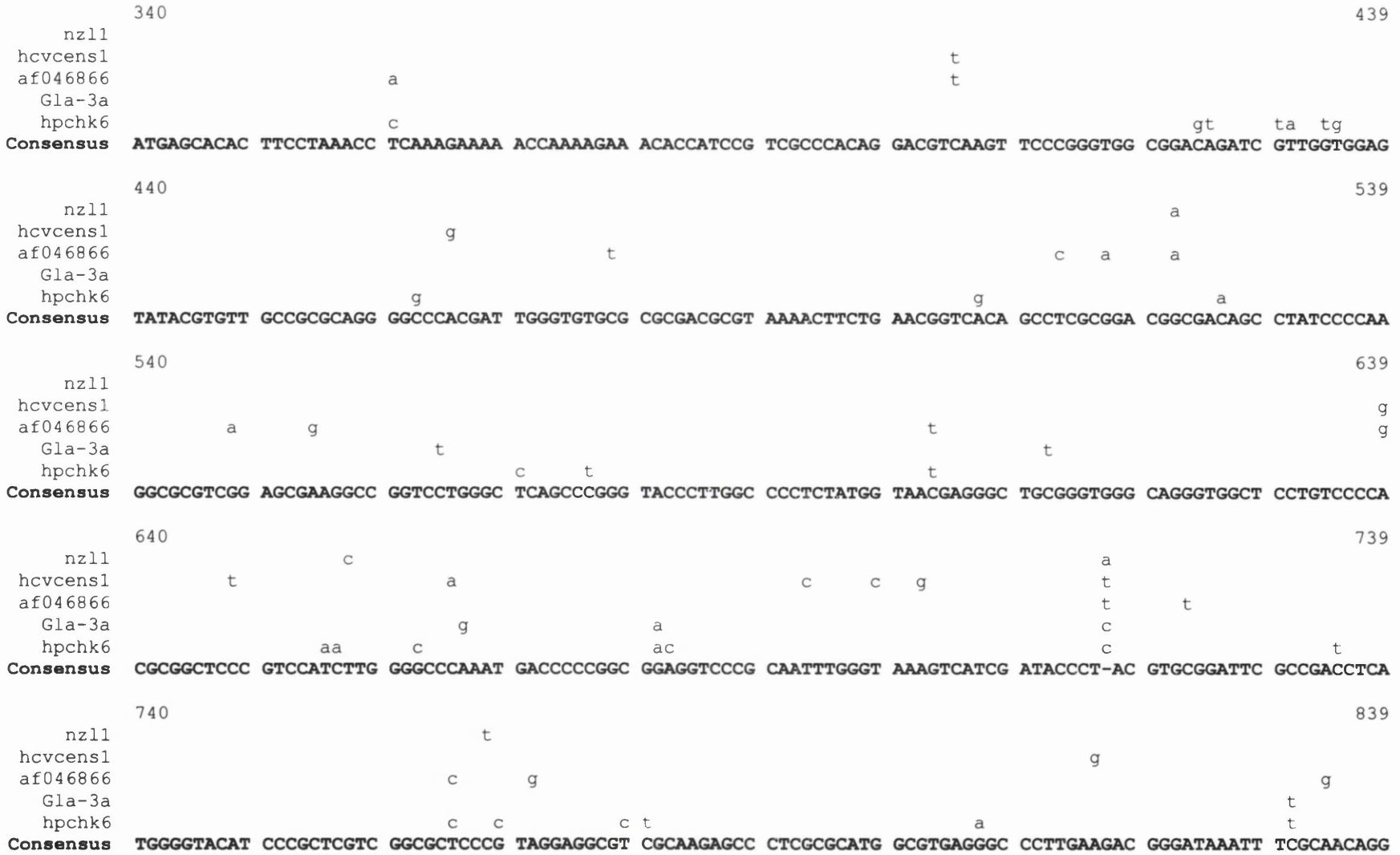


Figure 3.3 Part 1

Figure 3.3 Part 2

Figure 3.3 Part 3

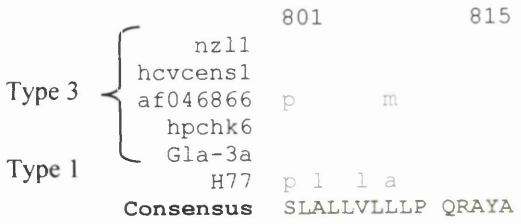
Figure 3.3 Part 4

Figure 3.3 Part 5

Figure 3.4. Amino acid sequence derived from the amplified core, E1 and E2 genes (Gla-3a), aligned with the equivalent sequence from four published genotype 3a isolates (NZL1, HPCHK6, HCVCENS1 and AF046866) and a genotype 1a reference strain (H77). Core sequence is shown in black, E1 in red, E2 in blue and p7 in green. The hypervariable region is highlighted, glycosylation sites are underlined and conserved cysteine residues are indicated with a "*". "." denotes a missing amino acid and "-" indicates there is no consensus at that position.

Figure 3.4 Part 1

Figure 3.4 Part 2



KEY

XXX – Core sequence

XXX – E1 sequence

XXX – E2 sequence

XXX – Hypervariable region in E2

XXX – p7 sequence

XXX – Glycosylation site in E1&E2

* – Conserved cysteine residue in E1&E2

Figure 3.4 Part 3

such differences in E1; T223A (i.e threonine in consensus and alanine in Gla-3a at position 223), D232V (although this was an asparagine in AF046866) and P241S. Ignoring the hypervariable region and any position where there was no consensus, there were 13 positions in E2 where Gla-3a differed from the other four. They were E431D, I433L, F442I, F464S, P480S (although this was an alanine in HCVCENS1), A491P, I497T, A532G (this was glutamic acid in NZL1 and HCVCENS1), T559S, F586L, Q668L, K721R and K787R.

Four predicted glycosylation sites at positions 196, 209, 234, and 305 (underlined in Figure 3.4) in the E1 protein were present in both genotypes 1 and 3. A fifth site was predicted at position 325 but it had previously been demonstrated that only 50% of such sites (with a C-terminal proline) were used (Gavel & von Heijne, 1990). The E2 protein had 7 predicted glycosylation sites that were conserved between H77 and all type 3 isolates, at positions 417, 423, 430, 448, 476, 557 and 651. Two further glycosylation sites at positions 533 and 629 were present in H77 and all type 3 isolates except NZL1. Due to the insertion present in genotype 3, the glycosylation site at position 578 in genotype 1 was disrupted in genotype 3 but an alternative had been introduced at position 581. H77 had an additional site at position 541 that was not present in genotype 3.

3.4. Cloning of the HCV structural genes into the SFV vector

The Semliki Forest Virus (SFV) expression system was chosen for the purposes of producing the protein products of the HCV structural genes. The pSFV1 expression vector, which forms the basis of the SFV expression system, contains a polylinker cassette with a *BamHI* restriction site downstream of the transcription initiation site (Figure 3.5). This site was used for insertion of the PCR product encoding core, E1 and E2. As mentioned previously, *BamHI* sites were added to the 5' ends of the PCR primers and hence into the PCR product as well. Both the PCR product and the pSFV1 vector were cleaved with *BamHI*, purified and then ligated together and electroporated into *E.coli.*

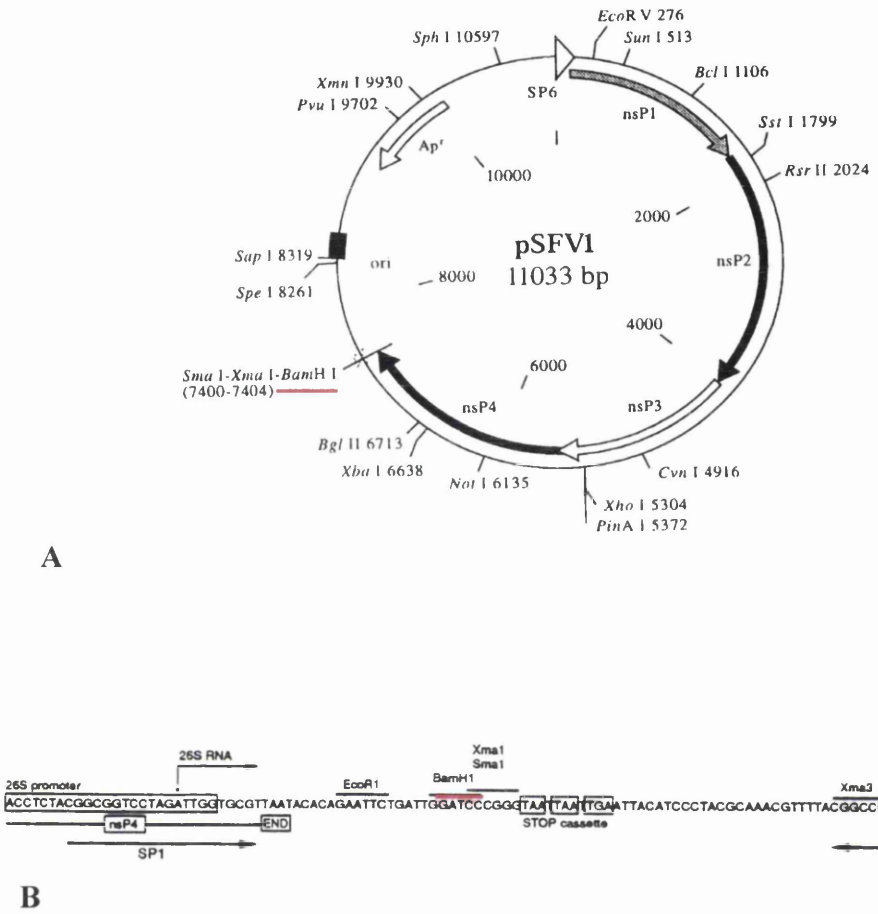


Figure 3.5. Schematic showing the pSFV1 plasmid (A) and the polylinker region (B). The multiple cloning site contains a unique *BamH*I restriction site (underlined in red) which allowed insertion of the PCR product encoding the HCV structural genes. The ampicillin resistance gene (Ap^r) and the SFV nonstructural proteins (nsP1-4) are shown (A). The 26S RNA promoter is boxed and the first transcribed nucleotide is marked with an asterix (B). The figures were taken from the Stratagene catalogue.

Clones were selected for presence of the inserted PCR product by digestion with *BamH*I or *Bgl*II (internal to the *BamH*I sites on the PCR product). This released a 2.5 kb fragment. The orientation of the insert within the vector was determined by cleavage with *Sma*I. From the PCR nucleotide sequence, three *Sma*I sites were identified and with one site present in the vector, it was possible to determine orientation of the insert by simple restriction endonuclease digestion. It was noticed that only 20% of the clones analysed had an insert in the correct orientation. The selected clone was named pSFVF2CE1E2. As with the PCR product, the clone was sequenced across the length of the insert. The derived amino acid sequence, which was aligned with that of the PCR product, is shown in Figure 3.6. There were 9 differences between these two sequences.

3.5. Site -directed mutagenesis of pSFVF2CE1E2

The nine amino acid mutations found in pSFVF2CE1E2 were all due to single nucleotide exchanges with one in core, two in E1, five in E2 and one in p7 (Figure 3.6). To correct these by site-directed mutagenesis, it was first necessary to subclone the HCV structural genes into another smaller vector to reduce the chances of introducing mutations into the SFV vector. The selected vector for this purpose was pGEM-1 (Figure 3.7). Due to the size of the HCV coding region, it was decided to divide this in two at the *Pst*I site within the E2 gene. Therefore pSFVF2CE1E2 was cleaved with *Xba*I, *Hind*III (both in the vector) and *Pst*I to release two fragments of 2022 bp and 1960 bp (Figure 3.8). These were cloned into pGEM-1 at the appropriate sites to create pMS1 and pMS2 respectively. pMS1 contained exchanges 1, 2 and 3 and pMS2 contained exchanges 4,5,6,7,8 and 9 and these plasmids were the templates for site-directed mutagenesis.

The strategy for site-directed mutagenesis was that employed by the "QuikChange" kit from Stratagene (Figure 3.9). Oligonucleotide primers designed against both DNA strands and containing the desired mutation were made for each exchange. As exchanges 4 and 5 were situated close together, one pair of primers containing both mutations was designed. These mutations were obviously made simultaneously whereas the remainder were made consecutively. Following denaturation of the template plasmid, the specific


1
100
Gla-3a
pSFVF2
n
y
Consensus MSTLPKPQRK TKRNTIRRPQ DVKFPGGGQI VGGVYVLPRR GPRLGVRATR KTSERSQPRG RRQPIPKARR SEGRSWAQPG YPWPLYG-EG CGWAGWLLSP
1
101
200
Gla-3a
pSFVF2
Consensus RGSRPSWGPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPVGGVARA LAHGVRALED GINFATGNLP GCSFSIFLLA LFSCLIHPAA SLEWRNTSGL
201
300
Gla-3a
s
l
pSFVF2
p
p
Consensus YILTNDCSNS SIVYEADDVI LHAPGCVPCV QVGNT-TCWT SVTPTVAVRY VGATTASIRS HVD-LVGAAT MCSALYVGDM CGAVFLVGQA FTFRPRRHQT
2
3
301
400
Gla-3a
pSFVF2
Consensus VQTCNCSLYP GHLSGHRMAW DMMMNWSPAV GMVVAHVLRL PQTLFDIIAG AHWGILAGLA YYSMQGNWAK VAIIMVMFSG VDAATYTTGG SVARGASILT
401
500
Gla-3a
a
p
pSFVF2
t
s
Consensus GMFSQGARQN LQLIKTNGSW HINRTALNCN DSLNTGFIAG LIYYHKFNST GCPQRLSSCK PITSFRQGWG SLTDANITGS SDDKPYCWHY PPRPC-TV-A
4
5
501
600
Gla-3a
s
e
pSFVF2
p
g
Consensus STVCGPVYCF TPSPVVVGTT DAKGAPTFNW GGNWTDVFLL ESLRPPSGRW FGCAWMN-SG FVKTCGAPPC DIYGGGGNPG N-SDLLCPTD CFRKHPEATY
6
7
601
700
Gla-3a
e
pSFVF2
g
Consensus SRCGAGPWLT PRCMVDYPYR LWHYPCTVNF TLFKVRMFVG GFEHRFTAAC NWTRGERCDI -DRDRSELHP LLHSTTELAI LPCSFTPMPA LSTGLIHLHQ
8

```
          701                                                                                                  800
   Gla-3a                                                      t
   pSFVF2                                                      a
Consensus  NIVDVQYLYG VGSGMVGWAL RWEFVILVFL LLADARVCVA LWLMLMITQA EAALENLV-L NAVAAAGTHG IGWYLVAFCA AWHVRGRLVP LVTYSLTGLW
                                                               9

          801          815
   Gla-3a
   pSFVF2
Consensus  SLALLVLLLP QRAYA
```

KEY

XXX - Core sequence

XXX - E1 sequence

XXX - E2 sequence

XXX - p7 sequence

Figure 3.6. Amino acid sequence derived from the core, E1 and E2 genes of pSFVF2CE1E2, aligned with the equivalent sequence derived from the PCR product (Gla-3a). The differences between the sequences are numbered 1-9 and highlighted. Core sequence is shown in black, E1 in red, E2 in blue and p7 in green.

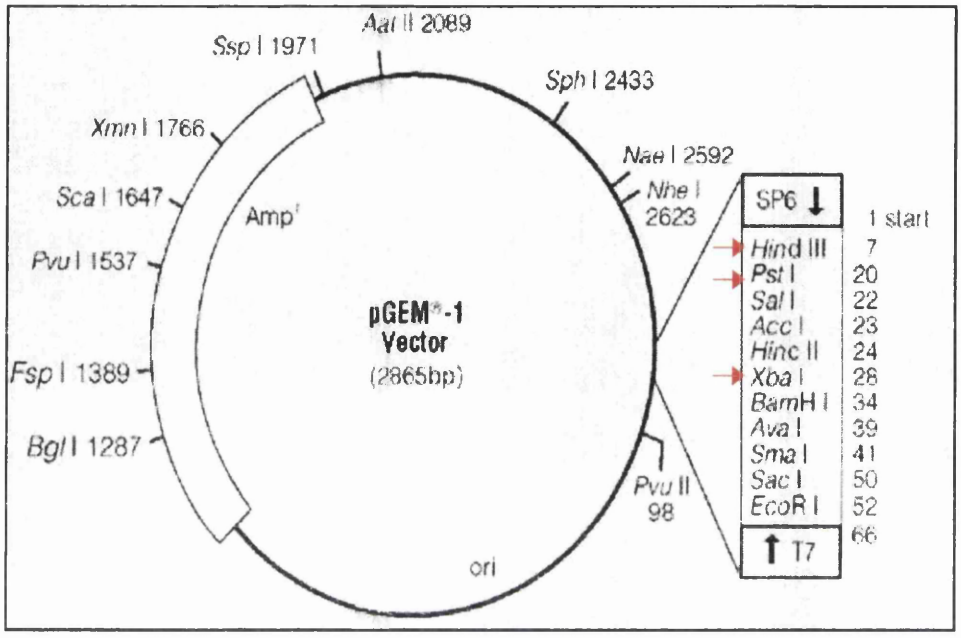
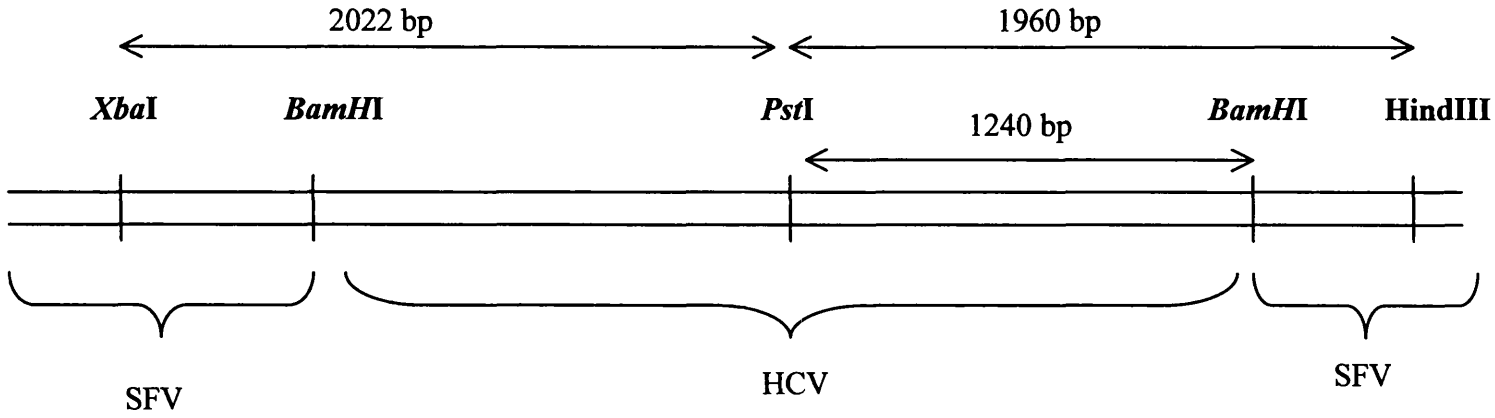


Figure 3.7 Schematic of the pGEM-1 plasmid. The *Hind* III, *Pst* I and *Xba* I restriction sites which were used for cloning are indicated with red arrows. The ampicillin resistance gene (Amp^r) is shown. The figure was taken from the Promega catalogue.



pSFVF2CE1E2

*Pst*I / *Xba*I digest

*Pst*I / *Hind*III digest

2022 bp fragment cloned into pGEM-1 to form **pMS1**

1960 bp fragment cloned into pGEM-1 to form **pMS2**

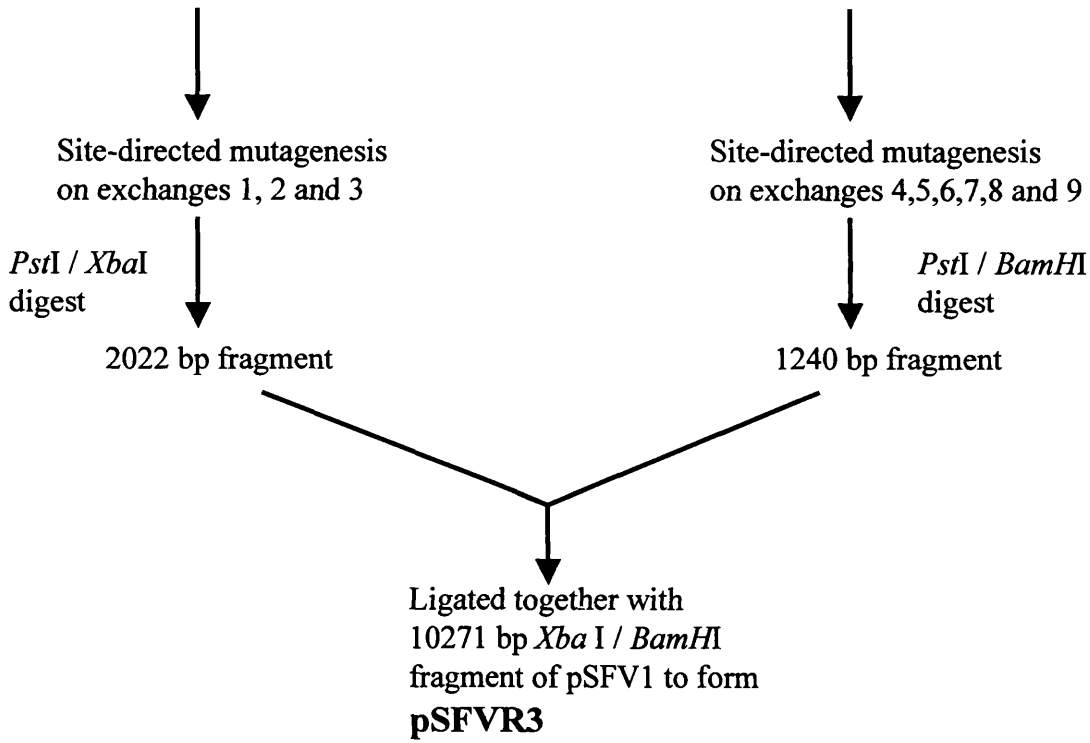


Figure 3.8. Illustration of cloning strategy prior to site-directed mutagenesis on exchanges 1-9 in pSFVF2CE1E2 and the construction of the repaired construct, pSFVR3CE1E2.

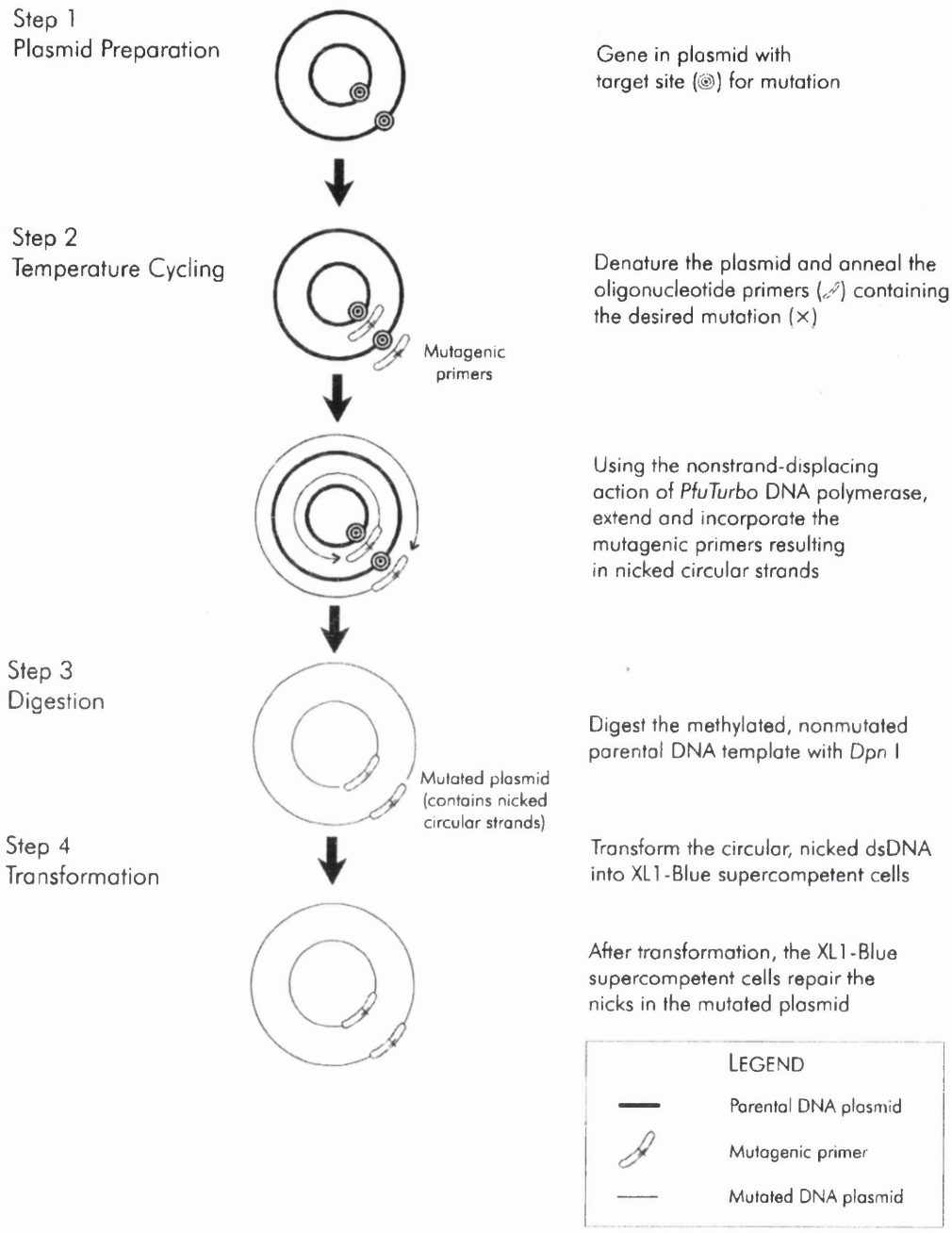


Figure 3.9. Illustration of the strategy used for site-directed mutagenesis, using the QuikChange kit from Stratagene.

primer pair was annealed and extended by *Pfu Turbo* DNA polymerase, which resulted in nicked, circular DNA strands. This was repeated for 12-24 cycles to amplify the mutated DNA. An example of the PCR product obtained during mutagenesis of mutation 9 in pMS2 can be seen in Figure 3.10. To remove the parental plasmid, the PCR product was digested with *Dpn*I, which is a 4 bp cutter and specifically cleaves methylated DNA. Only the parental DNA, which originated from *E.coli* would be methylated. The remaining, mutated DNA was transformed into *E.coli* where the nicked DNA was repaired. This plasmid DNA was then used as template for the next mutation, however I found that it was beneficial to ensure that this template consisted of supercoiled plasmid only as amplification of both DNA strands did not occur when nicked plasmid DNA was present. Therefore the plasmid DNA was first separated by gel electrophoresis and the supercoiled DNA extracted and purified before being used as template for the next round of mutagenesis.

After mutagenesis, the repaired HCV genes were reconstructed and cloned back into pSFV1 (Figure 3.8). The 2022 bp *Xba*I/*Pst*I fragment from pMS1 and the 1240 bp *Pst*I/*Bam*HI fragment from pMS2 were ligated together along with the 10271 bp *Xba*I/*BamH*I fragment of pSFV1. The resulting SFV clone was named pSFVR3CE1E2. This repaired clone was nucleotide sequenced over core, E1 and E2 to confirm that mutations 1-9 had been repaired and that no other mutations had been introduced. The derived amino acid sequence is shown in Figure 3.11 and was aligned with the amino acid sequence of Gla-3a. The amino acid sequence of the genotype 1a infectious clone, H77 (Yanagi *et al.*, 1997) was included in the alignment as this would be used for the expression of the representative type 1 proteins for comparison with those of Gla-3a. The derived amino acid sequence of the repaired clone was identical to that of the Gla-3a PCR majority sequence. The predicted transmembrane domains (TMD), which are responsible for ER retention, are boxed. The C-terminal halves of the E1 and E2 TMD sequence are thought to function as signal sequences for E2 and p7 respectively (Cocquerel *et al.*, 1998 & 1999) and are underlined in Figure 3.11. Amino acids within these TMDs, which were identified by Cocquerel *et al.* (2000) as being fully conserved between HCV variants, were also conserved in Gla-3a (indicated with "*"). These included charged residues in

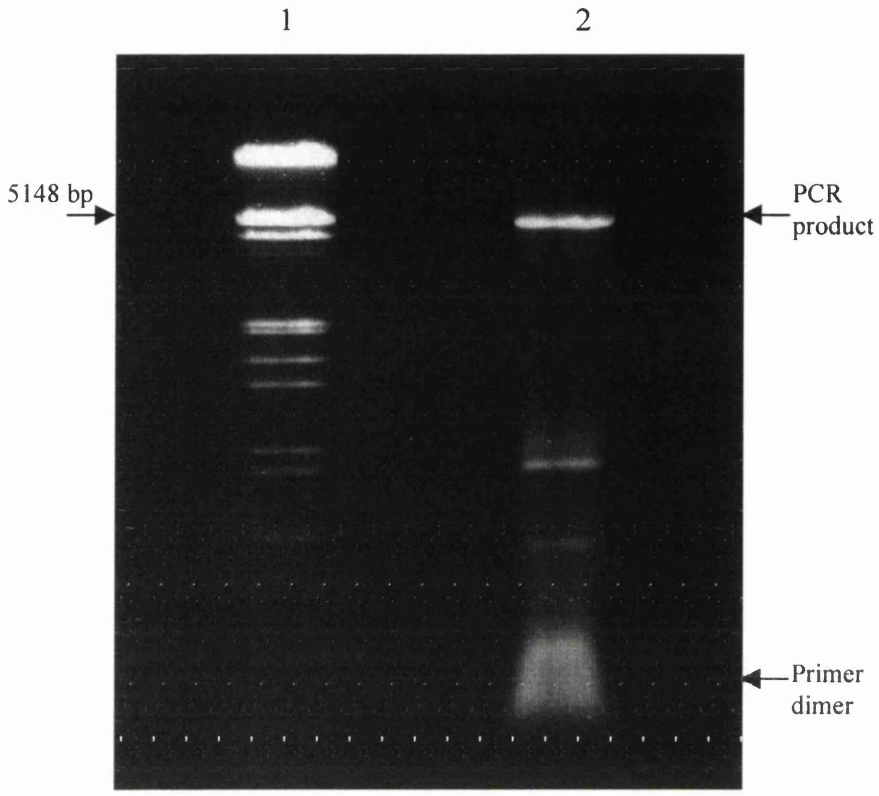
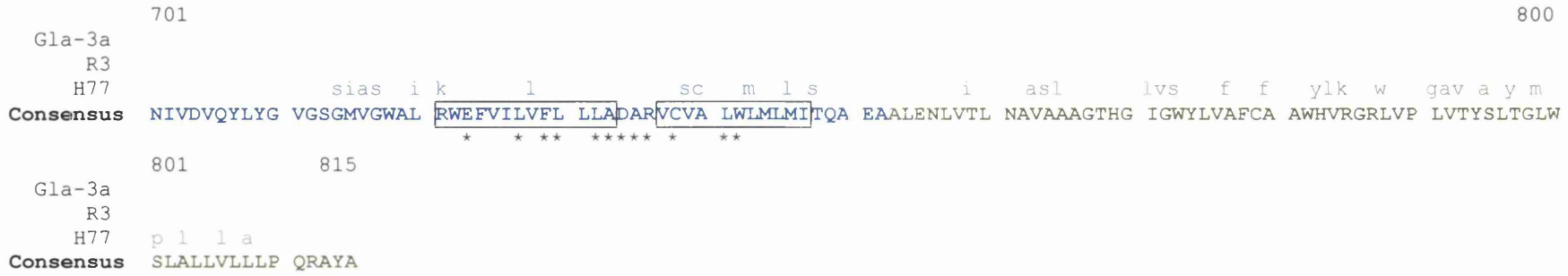


Figure 3.10. Agarose gel electrophoresis of the pMS2 PCR product produced in the temperature cycling step during mutagenesis of mutation number 9. Lane 2 shows the major PCR product of approximately 5 kb and some smaller, non-specific products. The smear at the bottom is likely to be due to formation of primer dimers. Lane 1 contains the molecular weight marker III (Roche).


1
100
Gla-3a
R3
H77
Consensus
MSTLPKPQRK TKRNTIRRPQ DVKFPGGGQI VGGVYVLPRR GPRLGVRATR KTSERSQPRG RRQPIPKARR SEGRSWAQPG YPWPLYGNEG CGWAGWLLSP
101
200
RGSRPSWGPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPVGGVARA LAHGVRALED GINFATGNLP GCSFSIFLLA LFSCLIHPAA SLEWRNTSGL
201
300
YILTNDCSNS SIVYEADDVI LHAPGCVPCV QVGNTSTCWT SVTPTVAVRY VGATTASIRS HVDLLVGAAT MCSALYVGDM CGAVFLVGQA FTFRPRRHQT
301
400
VQTCNCSLYP GHLSGHRMAW DMMMNWSPAV GMVVAHVLRL PQTLFDIIAG AHWGILAGLA YYSMQGNWAK VAIIMVMFSG VDAATYTTGG SVARGASILT
401
500
GMFSQGARQN LQLIKTNGSW HINRTALNCN DSLNTGFIAG LIYYHKFNST GCPQRLSSCK PITSFRQGWG SLTDANITGS SDDKPYCWHY PPRPCATVPA
501
600
STVCGPVYCF TPSPVVVGTT DAKGAPTFNW GGNWTDVFLL ESLRPPSGRW FGCAWMNSSG FVKTCGAPPC DIYGGGGNPG NESDLLCPTD CFRKHPEATY
601
700
SRCGAGPWLT PRCMVDYPYR LWHYPCTVNF TLFKVRMFVG GFEHRFTAAC NWTRGERCDI EDRDRSELHP LLHSTTELAI LPCSFTPMPA LSTGLIHLHQ



KEY

XXX - Core sequence

XXX - E1 sequence

XXX - E2 sequence

XXX - Hypervariable region in E2

XXX - p7 sequence

XXX – Predicted transmembrane domains (TMD) for E1 and E2 (Cocquerel *et al.*, 2000)

XXX – Predicted signal sequence for E1 (Santolini *et al.*, 1994) and E2 (Patel *et al.*, 1999b)

* - Conserved residue in TMD as identified by Cocquerel *et al.* (2000)

Figure 3.11. Amino acid sequence derived from the type 3 core, E1 and E2 genes amplified by PCR (Gla-3a) aligned with the equivalent sequence from the type 3 clone, pSFVR3CE1E2 (R3) and the infectious type 1 clone, H77 (H77). Core sequence is shown in black, E1 in red, E2 in blue and p7 in green. The hypervariable region in E2 is highlighted. "." indicates a missing amino acid.

the hydrophilic region between the two hydrophobic regions. Specifically, a lysine residue at position 370 in E1 and aspartic acid and arginine residues at positions 734 and 736 of E2 which had been shown to be crucial for the maintenance of structure and function of the TMDs (Cocquerel *et al.*, 2000), were all present.

3.6. Construction of genotype 3 E2 SFV clone

For some purposes it was necessary to observe the characteristics of E2 in isolation and to compare them to the clone expressing E2 in the context of E1. Therefore the E2 gene was subcloned from pSFVR3CE1E2 into the pSFV1 vector (Figure 3.12). Previous experience with expression of the HCV genotype 1 structural proteins with the SFV system had shown that the authentic HCV start codon and the sequence surrounding it, was required for good expression levels of any of the downstream genes (J. McLauchlan, personal communication). Therefore the strategy for subcloning E2 had to include the first few amino acids of core. Conveniently there were *Sma*I sites in pSFVR3CE1E2 which cleaved the nucleotide sequence at positions equivalent to 24 amino acids into core and at the end of E1 but before the E2 signal sequence. Cleavage with *Sma*I resulted in four fragments of 163 bp, 692 bp, 1577 bp and 11127 bp as shown in Figure 3.12. The 11127 bp fragment was dephosphorylated and ligated to the 1577 bp fragment to form pSFVR3E2. The orientation of the E2 gene within the vector was confirmed with *Eco*RV and *Pvu*II digestions.

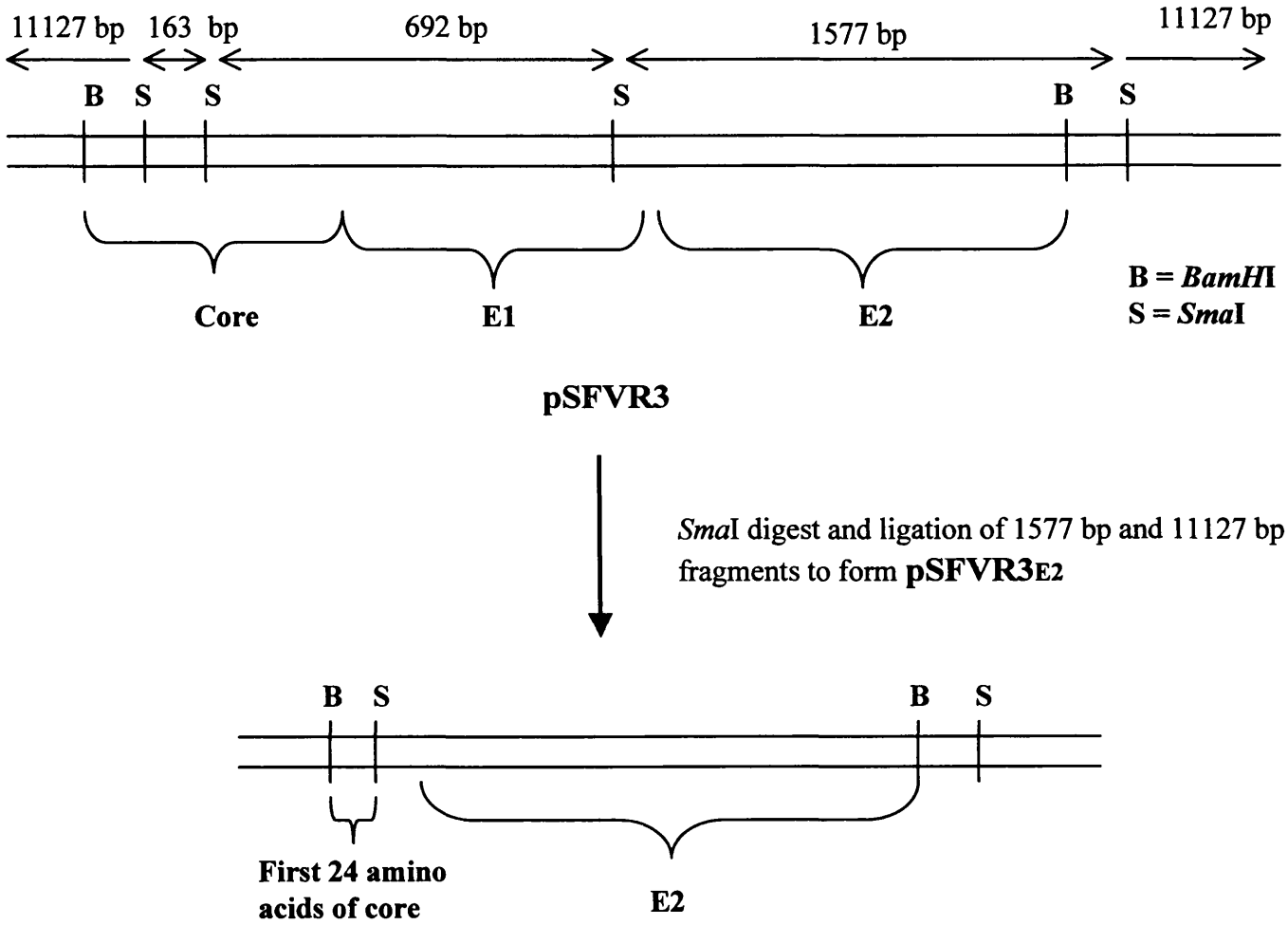

11127 bp
163 bp
692 bp
1577 bp
11127 bp
B
S
S
S
B
S
Core
E1
E2
B = BamHI
S = SmaI
pSFVR3
SmaI digest and ligation of 1577 bp and 11127 bp fragments to form pSFVR3E2
B
S
B
S
First 24 amino acids of core
E2

Figure 3.12. Illustration of strategy for subcloning E2 from pSFVR3CE1E2 to form pSFVR3E2. In addition to E2, pSFVR3E2 encodes the first 24 amino acids of core and the E2 signal sequence located at the end of E1.

3.7. Discussion

There are four complete genome sequences available for HCV genotype 3a in the EMBL database. They are NZL1 (Sakamoto *et al.*, 1994), HPCHK6 (Yamada *et al.*, 1994) and two unpublished sequences (AF046866 and HCVCENS1). A few other published sequences are available for genotype 3a core and E1 but none for the complete E2 gene and there are no reports of detailed analysis of the E2 gene from the different genotypes. When the PCR majority derived amino acid sequence of Gla-3a is compared to the four reference isolates (Figure 3.4), the percentage homology in the E1 protein is 91.6%. According to Bukh *et al.* (1993) this amino acid identity within E1 is consistent with these isolates being of the same genotype. In contrast, the 69.1% homology in E1 between Gla-3a and the genotype 1a reference strain, H77, confirms that these two isolates are different genotypes. The degree of homology in the core and E2 proteins (93.2% vs 80.9%) between the five genotype 3 isolates, is consistent with previous data showing that core is one of the most conserved regions of the genome whereas E2 is one of the most variable (Simmonds, 1995).

The derived E1 amino acid sequence shows that the unmodified proteins of genotypes 1 and 3 are identical in length and also share all 8 cysteine residues, which are likely to be involved in the formation of disulphide bonds and hence in maintenance of the correct conformational structure of the protein. Proline residues are also considered important for protein structure and are conserved in both genotypes in 7 positions, although additional prolines in some isolates could reflect structural differences. The unmodified E2 protein of genotype 1a is 6 amino acids shorter than that of genotype 3. Two of the insertions in genotype 3 involve a proline residue, which could alter the structure of the protein, although these residues are not conserved in all 5 of the genotype 3 isolates reported here. As with E1, the E2 proteins of genotypes 1 and 3 share several cysteine residues that are crucial for the correct conformation of the protein. Genotype 1 E2 has an additional cysteine residue, which could lead to differences in the structure of the E2 proteins of genotypes 1 and 3 through differences in disulphide bond formation. Proline residues are also highly conserved in E2 but genotype 3 has an additional conserved

proline residue which is absent in genotype 1 and there are other prolines which are not conserved within the genotypes. Therefore variation in E2 structure is expected, even between isolates of the same genotype, however the highly conserved nature of some residues between genotypes implies that conservation of structure in certain regions is necessary for protein function.

The HCV envelope proteins are heavily modified by *N*-linked glycosylation. The degree of glycosylation could alter the ability of E1 and E2 to interact with cellular proteins due to steric hindrance. Therefore, differences in glycosylation states between genotypes could impart different functional characteristics on the proteins. Both genotype 1 and genotype 3 E1 have four (or five) predicted glycosylation sites. The situation with E2 is more complicated, but if the two absent sites in NZL1 are ignored, genotype 3 has 10 predicted glycosylation sites and genotype 1 has 11. This is in agreement with that reported by Miyamura & Matsuura (1993) for genotype 1. The site in genotype 1, which is disrupted by the insertion in genotype 3, is reintroduced in the insertion implying that glycosylation in this region of E2 is important for protein function.

As the characteristics of the E2 protein expressed from Gla-3a will be analysed as a representative type 3 E2, it is important to note any amino acids which differ from the four other type 3 isolates. Possibly the most significant of these differences are those that involve proline residues, at positions 241, 480 and 491. At position 241 the consensus is a proline but this is a serine in Gla-3a. At position 480 the consensus residue is also a proline but it is also a serine in Gla-3a and an alanine in HCVCENS1. Interestingly this residue is absent in genotype 1. At position 491 the consensus residue is alanine and this is a proline in Gla-3a as well as in genotype 1. The structure of proline is very rigid due to the side chain being covalently bonded to the nitrogen atom of the peptide backbone. This prevents rotation and therefore has a significant effect on the conformation of the protein around these residues. It is likely that these differences at positions 241, 480 and 491 do induce a structural change in this region of E2.

Following cloning of the PCR product into pSFV1, the resulting clone, pSFVF2CE1E2 was sequenced across the core, E1 and E2 genes and 9 amino acid differences were discovered between the clone and the PCR product (Figure 3.6). These were all due to single nucleotide exchanges and were present throughout the insert, with exchange 1 in core, exchanges 2 and 3 in E1, 4-8 in E2 and exchange 9 in p7. Exchanges 2, 3, 5 and 6 all involve proline residues, which could have an effect on the protein structure as described previously. In addition exchanges 2, 6 and 7 abolish conserved glycosylation sites. Exchange 5 in E2 removes a conserved proline residue. Therefore I predicted that these changes would have an effect on the overall structure of E1 and E2. Construction of the infectious H77 clone revealed that to be infectious, the injected RNA transcripts had to match the consensus sequence (Yanagi *et al.*, 1997). These authors reported finding a high proportion of clones which were either defective for protein production or had multiple amino acid mutations. There are several reasons for the differences in our clone. They could have been introduced during the RT-PCR procedure due to the infidelity of the *Taq* polymerase. The cloned product could also represent a variant that was present as a minority in the patient serum, although if this were the case it would be unexpected to find no differences in the HVR1. In a report directly relevant to this study, in which the genotype 1 structural region was amplified from serum and expressed, Forns *et al.* (1997) reported that there was evidence for selection of defective genomes during the cloning procedure. All of their clones were dysfunctional for polyprotein production. It was unlikely that such a high proportion of defective genomes was present in the viral population and the authors concluded that the explanation was due to spurious translation from within the HCV sequence, resulting in a protein product that is toxic to *E.coli*. By introducing extra nucleotides directly upstream of the HCV start codon, the proportion of functional clones was increased. It is not known if the same explanation is relevant to this present study as too few clones were analysed for protein expression. In fact, in our study, it was the ability to obtain a clone with an insert in the correct orientation that was more problematic. This could also reflect the bias of *E.coli* for translationally defective clones.

Having corrected all 9 differences and cloned the inserts into pSFV1 once again, the resulting construct, pSFVR3CE1E2, was sequenced across core, E1 and E2. All nine amino acid differences had been corrected and no other exchanges had been introduced. Therefore this construct could be used confidently for the purposes of expressing the genotype 3 HCV structural proteins encoded by the genome circulating as the majority in this particular patient.

Chapter 4

Expression and characterisation of the HCV genotype 3 glycoproteins

4.1. Introduction

The Semliki Forest virus (SFV) expression system directs high levels of expression of a heterologous gene product in a wide range of animal cell hosts (Liljestrom & Garoff, 1991). It has advantages over other eukaryotic expression systems, such as the baculovirus and vaccinia systems, in that it does not rely on the construction of recombinant virus stocks. SFV can infect a broad range of cultured mammalian cells, including human, so this expression system is appropriate for the expression of mammalian viral glycoproteins.

The SFV vector, pSFV1, is a modified cDNA copy of the viral genome. The region encoding the structural proteins (produced by the 26S RNA) has been replaced by a polylinker containing a *Bam*HI-*Xma*I-*Sma*I polylinker sequence and three stop codons downstream of the 26S promoter (Figure 3.5). The gene of interest is inserted into one of these unique restriction sites. The recombinant SFV genome is synthesised by *in vitro* transcription and electroporated into cells. The 5' and 3' untranslated regions and the nsP1-4 genes, which encode the RNA replicase complex, are retained and can therefore drive the replication and transcription of recombinant SFV RNA. The translational machinery of the host cell is responsible for protein synthesis and, due to the high RNA

levels produced by the SFV replicase complex, synthesis of the viral proteins overtakes host protein synthesis, with levels of the heterologous protein reaching 25% of total cellular protein 75 hours post infection (Liljestrom & Garoff, 1991). The SFV expression system has been used successfully for the expression of human, murine, avian, bovine, bacterial and viral proteins (Liljestrom & Garoff, 1991, Ciccaron *et al.*, 1993, Sjoberg *et al.*, 1994). Viral proteins include herpes simplex virus-encoded DNase (Kehm *et al.*, 1998), influenza nucleoprotein (Berglund *et al.*, 1999), HIV gp120 and gp41 (Altmeyer *et al.*, 1999), human papilloma virus E6 and E7 proteins (Daemen *et al.*, 2000) and the HCV genotype 1 structural proteins (Patel *et al.*, 1999a & 2001).

Numerous studies have reported complex formation between the HCV glycoproteins, E1 and E2 (Grakoui *et al.*, 1993a, Ralston *et al.*, 1993, Selby *et al.*, 1994). Due to lack of a cell culture system for HCV and the low levels of virus particles in patient sera, the only means to study the glycoprotein complexes has been through expression from recombinant cDNA templates in a variety of systems. These include vaccinia virus (Ralston *et al.*, 1993, Dubuisson *et al.*, 1994), Baculovirus (Baumert *et al.*, 1998) Sindbus virus (Dubuisson *et al.*, 1994) and Semliki Forest virus expression systems (Patel *et al.*, 1999a). All of these studies have reported the presence of two forms of the E1E2 complex. They form a heterodimer that is stabilised by non-covalent interactions and a high molecular weight, disulphide-linked aggregate (Dubuisson *et al.*, 1994). The non-covalent interactions are thought to be authentic and represent the native prebudding form of the glycoprotein complex (Deleersnyder *et al.*, 1997). The aggregate is thought to be formed by the misfolding of the proteins but it is unclear if these complexes are formed during the natural HCV life cycle or whether they are an artefact of the expression systems used. These observations are all based on analysis of genotype 1 E1 and E2 proteins and before firm conclusions can be drawn on the authentic formation of the glycoprotein complex in the virion, it is important to see if these findings apply to other genotypes.

4.2. Reactivity of genotype 1 antibodies against genotype 3 glycoproteins

4.2.1. Expression of genotype 3 structural proteins using the SFV system

For expression of recombinant proteins with the SFV system, the gene of interest is first cloned into the SFV plasmid vector, pSFV1. This process was described in chapter 3 for the construction of pSFVF2CE1E2, pSFVR3CE1E2 and pSFVR3E2, which encode the genotype 3 proteins. For comparative purposes, SFV constructs encoding the equivalent genotype 1 proteins were kindly provided by Janisha Patel (MRC Virology Unit, Glasgow). These constructs are pSFVH77CE1E2 and pSFVH77E2, which encode the structural proteins of the infectious genotype 1a clone, H77 (Yanagi *et al.*, 1997) and pSFVGlaCE1E2 which encodes the structural proteins of a local genotype 1a strain known as HCV-1a Gla (Gla) (Patel *et al.*, 1999a & 2000). A summary of these constructs and the proteins they encode can be seen in Table 4.1.

In preparation for *in vitro* transcription, the SFV constructs were linearised by cleavage with restriction endonuclease *Spe*I. There was a unique *Spe*I site beyond the *BamH*I cloning site in pSFV1 (see Figure 3.4) and there were no *Spe*I sites in the genes encoding the structural proteins of either genotype 1 or 3. This linearised DNA served as a template for *in vitro* synthesis of recombinant RNA directed by SP6 RNA polymerase. One twelfth of the total recombinant RNA was analysed by agarose gel electrophoresis (Figure 4.1). The template plasmid DNA was visible as a faint band at approximately 13.5 kb for the recombinant plasmids (lanes 1 and 2) and 11 kb for the SFV vector (lane 3). The brighter bands below this were the transcribed RNA molecules and the smear below this main RNA band was due to RNA degradation products. The remainder of the transcription reaction was electroporated into mammalian cells for expression of the recombinant protein. Initially Baby Hamster Kidney (BHK) cells were used, as these

Table 4.1. Summary of SFV constructs.

SFV construct	Encoded protein	Genotype	Strain	comment
pSFVF2CE1E2	Core, E1, E2	3a	Gla-3a	9 aa differences from R3
pSFVR3CE1E2	Core, E1, E2	3a	Gla-3a	Corrected sequence matches PCR majority sequence
pSFVR3E2	E2	3a	Gla-3a	Subcloned from pSFVR3CE1E2
pSFVH77CE1E2	Core, E1, E2	1a	H77	From infectious clone
pSFVH77E2	E2	1a	H77	From infectious clone
pSFVGlaCE1E2	Core, E1, E2	1a	Gla	

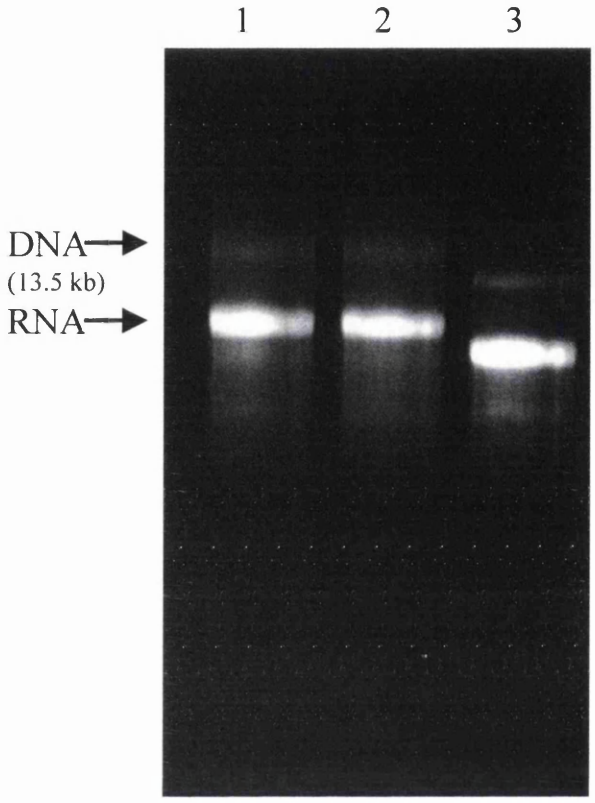


Figure 4.1. *In vitro* transcribed RNA from pSFVR3CE1E2 (lane 1), pSFVH77CE1E2 (lane 2) and pSFV1 (lane 3). Linearised plasmid DNA was transcribed *in vitro* from an SP6 promoter and 2 µl of the reaction was separated on a 1% agarose gel. A small amount of the DNA template is visible along with the transcribed RNA. The smear underneath the main RNA band is indicative of degradation products. The remainder of the transcription reaction was electroporated into cells for protein expression.

were the cells used in the first description of this expression system by Liljestrom & Garoff (1991). However, human hepatocytes are a more relevant cell type for the expression of HCV proteins and hence, conditions were adjusted to allow electroporation of the constructs into the human hepatocellular carcinoma cell line, Huh7.

4.2.2. Recognition of genotype 3 E2 by genotype 1 E2 antibodies

There have been no previous reports describing expression of the genotype 3 envelope proteins and as far as is known, there are no genotype 3 specific antibodies available. A large panel of genotype 1 specific antibodies to E1 and E2 had been generated by Dr Arvind Patel (MRC Virology Unit, Glasgow) (Patel *et al.*, 2000, Owsianka *et al.*, 2001) and these were kindly made available for this study. The majority of these antibodies were raised against the Gla strain proteins and not all recognised strain H77 E2 (A. Patel, personal communication). Attention was focussed on those antibodies that did recognise H77 E2, as RNA from this strain had been shown to be infectious. Three genotype 1 anti-E1 antibodies, raised against the Gla strain, were also available. Details of the antibodies tested are shown in Table 4.2. The reactivity of these genotype 1 antibodies against genotype 3 was analysed in a variety of immunological assays.

Western blot analysis identified four antibodies that recognised both type 1 and type 3 E2 (Figure 4.2A). They were ALP98 (a), AP266 (b), ALP1 (c) and R646 (d). All four antibodies recognised a protein of approximately 66 kDa which was specific to cells expressing the structural proteins, not being seen in the lanes containing the SFV vector only (lane 5) or Huh7 cells only (lane 6). Interestingly, ALP1 failed to recognise F2 E2 and AP266 had less affinity for F2 E2 when compared to R3 E2. The polyclonal antibody, R646, had more activity against the type 1 proteins (H77 and Gla) than the type 3 proteins (R3 and F2) and many additional cellular proteins were also recognised by this antibody. This was possibly due to over-exposure to allow for detection of genotype 3 E2. In addition to the main protein band, ALP98 also recognised protein bands at

Table 4.2. Details of genotype 1 E1 and E2 antibodies used in this study

Antibody	Type	Epitope Type	Specificity	Dilution	Original Source
ALP98*	Mouse monoclonal	linear	E2 (aa 650-657)	1:1000	A. Patel
AP266*	Mouse monoclonal	linear	E2 (aa 650-657)	1:10	A. Patel
ALP1*	Mouse monoclonal	linear	E2 (aa 654-665)	1:10	A. Patel
AP320*	Mouse monoclonal	linear	E2 (aa 464-471)	1:500	A. Patel
AP33*	Mouse monoclonal	linear	E2 (aa 412-423)	1:500	A. Patel
R646§	Rabbit polyclonal	linear	E2	1:1000	A. Patel
H30#	Mouse monoclonal	conform-ational	E2	1:200	J. Dubuisson
H50#	Mouse monoclonal	conform-ational	E2	1:200	J. Dubuisson
H53#	Mouse monoclonal	conform-ational	E2	1:200	J. Dubuisson
H60#	Mouse monoclonal	conform-ational	E2	1:200	J. Dubuisson
H61#	Mouse monoclonal	conform-ational	E2	1:200	J. Dubuisson
AP21.010*	Mouse monoclonal	linear	E1 (aa 213-235)	1:500	A. Patel
AP119*	Mouse monoclonal	linear	E1 (aa 235-260)	1:500	A. Patel
R528*	Rabbit polyclonal	linear	E1	1:1000	A. Patel

* raised against HCV strain Gla (Patel *et al.*, 2000, Owsianka *et al.*, 2001)

§ raised against HCV strain H77 (Owsianka *et al.*, 2001)

raised against strain HCV-H (Cocquerel *et al.*, 1998)

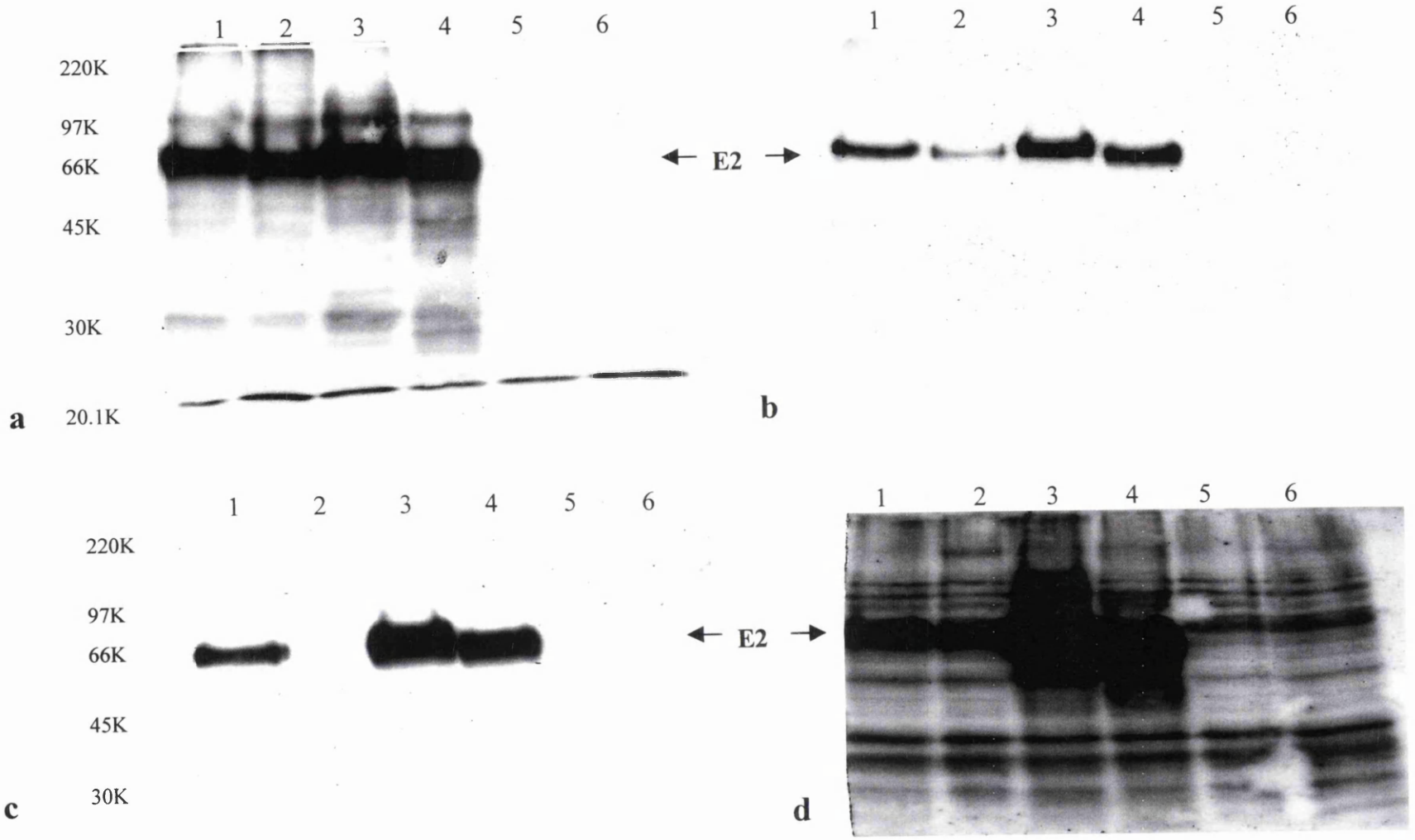

1 2 3 4 5 6
220K
97K
66K
45K
30K
20.1K
a
1 2 3 4 5 6
E2
b
1 2 3 4 5 6
220K
97K
66K
45K
30K
c
E2
1 2 3 4 5 6
d

Figure 4.2A. Reactivity of type 1 E2 antibodies against type 3 E2 as determined by Western blot analysis. RNA was transcribed *in vitro* from pSFVR3CE1E2 (lane 1), pSFVF2CE1E2 (lane 2), pSFVH77CE1E2 (lane 3), pSFVGlaCE1E2 (lane 4) and pSFV1 (lane 5) and was electroporated into Huh7 cells. Lane 6 contains extract from mock electroporated cells. Proteins were separated by SDS-PAGE and analysed by Western blot using mouse monoclonals ALP98 (a), AP266 (b), ALP1 (c) and rabbit polyclonal R646 (d). The position and size of E2 is indicated.

approximately 97 kDa and 30 kDa (although to a much lesser extent). These bands were specific to E2-expressing cells. The ~21 kDa protein band was present in all 6 lanes, indicating it was a cellular protein. Neither AP266 nor ALP1 recognised any additional proteins. Close inspection of the migration patterns of the E2 proteins, revealed that there may be slight differences in mobility between the E2 proteins of different constructs. Two antibodies that recognised only genotype 1 E2 are shown in Figure 4.2B. Western blot with AP320 (a) indicated a major protein band in the type 1 E2-expressing cells only. There was a minor band at the same position in all the other lanes, which indicated recognition of a cellular protein. AP33 (b) recognised only H77 E2. The antibody was raised against Gla E2 and did recognise this protein as well but the pSFVGla$_{CE1E2}$ construct used here had a tag that disrupted the epitope recognised by AP33.

Western blot analysis detects antibodies that recognise denatured proteins. Some antibodies may recognise an epitope that is conformational and would be destroyed when the protein is denatured. To detect reactivity of such an antibody, an alternative assay, such as immunofluorescence can be used, in which the native structure of proteins may be retained. To verify the reactivity of the antibodies identified by Western blot analysis, they were also analysed by indirect immunofluorescence and the results are summarised in Table 4.3. The results were consistent with those for Western blot analysis. Those antibodies (AP320 & AP33) which did not recognise type 3 E2 by Western blot did not recognise type 3 E2 by immunofluorescence either. The patterns of recognition by ALP98 and R646 are shown in Figure 4.3 A and B, respectively. ALP98 recognised E2 of R3 (a), F2 (b) and H77 (c) but no specific fluorescence was visible in cells expressing the SFV proteins only (d) or Huh7 cells only (e). The pattern of fluorescence was cytoplasmic and had a web-like nature. Similar fluorescence patterns were seen with all reactive antibodies. In Western blot analysis, R646 showed more reactivity to type 1 E2 and this was even more pronounced in immunofluorescence assays (Figure 4.3B). In fact, the reactivity against type 3 E2 was scored as negative (a). If the concentration of the antibody was increased, the same cytoplasmic staining pattern could be seen with type 3 E2 through the general high background fluorescence. It was concluded that R646 was not an optimal antibody for detecting type 3 E2 in an immunofluorescence assay.

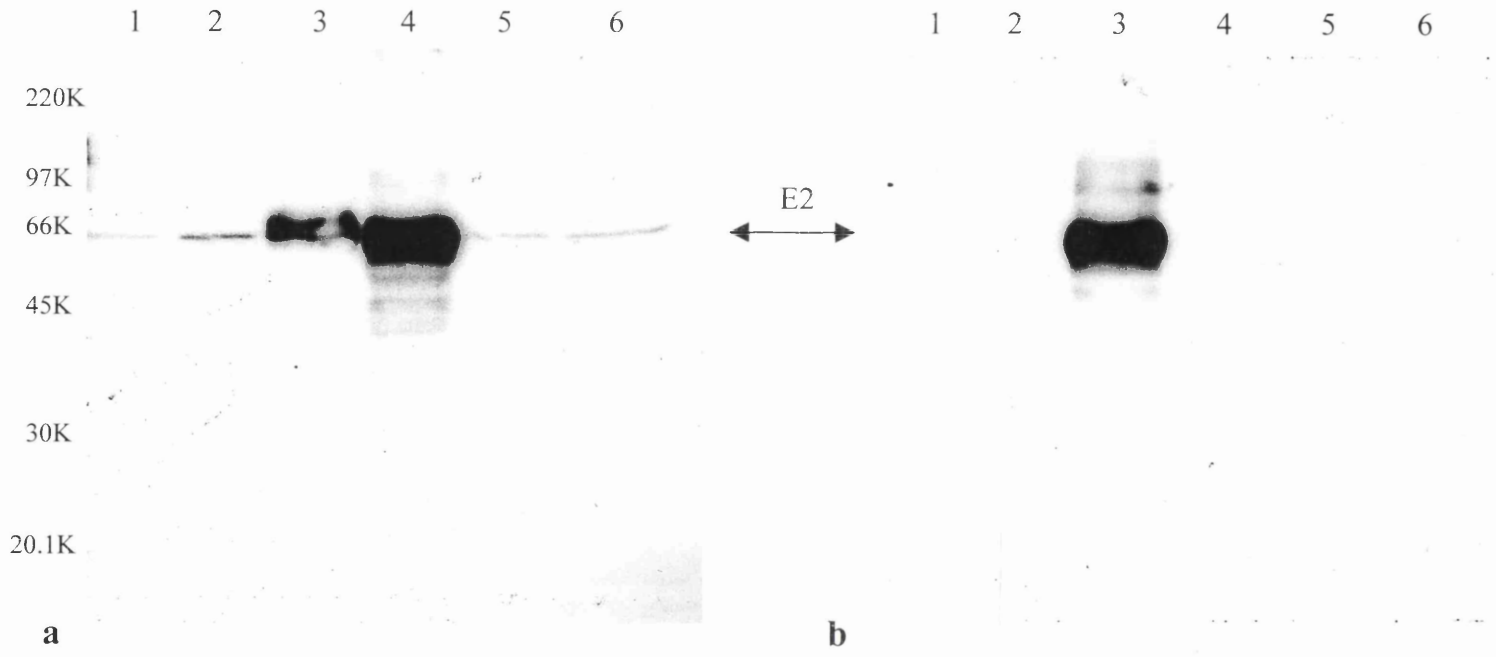


Figure 4.2B Western blot analysis of genotype 1 E2 antibodies which do not cross-react with genotype 3 E2. RNA was transcribed *in vitro* from pSFVR3CE1E2 (lane 1), pSFVF2CE1E2 (lane2), pSFVH77CE1E2 (lane3), pSFVGlaCE1E2 (lane 4) and pSFV1 (lane 5) and was electroporated into Huh7 cells. Lane 6 contains extract from mock electroporated cells. Proteins were separated by SDS-PAGE and analysed by Western blot using mouse monoclonals AP320 (a) or AP33 (b). The position and size of E2 is indicated.

Table 4.3. Summary of reactivity of genotype 1 E2 antibodies against both genotype 1 and genotype 3 E2 as determined by indirect immunofluorescence.

Antibody	Genotype 1 E2		Genotype 3 E2	
	H77	Gla	R3	F2
ALP98 (MAb)	++	++	++	++
AP266 (MAb)	++	++	++	++
ALP1 (MAb)	++	++	++	-
AP33 (MAb)	++	-	-	-
AP320 (MAb)	++	++	-	-
R646 (PolyAb)	++	++	+/-	+/-

Table 4.4. Summary of reactivity of genotype 1 E1 antibodies against both genotypes 1 and 3 E1 as determined by indirect immunofluorescence.

Antibody	Genotype 1 E1 - Gla	Genotype 3 E1 – F2
21.010	++	-
AP119	++	-
R528	++	-

"++" indicates positive reactivity

"-" indicates negative reactivity

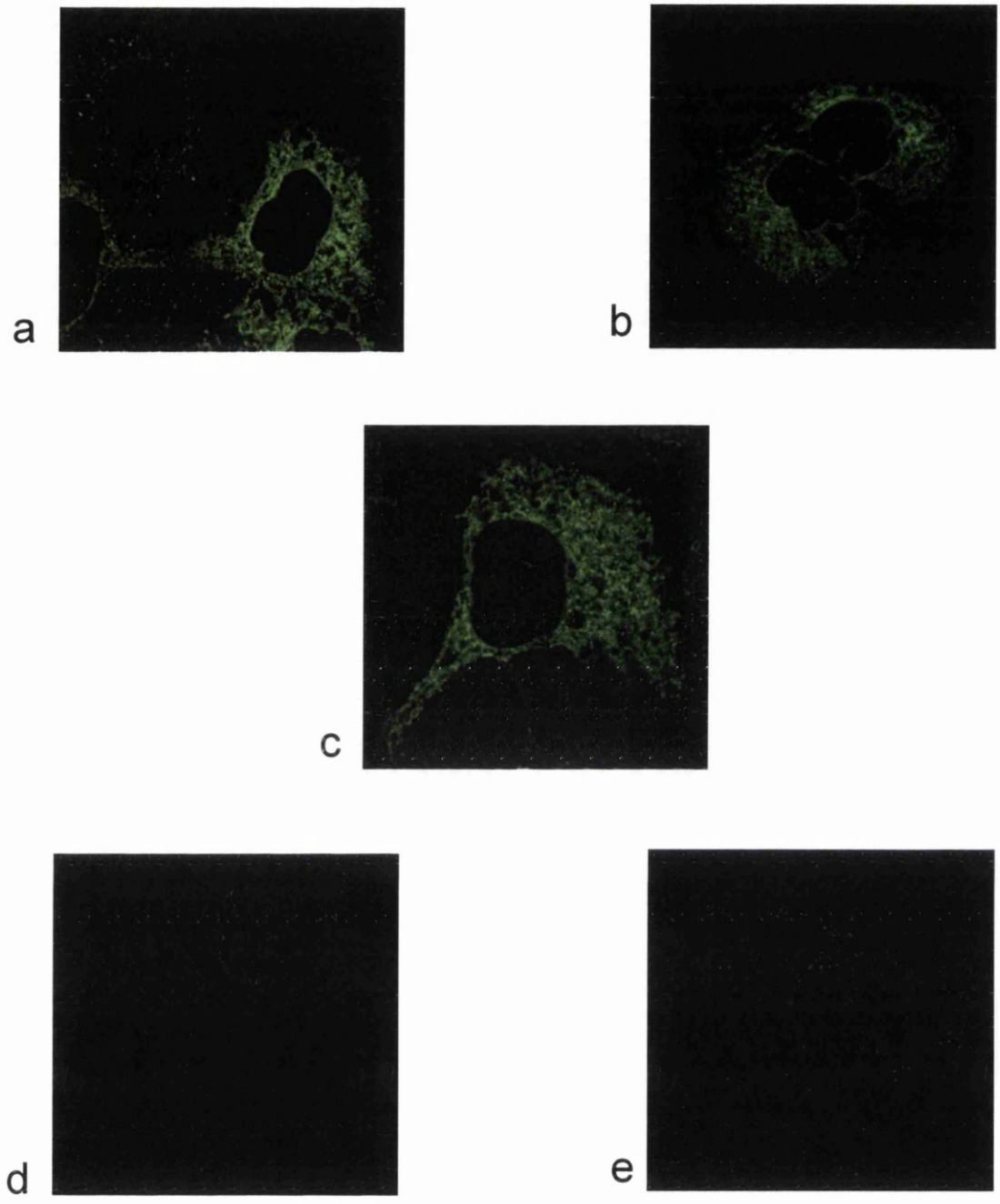


Figure 4.3A Reactivity of anti-E2, ALP98, against genotype 1 and genotype 3 E2 by indirect immunofluorescence in Huh7 cells. Huh7 cells were electroporated with *in vitro* transcribed RNA from pSFVR3CE1E2 (a), pSFVF2CE1E2 (b), pSFVH77CE1E2 (c) and pSFV1 (d). Panel (e) shows mock electroporated cells. Cells were fixed in methanol and prepared for indirect immunofluorescence with ALP98 and anti-mouse FITC. Images were captured with a 63X oil immersion lens.

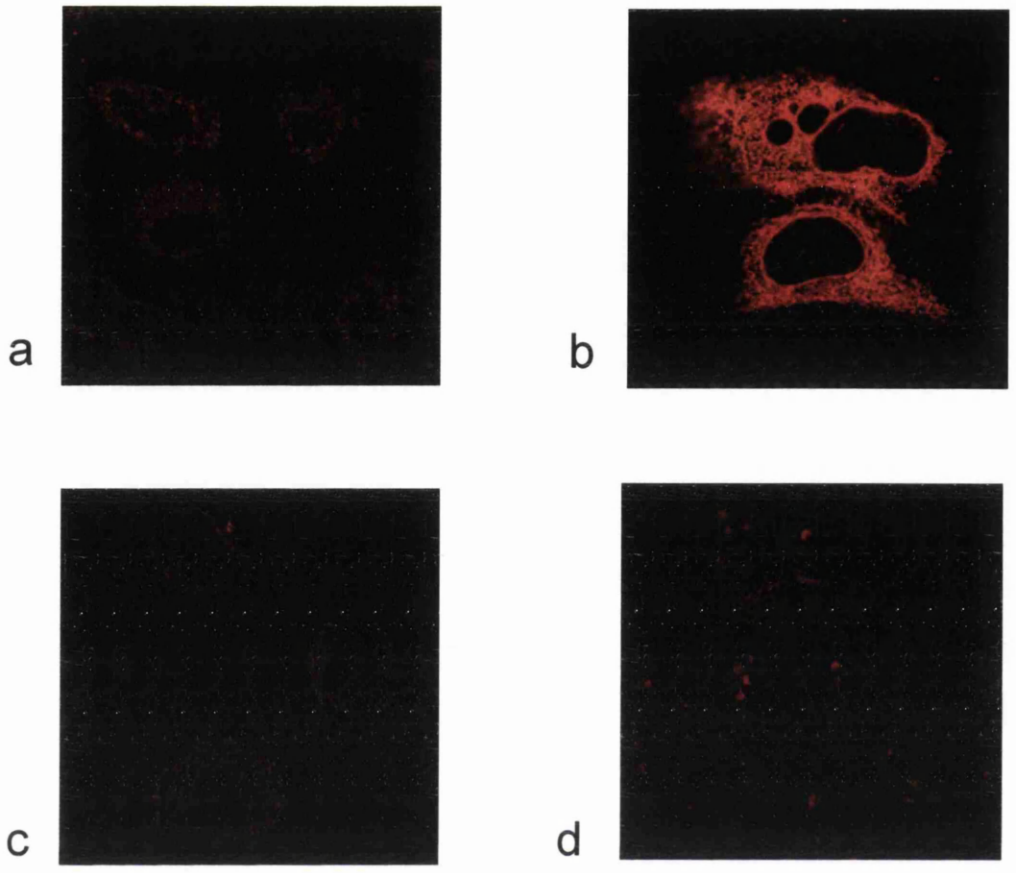
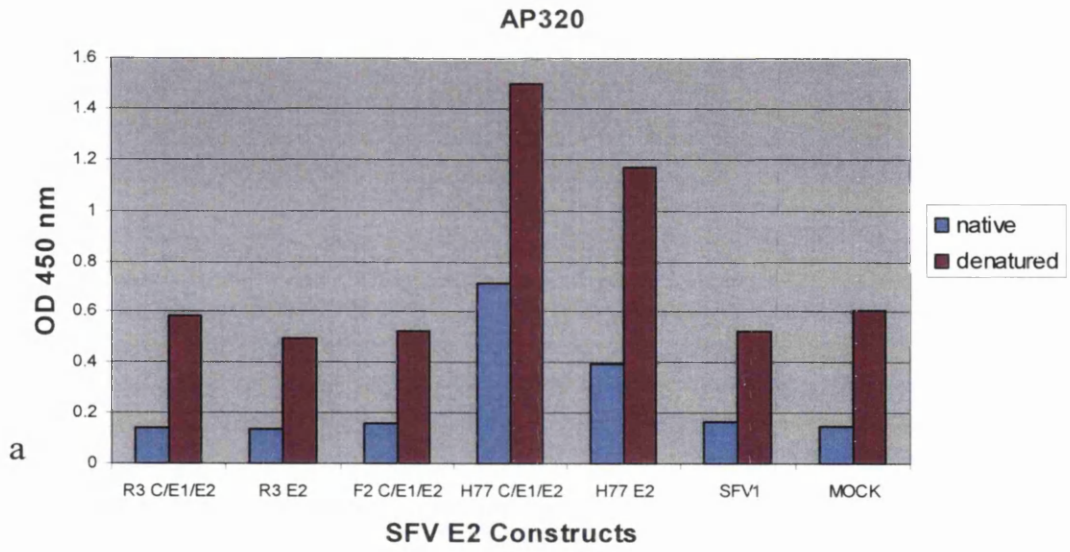


Figure 4.3B Reactivity of anti-E2, R646, to genotype 1 and genotype 3 E2 by indirect immunofluorescence in Huh7 cells. Huh7 cells were transfected with pSFVR3CE1E2 (a), pSFVH77CE1E2 (b) or pSFV1 (c) and fixed in methanol before being prepared for indirect immunofluorescence with R646 and anti-rabbit Cy3. Panel (d) shows mock transfected cells. Images were captured using a 63X oil immersion lens.

The reactivity of three genotype 1 E1 antibodies against genotype 3 E1 was also analysed by immunofluorescence (Table 4.4). Neither of the monoclonal antibodies (21.010 and AP119) nor the polyclonal antibody (R528) were able to recognise type 3 E1.

As a final test of antibody reactivity, both native and denatured forms of E2 were captured by GNA and the ability of the antibodies to recognise the different forms was analysed by ELISA. In addition to those constructs expressing all the structural proteins (i.e. core, E1, E2), the constructs expressing E2 only were also included to examine whether antibody reactivity was modified by the presence or absence of core and E1. *Galanthus nivalis* agglutinin (GNA) is a lectin that captures the carbohydrate portion of the glycoprotein. Prior to GNA capture, half of the E2-expressing cellular extract was treated with SDS and DTT to denature the protein. Both native and denatured E2 were bound to GNA and antibody reactivity was measured in terms of absorbance at 450 nm. These figures were plotted for each E2 construct and are shown in Figure 4.4 A and B. The ELISA results for those antibodies that were previously shown not to recognise genotype 3 E2 are shown in Figure 4.4A. For both AP320 (a) and AP33 (b), the absorbance readings for the genotype 3 proteins (R3 and F2) were in the same range as those of the SFV vector or cells only. Therefore, as with the Western blot and immunofluorescence assays, AP320 and AP33 did not recognise genotype 3 E2 in ELISA. The reactivities of those antibodies that were previously shown to recognise type 3 E2 are shown in Figure 4.4B. As before, ALP98 (a), ALP1 (b) and AP266 (c) recognised both genotype 1 and genotype 3 E2 in ELISA. ALP98 showed similar levels of reactivity to E2 of both genotypes, whether expressed alone or along with core and E1. ALP1 did not recognise genotype 3 F2 E2, which agreed with the previous Western blot and immunofluorescence findings. With ALP1, similar absorbance readings were seen for the equivalent genotype 3 (R3) and genotype 1 (H77) constructs, whereas with AP266, higher readings were seen with genotype 3 than genotype 1. The same cell extracts were used for all ELISAs in this figure, therefore the degree of reactivity of different antibodies against proteins of the same construct could be compared. With both ALP1 and AP266, the absorbance readings for E2 expressed alone were lower than for E2 co-expressed with core and E1. This applied to both type 1 and type 3 but it was not possible to guarantee that the expression

a


AP320
OD 450 nm
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
native
denatured
R3 C/E1/E2
R3 E2
F2 C/E1/E2
H77 C/E1/E2
H77 E2
SFV1
MOCK
SFV E2 Constructs

b

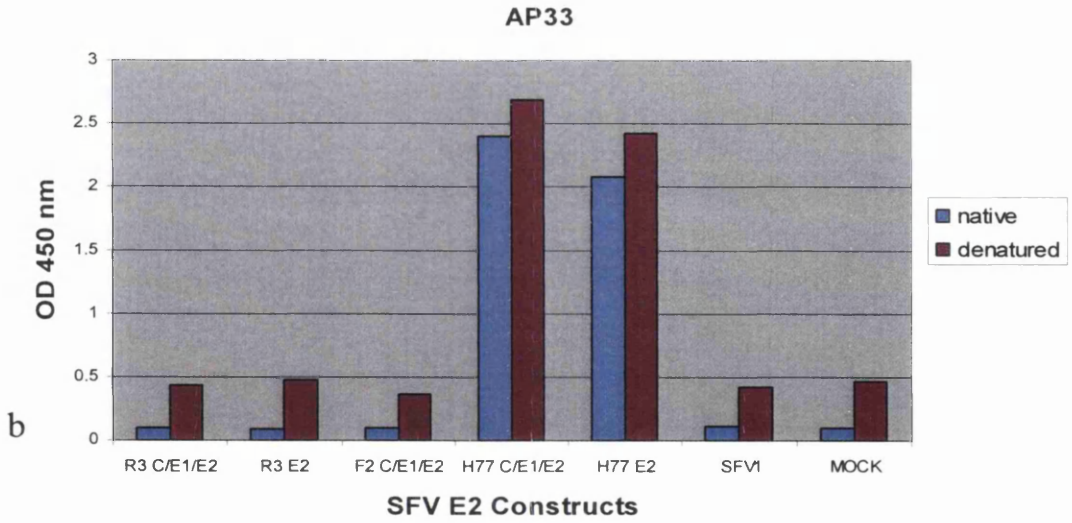

AP33
OD 450 nm
3
2.5
2
1.5
1
0.5
0
native
denatured
R3 C/E1/E2
R3 E2
F2 C/E1/E2
H77 C/E1/E2
H77 E2
SFV1
MOCK
SFV E2 Constructs

Figure 4.4A Genotype 1 E2 antibodies which do not recognise genotype 3 E2. E2 expressing cell extracts were prepared from BHK cells which had been electroporated with *in vitro* transcribed RNA from pSFVR3$_{CE1E2}$, pSFVR3$_{E2}$, pSFVF2$_{CE1E2}$ (all genotype 3), pSFVH77$_{CE1E2}$, pSFVH77$_{E2}$ (genotype 1) and pSFV1. Denatured extracts were treated with SDS and DTT, while the native extracts were untreated. E2 was captured in GNA-coated wells and anti-E2, AP320 (a) or AP33 (b) was added. Antibody reactivity was measured by the addition of anti-mouse HRP conjugate and TMB substrate. The reaction was stopped by the addition of H_2SO_4 and the absorbance at 450 nm was measured.

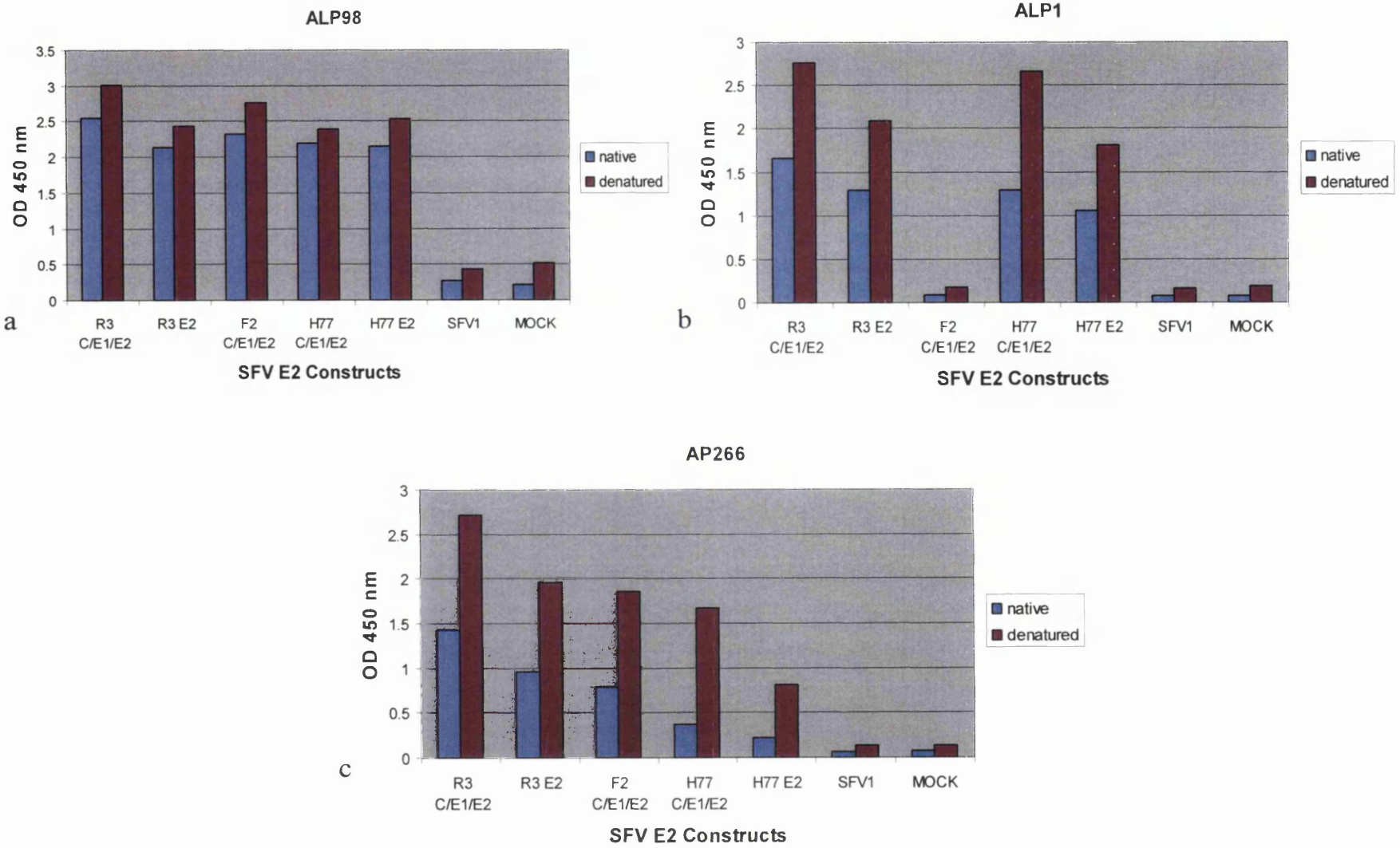

ALP98
OD 450 nm
3.5
3
2.5
2
1.5
1
0.5
0
native
denatured
R3 C/E1/E2
R3 E2
F2 C/E1/E2
H77 C/E1/E2
H77 E2
SFV1
MOCK
SFV E2 Constructs
a
ALP1
OD 450 nm
3
2.5
2
1.5
1
0.5
0
native
denatured
R3 C/E1/E2
R3 E2
F2 C/E1/E2
H77 C/E1/E2
H77 E2
SFV1
MOCK
SFV E2 Constructs
b
AP266
OD 450 nm
3
2.5
2
1.5
1
0.5
0
native
denatured
R3 C/E1/E2
R3 E2
F2 C/E1/E2
H77 C/E1/E2
H77 E2
SFV1
MOCK
SFV E2 Constructs
c

Figure 4.4B Genotype 1 E2 antibodies which recognise both genotype 1 and 3 E2. E2 expressing cell extracts were prepared from BHK cells which had been electroporated with *in vitro* transcribed RNA from pSFVR3$_{CE1E2}$, pSFVR3$_{E2}$, pSFVF2$_{CE1E2}$ (all genotype 3), pSFVH77$_{CE1E2}$, pSFVH77$_{E2}$ (genotype 1) and pSFV1. Denatured extracts were treated with SDS and DTT, while the native extracts were untreated. E2 was captured in GNA-coated wells and anti-E2, ALP98 (a), ALP1 (b) or AP266 (c) was added. Antibody reactivity was measured by the addition of anti-mouse HRP conjugate and TMB substrate. The reaction was stopped by the addition of H_2SO_4 and the absorbance at 450 nm was measured.

levels of all the constructs were equivalent. The results shown in this figure were from one experiment but this was repeated 3 times from the protein expression stage. As equivalent levels of expression could not be ensured, the results of these 3 experiments could not be incorporated into one figure but the other experiments all showed similar trends to those reported above.

None of the antibodies tested above displayed any tendency towards recognising conformational epitopes only. It would be of interest to know if a conformation-dependent antibody would be able to recognise both genotypes 1 and 3 E2, even if there are differences in the primary sequence. Small quantities of known conformation-dependent E2 antibodies were available from J. Dubuisson. These were raised against genotype 1 (strain HCV-H) E2 as described by Deleersnyder *et al.* (1997). The reactivity of these antibodies against native forms of both genotypes 1 and 3 E2 was analysed by ELISA (Figure 4.5). A non-conformation-dependent antibody, AP266, was included and reactivity with all E2-expressing extracts indicated that expression levels were optimal. However, the conformational antibodies recognised only H77 E2. Reactivity was higher with H50, H53 and H60 than H30 and H61. Due to limitations on the quantities of these antibodies, it was only possible to repeat this experiment twice, each from the protein expression stage. Both experiments showed reactivity with H77 E2 only.

The epitopes of some of these antibodies had been mapped by Arvind Patel, using peptides (Figure 4.6) based on the amino acid sequence of the HCV genotype 1a strains Gla or HCV-H (personal communication). The amino acid sequences of E1 and E2 of genotype 1 (H77) and genotype 3 (R3) are aligned in Figure 4.6. The proposed antibody epitopes were marked on the sequence and distinction was made between those antibodies that recognised both types 1 and 3 and those that recognised only type 1. The epitopes of ALP98 and AP266 (aa 650-657) were identical and conserved between R3 and H77. ALP1 recognised an overlapping epitope (aa 654-665) and there was one difference between R3 and H77 within this region. R3 had an isoleucine residue at position 660 compared to a leucine residue in H77. There was also only one amino acid difference in the epitope recognised by AP33 (aa 412-423). This was a lysine residue at position 415

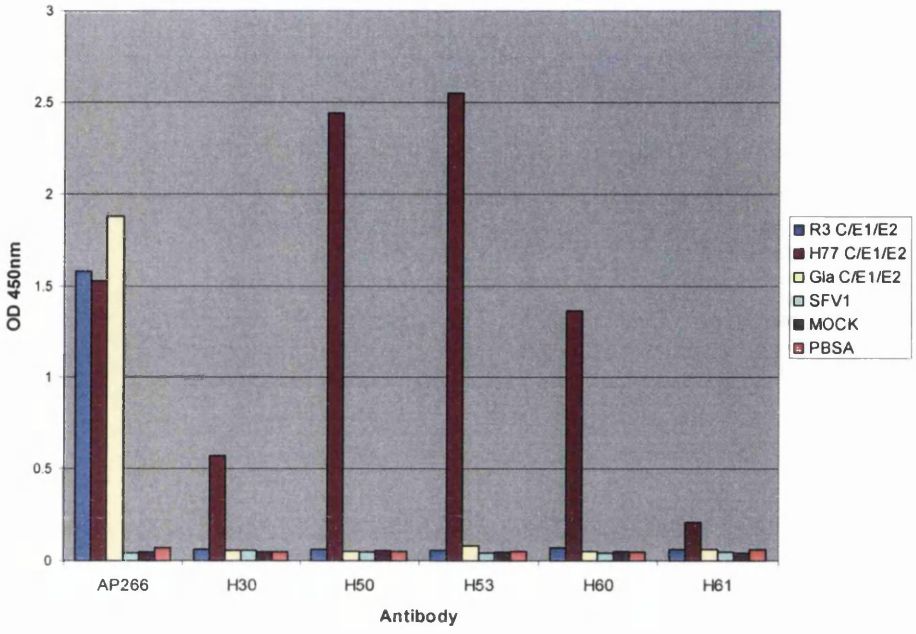
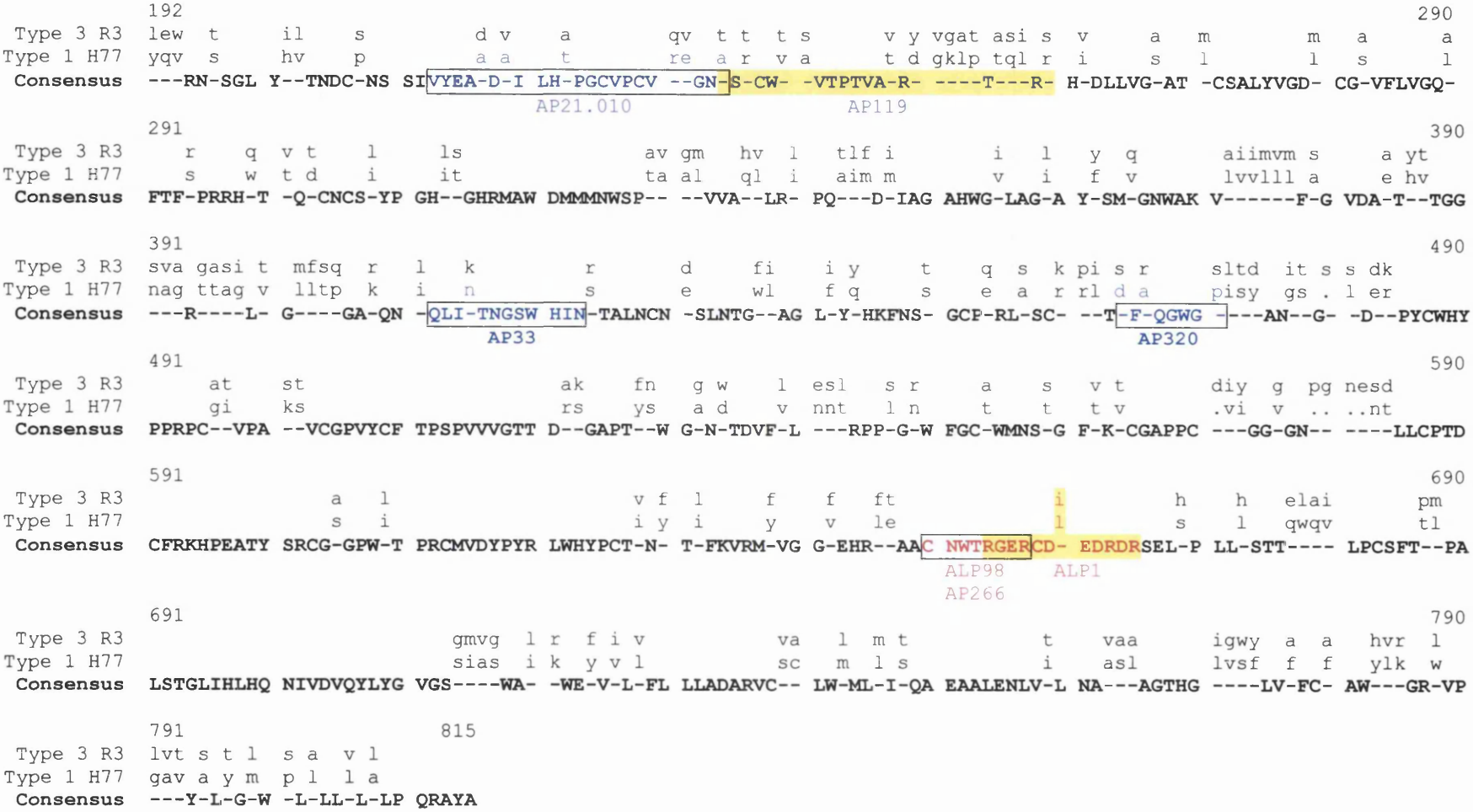


Figure 4.5. Reactivity of genotype 1 conformation-dependent antibodies against genotype 1(H77 &Gla) and 3 (R3) E2. E2-expressing cell extracts were prepared by electroporating BHK cells with *in vitro* transcribed RNA from pSFVR3CE1E2 (type 3), pSFVH77CE1E2, pSFVGlaCE1E2 (both type 1) and pSFV1. Native E2 was captured in GNA-coated wells and linear-epitope antibodies (AP266) or conformational-epitope antibodies (H30, H50, H53, H60 & H61) were added. Antibody reactivity was measured by the addition of anti-mouse HRP conjugate and TMB substrate. The reaction was stopped by the addition of H_2SO_4 and the absorbance measured at 450 nm.


192
290
Type 3 R3 lew t il s d v a qv t t t s v y vgat asi s v a m m a a
Type 1 H77 yqv s hv p a a t re a r v a t d gklp tql r i s l l s l
Consensus ---RN-SGL Y--TNDC-NS SIVYEA-D-I LH-PGCVPCV --GN-S-CW- -VTPTVA-R- ----T---R- H-DLLVG-AT -CSALYVGD- CG-VFLVGQ-
AP21.010
AP119
291
390
Type 3 R3 r q v t l ls av gm hv l tlf i i l y q aiimvm s a yt
Type 1 H77 s w t d i it ta al ql i aim m v i f v lvvlll a e hv
Consensus FTF-PRRH-T -Q-CNCS-YP GH--GHRMAW DMMMNWSP-- --VVA--LR- PQ---D-IAG AHWG-LAG-A Y-SM-GNWAK V------F-G VDA-T--TGG
391
490
Type 3 R3 sva gasi t mfsq r l k r d fi i y t q s k pi s r sltd it s s dk
Type 1 H77 nag ttag v lltp k i n s e wl f q s e a r rl d a pisy gs . l er
Consensus ---R----L- G----GA-QN -QLI-TNGSW HIN-TALNCN -SLNTG--AG L-Y-HKFNS- GCP-RL-SC- --T-F-QGWG -----AN--G- -D--PYCWHY
AP33
AP320
491
590
Type 3 R3 at st ak fn g w l esl s r a s v t diy g pg nesd
Type 1 H77 gi ks rs ys a d v nnt l n t t t v .vi v .. ..nt
Consensus PPRPC--VPA --VCGPVYCF TPSPVVVGTT D--GAPT--W G-N-TDVF-L ---RPP-G-W FGC-WMNS-G F-K-CGAPPC ---GG-GN-- ----LLCPTD
591
690
Type 3 R3 a l v f l f f ft i h h elai pm
Type 1 H77 s i i y i y v le l s l qwqv tl
Consensus CFRKHPEATY SRCG-GPW-T PRCMVDYPYR LWHYPCT-N- T-FKVRM-VG G-EHR--AAC NWTRGERCD- EDRDRSEL-P LL-STT---- LPCSFT--PA
ALP98
ALP1
AP266
691
790
Type 3 R3 gmvg l r f i v va l m t t vaa igwy a a hvr l
Type 1 H77 sias i k y v l sc m l s i asl lvsf f f ylk w
Consensus LSTGLIHLHQ NIVDVQYLYG VGS----WA- -WE-V-L-FL LLADARVC-- LW-ML-I-QA EAALENLV-L NA---AGTHG ----LV-FC- AW---GR-VP
791
815
Type 3 R3 lvt s t l s a v l
Type 1 H77 gav a y m p l l a
Consensus ---Y-L-G-W -L-LL-L-LP QRAYA

Antibody	Specificity	Epitope tested (strain Gla)
AP21.010	Type 1 E1 only	VYEAVDAILHTPGCVPCVREGNA
AP119	Type 1 (Gla) E1 only	ASRCWVAMTPTVATRDGRLPTTQLRR
AP33	Type 1 E2 only	QLINTNGSWHVN
AP320	Type 1 E2 only	DFDQGWGP
ALP98	Type 1 & 3 E2	CNWTRGER
AP266	Type 1 & 3 E2	CNWTRGER
ALP1	Type 1 & 3 (R3) E2	RGERCDLEDRDR

Figure 4.6. Alignment of derived amino acid sequences from genotype 1 (H77) and genotype 3 (R3) E1 and E2. Dashes indicate non-conserved amino acids and dots represent missing amino acids. The epitopes recognised by the genotype 1 specific antibodies are marked in blue for those that recognise type 1 only and in red for those that recognise both type 1 and type 3. Epitopes were mapped by PEPSCAN using peptides based on the Gla or HCV-H amino acid sequence (Owsianka *et al.*, 2001).

in R3 and an asparagine residue in H77. The epitope recognised by AP320 (aa 464-471) was more variable and there were 3 such differences. At position 464 there was a serine in R3 and an aspartic acid in H77; at position 466 there was an arginine in R3 and an alanine in H77 and finally, at position 471 there was a serine in R3 and a proline in H77. The epitopes of the E1 antibodies had not been mapped down to the minimum number of amino acids, therefore just the general region recognised by each antibody is shown. It was evident that these epitopes for AP21.010 and AP119 were highly variable between genotype 1 and 3. These antibodies were only tested against the genotype 1 Gla isolate and it has since been shown that AP119 does not recognise H77 E1 (A. Patel, personal communication). There are only two differences in the H77 sequence that could account for this loss of reactivity, a valine at position 242 instead of a methionine and a lysine at position 252 instead of an arginine (compare epitope tested to H77 sequence in Figure 4.6).

4.3. Localisation of genotype 3 E2

The immunofluorescence data suggested that genotype 3 E2 showed a similar staining pattern to genotype 1 E2 (Figure 4.3A) which had been shown to localise to the endoplasmic reticulum (ER) (Dubuisson *et al.*, 1994). Calnexin is an ER chaperone protein and can therefore be used as an ER marker. To investigate whether genotype 3 E2 was localised to the ER, Huh7 cells expressing genotypes 1 and 3 E2 were probed with both anti-E2 (ALP98) and anti-calnexin antiserum in an indirect immunofluorescence assay. Localisation of the two proteins was analysed by confocal microscopy (Figure 4.7). The red staining indicated E2 and the green staining indicated calnexin. The E2-expressing cells in panels c, g and k were probed for E2 only, whereas those in panels d, h and l were probed for calnexin only, however both secondary antibodies were used in all reactions. The fact that no staining was visible in panels g and d ensured that there was no cross-reaction of antibodies which could have led to false positives. Cells expressing type 3 E2 (panels a, e and i) and type 1 E2 (panels b, f and j) were probed for both E2 and calnexin and the merged images are shown in panels i and j. The yellow staining indicated regions where E2 and calnexin were found in the same area of the cell i.e.

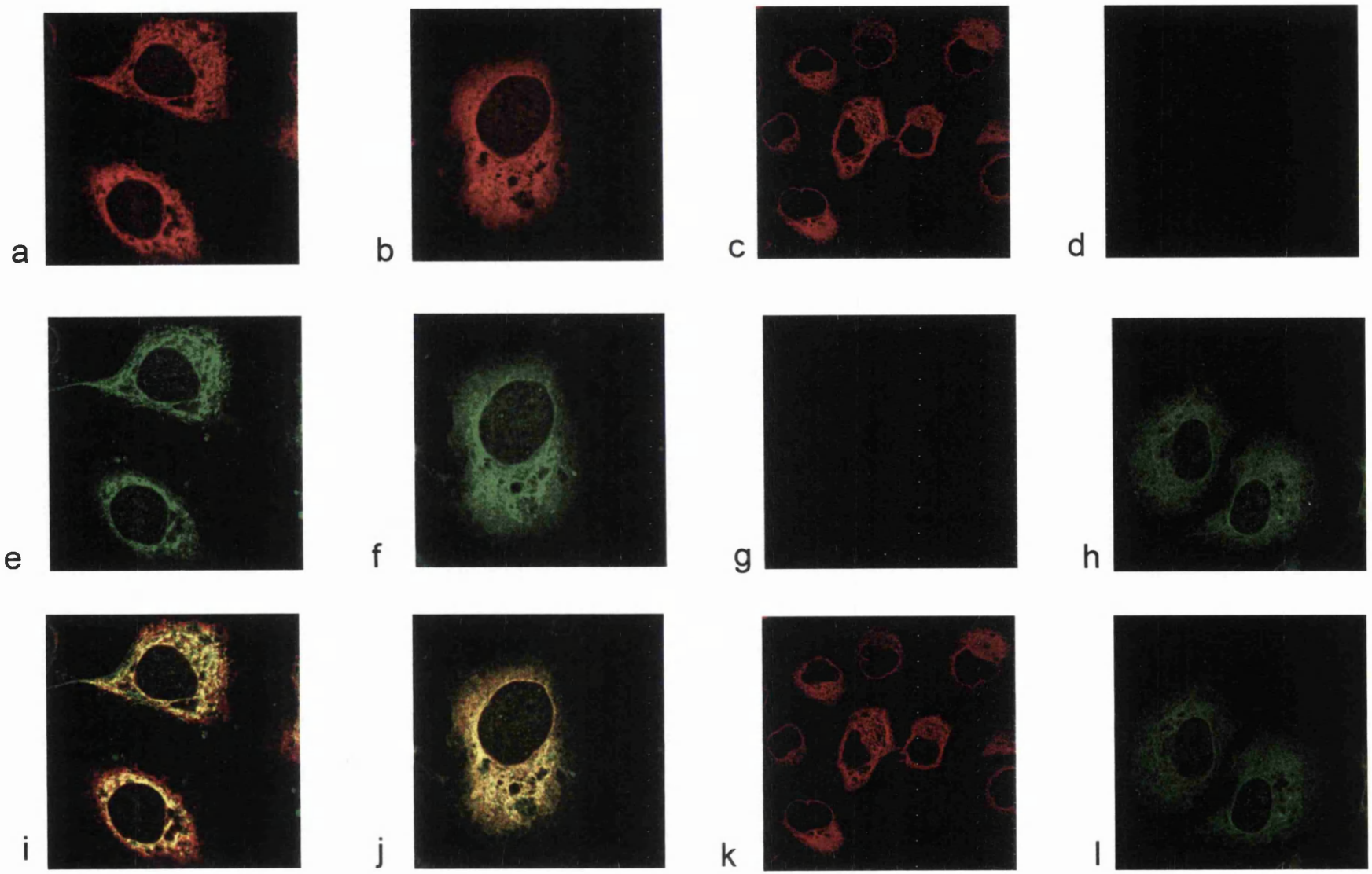

a
b
c
d
e
f
g
h
i
j
k
l

Figure 4.7. Colocalisation of E2 with calnexin. Huh7 cells were electroporated with *in vitro* transcribed RNA from pSFVR3CE1E2 (panels a, e and i) or pSFVH77CE1E2 (remainder of panels). Cells were fixed and processed for indirect immunofluorescence. Cell in panels (a), (b), (e) and (f) were doubly probed with mouse anti-E2 (ALP98) and rabbit anti-calnexin antiserum, whereas cells in panels (c), (g) and (d), (h) were probed singly with mouse anti-E2 (ALP98) and rabbit anti-calnexin antiserum, respectively. In all panels, bound antibodies were detected using goat anti-mouse IgG Cy5 and goat anti-rabbit IgG FITC. Merged images are shown in panels (i), (j), (k) and (l). Images were captured using a 63X or 40X (for c, g, k) oil immersion lens.

regions of colocalisation for E2 and calnexin. Extensive colocalisation was seen for both genotypes 1 and 3 E2 with calnexin.

4.4. Interaction of genotype 3 E1 and E2

4.4.1. Analysis of glycoprotein complex formation

Two methods have been described for distinguishing between the two forms of HCV glycoprotein complex. Some studies have made use of a conformation-sensitive E2 antibody that selectively recognises the native complex (Deleersnyder *et al.*, 1997, Cocquerel *et al.*, 1998). The majority of studies have utilised non-reducing electrophoresis conditions to differentiate between the aggregate and native forms of glycoprotein complex (Dubuisson *et al.*, 1994, Dubuisson & Rice, 1996). In the absence of reducing agents such as β-mercaptoethanol or dithiothreitol (DTT), the disulphide-linked aggregates remain at the top of the gel, while the non-covalently linked proteins, which constitute the native complex, migrate as monomers.

The genotype 3 (R3 and F2) and genotype 1 (H77) structural proteins were expressed in Huh7 cells and proteins were labelled with ^{35}S-methionine. This was particularly necessary for detection of genotype 3 E1 for which there was no available antibody. Radiolabelled type 3 E1 could be detected upon immunoprecipitation with anti-E2 antibody, as opposed to detection with anti-E1 antibody by Western blot analysis. Radiolabelling also allows for detection of all forms of the protein, some of which may not be recognised by a certain antibody. The E2 protein and any interacting proteins, were immunoprecipitated from the cellular extracts with an anti-E2 antibody (ALP98). For analysis of the two forms of glycoprotein complex, the precipitated proteins were divided into two equal portions and one portion was treated with the reducing agent, dithiothreitol (DTT) which reduces disulphide bonds. The remaining precipitated proteins were not treated to allow for distinction between the aggregates and native complex. These reduced and non-reduced proteins were separated by SDS-PAGE and are

shown in Figure 4.8. In lanes 1-3 of the non-reduced samples there was a strong protein band at the top of the gel, which was thought to be the aggregate. Below this, bands representing the monomeric forms of E2 at approximately 66 kDa and E1 at approximately 30 kDa were visible. The E1 and E2 proteins of the different constructs appeared to be slightly different sizes and this will be discussed further in section 4.4.2. This also provided the first proof of E1 expression from the genotype 3 constructs, R3 and F2. In spite of the addition of DTT, this high molecular weight aggregate was still visible in lanes 1-3 of the "reduced" samples. The monomeric forms of E1 and E2 were also visible but there appeared to be a greater amount present here than in the non-reduced samples (compare lane 1 of "non-reduced" with lane 1 of "reduced"). An additional protein band of approximately 97 kDa was present in both non-reduced and reduced samples and the size suggested it was the E1E2 complex. With all three HCV constructs there was a fainter protein band visible below the dominant E1 protein band. This could be the oxidised form of E1 that migrates faster than the reduced form (Dubuisson & Rice, 1996). The ability of the recombinant proteins expressed from the SFV vector to shut off host protein synthesis is evident from the lack of cellular protein in lanes 1-4 compared to lane 5.

The data presented above were obtained from proteins expressed in Huh7 cells. However, when this was repeated in BHK cells, it was noted that there was considerably less formation of native complex. In order to assess this difference properly, the immunoprecipitation experiment was performed in duplicate, in Huh7 cells and BHK cells (Figure 4.9). The non-reduced protein samples are shown in (a) and the DTT-treated reduced samples in (b). In lanes 1-3 of the non-reduced samples expressed in Huh7 cells, the protein bands representing the E1E2 complex (~97 kDa), E2 (~66 kDa) and E1 (~30 kDa) were clearly visible. However, in the equivalent lanes for BHK cells, the bands representing E1E2 and E2 were barely visible. The E1 protein bands were more evident but were much fainter than those from the Huh7 cells. Aggregates were present in both cell types. In the reduced samples, there was no apparent difference between the intensities of the protein bands in Huh7 and BHK cells. In addition to the R3 (genotype 3) and H77 (genotype 1) constructs that were previously analysed in Figure 4.8, this

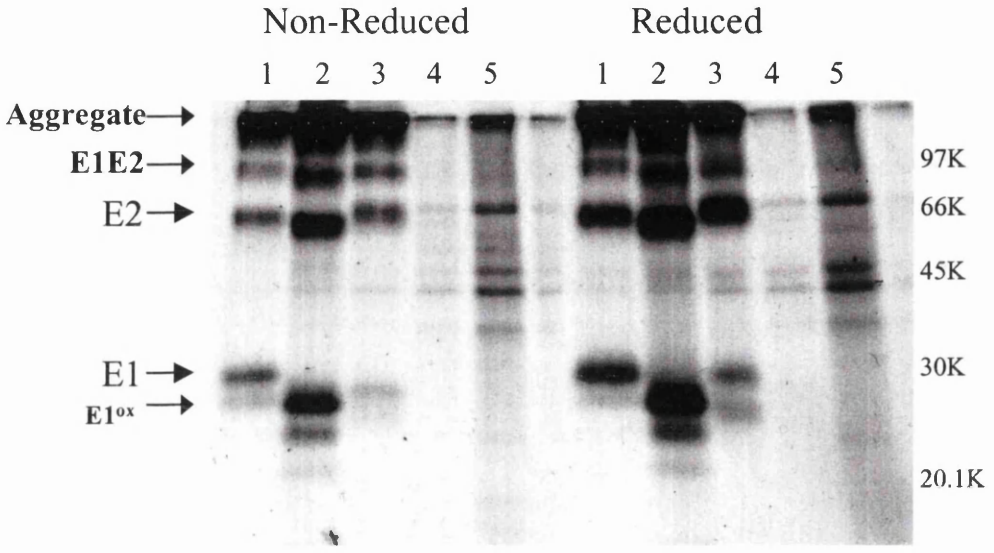
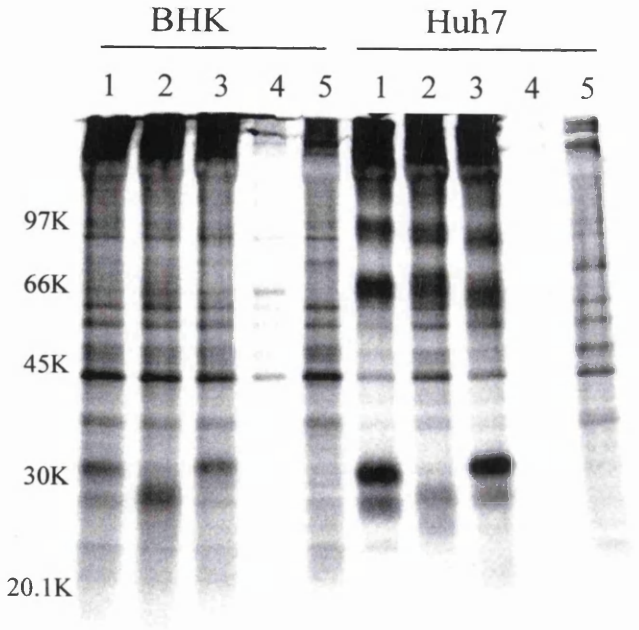
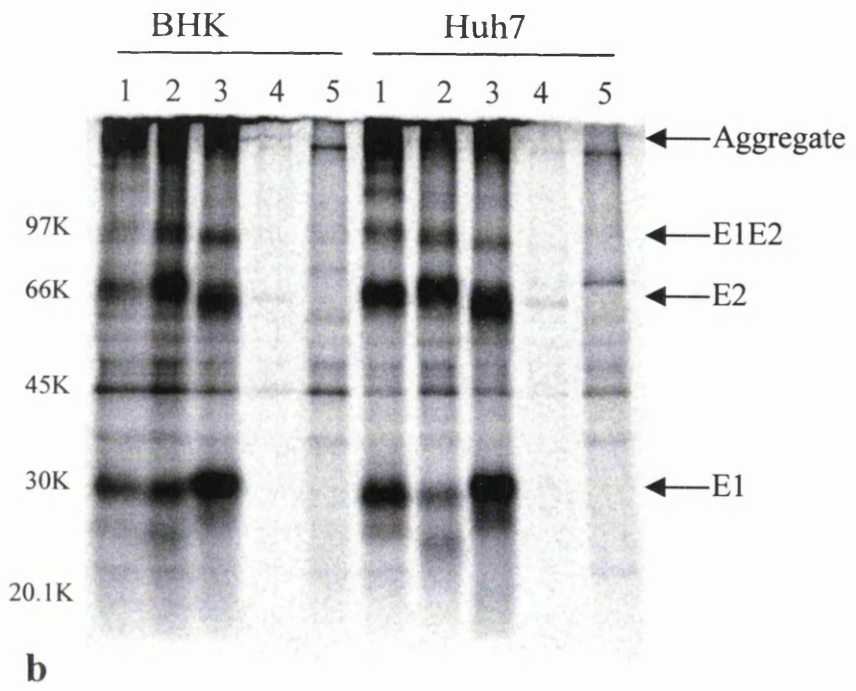


Figure 4.8. Immunoprecipitation of E1 and E2 with anti-E2, ALP98. Huh7 cells were electroporated with *in vitro* transcribed RNA from pSFVR3CE1E2 (lane 1), pSFVF2CE1E2 (lane 2), pSFVH77CE1E2 (lane 3) and pSFV1 (Lane 4). Lane 5 contains extract of mock electroporated cells. Proteins were radiolabelled with ^{35}S-methionine and immunoprecipitated with anti-E2 attached to protein A sepharose beads via rabbit anti-mouse IgG. Bound proteins were removed and half of the sample was treated with DTT (reduced samples) while the other half remained untreated (non-reduced samples). The proteins were separated on a 10% polyacrylamide gel. The positions and sizes of E1 and E2 are indicated.


BHK
Huh7
1 2 3 4 5 1 2 3 4 5
97K
66K
45K
30K
20.1K

a


BHK
Huh7
1 2 3 4 5 1 2 3 4 5
97K
66K
45K
30K
20.1K
Aggregate
E1E2
E2
E1

b

Figure 4.9. Immunoprecipitation of E1 and E2 with anti-E2 in both BHK and Huh7 cells, showing aggregate and native complex formation. Cells were electroporated with *in vitro* transcribed RNA from pSFVR3CE1E2 (lane 1), pSFVH77CE1E2 (lane 2), pSFVGlaCE1E2 (lane 3) and pSFV1 (lane 4). Lane 5 contains extract from mock electroporated cells. Proteins were radiolabelled and immunoprecipitated as described in Fig. 4.8. Non-reduced samples are shown in (a) and reduced samples in (b). The position and sizes of E1 and E2 are indicated.

experiment included the other genotype 1 strain, Gla (lane 3). This strain also shows evidence of both native complex and aggregate complex formation.

4.4.2. Comparison of E1 and E2 glycosylation states between genotypes 1 and 3.

The difference in sizes of E2 between the different constructs was first noted during Western blot analysis (Figure 4.2). However it became more evident in the immunoprecipitation experiments that E2 of the F2 construct migrated faster than that of the R3 construct which in turn migrated faster than E2 of H77 (Figure 4.8). The size of the E1 protein also differed between the constructs with E1 of F2 migrating fastest and E1 of R3 the slowest (Figure 4.8). As the E1 and E2 proteins were known to be heavily modified by *N*-linked glycosylation, the most likely explanation was that the proteins had different glycosylation states. As discussed in chapter 3, the protein sequence predicted 5 glycosylation sites in E1 for R3 and H77 and 4 for F2. E2 had 10 predicted glycosylation sites for R3, 11 for H77 and 9 for F2. To analyse these differences, the proteins were immunoprecipitated as before and were treated with glycosidases (Endo H and PNGase F) before separation by SDS-PAGE (Figure 4.10). These enzymes should remove the carbohydrate portion from the protein and therefore the protein band should represent that of the protein backbone alone. In lanes 1-3 of the glycosidase-treated samples, the unglycosylated form of the E1E2 complex was seen at approximately 60 kDa. Two forms of the unglycosylated E2 protein were seen, a 41 kDa form and a 36 kDa form. The larger protein was probably the uncleaved E2/p7 protein whereas the smaller was the fully cleaved E2 protein. In the PNGase F-treated samples, the unglycosylated form of E1 was approximately 18 kDa in all three constructs. This applied to the Endo H-treated forms of E1 for the R3 and F2 constructs as well but the Endo H-treated form of H77 E1 migrated slightly faster.

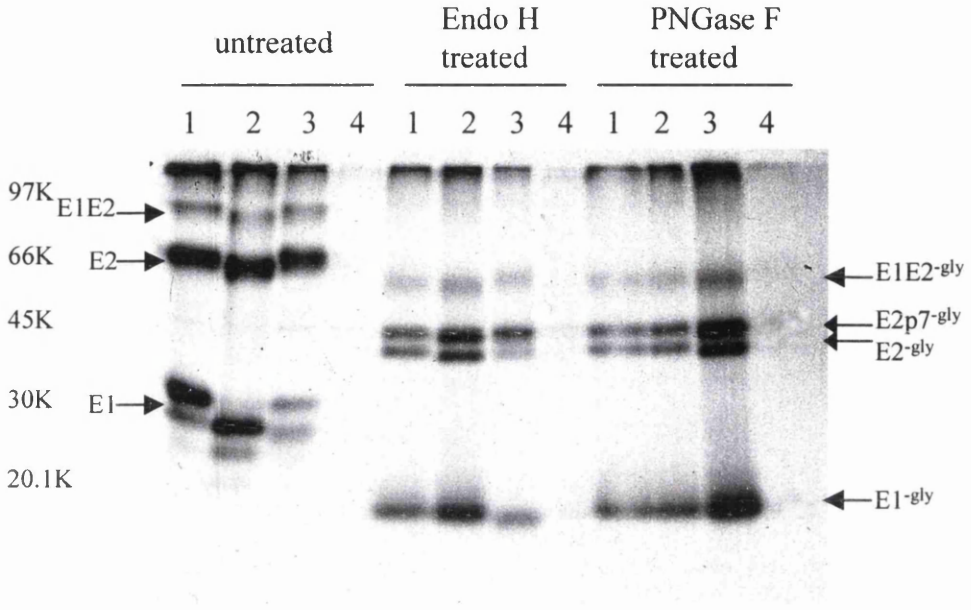


Figure 4.10. Immunoprecipitation of glycosylated and unglycosylated forms of E1 and E2 with anti-E2. Huh7 cells were electroporated with *in vitro* transcribed RNA from pSFVR3CE1E2 (lane 1), pSFVF2CE1E2 (lane 2), pSFVH77CE1E2 (lane 3) and pSFV1 (lane 4). Proteins were radiolabelled and immunoprecipitated as described in Fig. 4.8. Samples were divided into three equal parts and proteins treated with glycosidases (Endo H or PNGase F) or left untreated. Proteins were separated on a 10% polyacrylamide gel. The positions and sizes of the different forms of E1 and E2 are indicated.

4.5. Discussion

Of the 11 anti-E2 antibodies tested which recognised genotype 1 (H77) E2, only 4 (ALP98, AP266, ALP1 and R646) were able to recognise genotype 3 E2 as well. Western blot analysis (Figure 4.2) shows that the E2 protein is approximately 66 kDa for both genotype 1 and 3 isolates, with slight differences in migration detectable. These differences will be discussed in detail later but generally the size of genotype 3 E2 is in agreement with that of genotype 1 which has been reported as 68 kDa (Dubuisson *et al.*, 1994). In ELISA analysis of ALP98 reactivity (Figure 4.4B), there is evidence of more reactivity with the construct expressing both E1 and E2 than that expressing E2 alone but this is only evident with R3 and not H77. Also it is not possible to ensure that the expression levels of these constructs are equivalent, therefore the degree of reactivity is not directly comparable. The epitope on E2 recognised by ALP98 (Figure 4.6) is fully conserved between genotype 3 and genotype 1 and therefore it is not surprising that ALP98 recognises both types 1 and 3 E2. This epitope also contains one of the conserved cysteine residues and a conserved glycosylation site. The epitope recognised by AP266 has been mapped to the same position as that of ALP98 but Western blot analysis does not show an identical pattern of protein recognition (Figure 4.2). This may be simply because more antibody was present for ALP98 (which was used in its concentrated form) and therefore stronger activity is seen here than with AP266. Alternatively, it suggests that the complete epitopes recognised by these antibodies are different. Comparison of the reactivity of these antibodies by ELISA (Figure 4.4B) indicates that AP266 recognises denatured E2 better than native E2, whereas ALP98 recognises both equally well. This increased activity with ALP98 may reflect recognition of additional proteins as demonstrated in Western blot analysis. A second possibility is that the epitope recognised by ALP98 is equally accessible on the denatured and native protein, whereas the epitope recognised by AP266 is more accessible on the denatured form of E2 than the native form. This is speculative. Suffice to say there is no difference in reactivity of these antibodies with genotype 1 and genotype 3 E2.

ALP1 does not recognise E2 of the genotype 3 F2 construct in Western blots, immunofluorescence or ELISA. It does however recognise E2 of the genotype 3 R3 construct. Within the epitope recognised by ALP1 (Figure 4.6) is one of the amino acids that is mutated in R3 (mutation 8 in Figure 3.6). There is a glycine residue at position 661 in F2 and a glutamic acid residue in R3. This is the only difference between these constructs in this region and therefore must account for the difference in reactivity with ALP1. Glutamic acid has a larger side chain than glycine, so the loss of reactivity may be due to this local structural change. There is an additional amino acid difference between genotype 1 and genotype 3 within this epitope but it does not have an effect on reactivity. Genotype 1 has a leucine residue at position 660 and in genotype 3 this is an isoleucine. These residues are similar and there is no significant difference in the size of the side chain, which probably explains why there is no loss of reactivity with ALP1. It is interesting that the only monoclonal antibodies that recognise genotype 3 E2 all recognise epitopes within a 16 amino acid region. There are several stretches of conserved regions between genotypes 1 and 3 but interestingly, there are no antibodies specific to these regions. The only other antibody that showed potential for recognising genotype 3 E2, was the polyclonal R646. However, this could only be demonstrated in Western blot analysis (Figure 4.2). In immunofluorescence (Figure 4.3B) and ELISA (data not shown), the true reactivity was difficult to distinguish from the high background levels, though E2-specific staining was visible in immunofluorescence which is an advantage of this technique over ELISA. The activity of this antibody against type 3 E2 is therefore much less than with type 1 E2 but it is not absent. As with the data from the monoclonal antibodies, this indicates that there are differences between the epitopes on the type 1 and type 3 E2 proteins.

The epitopes have been mapped for two of the antibodies that recognise only genotype 1 E2 (Figure 4.6). Within the epitope recognised by AP33 there is only one amino acid difference between type 1 and type 3, a lysine residue in R3 and an asparagine residue in H77. Lysine has a larger side chain than asparagine. This could mean that the antibody is not able to bind and would explain the loss in reactivity. Apart from this residue, the epitope is conserved between R3 and H77, including a glycosylation site. This lysine

residue is also present in the genotype 3 HCVCENS1 isolate (accession no: X76918) but in the remaining published genotype 3 isolates there is an asparagine at this position as in H77 (Figure 3.4), therefore it is possible that AP33 would recognise these genotype 3 isolates. In contrast, AP320 recognises an epitope that is quite variable. Of the 3 positions where the sequence of R3 differs from that of H77 in this region (Figure 4.6), possibly the most influential is the change from a proline to a serine at position 471. As discussed in chapter 3, proline affects the local protein structure due to its rigid nature and loss of this structure could lead to loss of antibody reactivity. Additionally, the arginine residue at position 466 in R3 is much larger than the alanine in H77, which could also explain the loss of reactivity to type 3 E2. The epitopes of the conformation-dependent E2 antibodies have not been determined and therefore no conclusions can be drawn from their lack of reactivity to genotype 3 E2, other than it is likely to be due to a structural difference. Interestingly, the E2 protein of the genotype 1 Gla isolate is also not recognised by these antibodies, so even between isolates of the same genotype, the structure may differ.

None of the three E1 antibodies tested were able to recognise genotype 3 E1. Judging from the degree of variability of the sequence between type 1 and 3 within the epitopes recognised by AP21.010 and AP119 (Figure 4.6), this is not unexpected. However, the fact that even the polyclonal antibody (R528) does not show any reactivity with type 3 E1, suggests that the exposed regions of the E1 proteins are quite variable.

The cytoplasmic staining of genotype 3 E2-expressing cells seen in immunofluorescence with ALP98 (Figure 4.3A), was comparable to that of genotype 1 E2-expressing cells. This suggested that, as described for type 1 E2 (Dubuisson *et al.*, 1994), type 3 E2 is localised to the endoplasmic reticulum. Evidence of colocalisation between both genotypes 1 and 3 E2 with the ER marker, calnexin, indicates that genotype 3 E2 does localise to the ER (Figure 4.7). Typical reticular-like staining is seen for both E2 and calnexin. Comparison of the co-localisation patterns for genotypes 1 and 3 shows that with genotype 3 there is more evidence of E2 expression on the cell surface than with genotype 1. This could be followed up by analysing the medium of expressing cells for

evidence of E2 secretion or by creating a 3-dimensional image of the expressing cell, using the Z-stack function during confocal microscopy analysis. The co-localisation pattern also confirms that there must be an ER retention signal for type 3 E2. The ER retention signal for genotype 1 has been identified (Cocquerel *et al.*, 1998 & 2000) and the conservation with genotype 3 is shown in Figure 3.11.

Immunoprecipitation with the anti-E2 antibody, ALP98, confirmed that genotype 3 E1 and E2 interact with each other, as determined by the appearance of 66 kDa and 30 kDa protein bands representing E2 and E1 respectively (Figure 4.8). This is true for both R3 and F2 constructs, indicating that none of the 9 mutations in F2 have an effect on the E1/E2 interaction. Yi *et al.* (1997) mapped the region of E2 involved in the E1 interaction to amino acids 415-500, although the actual binding regions are thought to be discontinuous. Within this region there are two amino acid differences between F2 and R3 (Figure 3.6) at positions 496 and 499 but the immunoprecipitation results suggest that these residues are not important for the E1/E2 interaction. Furthermore, it was shown that the amino acid residues "WHY" from 488-490 are critical for E2 interaction with E1 (Yi *et al.*, 1997) and these residues are fully conserved between all genotype 3 isolates and genotype 1 (Figure 3.4), confirming the functional importance of these residues. The region of E1 involved in interacting with E2 has only been restricted to the N-terminal part (aa 192-238) (Yi *et al.*, 1997) in which there is one amino acid difference (at position 236) between F2 and R3, but again this does not appear to affect the E1/E2 interaction. This would suggest that the E1 and E2 proteins interact via their ectodomains. However, this is complicated by the fact that it has been shown that the native E1/E2 interaction is abolished when truncated forms of E1 and E2 are expressed (Michalak *et al.*, 1997) and Patel *et al.* (2001) have also demonstrated that the TMD of E2 is critical for correct folding of E1 and hence native complex formation. Truncated forms of genotype 3 E1 and E2 have not been analysed so it is not known if a similar function can be attributed to the type 3 TMDs.

Two different forms of the E1E2 complex have been widely reported. The heterodimer is stabilised by non-covalent interactions and is considered to be the native form of the HCV

glycoprotein complex (Deleersnyder *et al.*, 1997). In contrast, the aggregate contains intermolecular disulphide bonds and is thought to represent misfolded proteins (Dubuisson *et al.*, 1994). Both non-reduced and DTT-reduced proteins were analysed to determine the ability of the genotype 3 glycoproteins to form these different complexes (Figure 4.8). The presence of monomeric E1 and E2 in the non-reduced samples indicates that these proteins formed a non-covalently-bonded complex, as complexes containing covalent interactions remained intact and did not migrate through the gel. The 97 kDa heterodimer of E1 and E2 was also visible. Therefore, as described for genotype 1, the genotype 3 glycoproteins formed both aggregate and native complex. In addition, formation of native complex has been reported to be accompanied by an oxidised form of E1 which migrated faster (Dubuisson & Rice, 1996) and this form was visible for both genotype 3 constructs. A similar protein pattern was seen for the reduced samples except that the amounts of E2 and E1 monomer were greater than in the non-reduced samples. This was because as well as the proteins from the native complex, there were also proteins present which originated from the aggregated complex which was reduced by the addition of DTT. However, this explanation is over-simplified as Patel *et al.* (1999a) demonstrated that the aggregate also contained non-covalent interactions, as proteins expressed in the presence of DTT continued to precipitate as monomers under non-reducing conditions. This interpretation was based on the observation that the majority of complexes synthesised by the genotype 1 Gla strain were aggregated and under non-reducing conditions, very few monomeric forms of E1 and E2 were seen. These proteins were expressed in BHK cells using the SFV system and when I repeated this, the same observation was made (Figure 4.9a, lane 3, BHK). The same applied to genotype 1 H77 (Figure 4.9a, lane 2, BHK) and genotype 3 R3 (lane 1). In other words, evidence of native complex was indicated by the presence of faint bands representing E1 monomers but the neither the E2 monomer nor E1E2 complex was visible on the non-reducing gel. However, when the proteins were expressed in Huh7 cells, there was clear indication of both E1 and E2 monomers as well as the native E1E2 complex with all three constructs (Figure 4.9a, Huh7). No significant differences were observed between the proteins expressed in BHK and Huh7 cells when analysed under reducing conditions (Figure 4.9b). The reason for the different appearances of the "native" E1 and E2 monomers when

expressed in BHK and Huh7 cells and analysed by non-reducing SDS-PAGE is not clear. This difference was not been reported by a study using other expression systems. Dubuisson *et al.* (1994) examined complex formation of HCV glycoproteins expressed from recombinant vaccinia viruses in a human hepatoma cell line (HepG2), porcine kidney cells (PK-15) and BHK cells and reported similar results for all three cell types. They also examined formation of glycoprotein complexes when E1 and E2 were expressed from recombinant Sindbis viruses, which is related to SFV, in BHK cells only and no differences were observed between this system and the vaccinia virus expression system. Whatever the reasons, the differences observed in this study are a side-issue to the main objective, which was to determine if the genotypes 1 and 3 glycoproteins behaved similarly in terms of complex formation. In conclusion, genotypes 1 and 3 glycoproteins are able to form both aggregates and native complexes.

It is evident from Figures 4.8 and 4.9, that both the E1 and E2 proteins of the genotype 3 and genotype 1 isolates have different mobilities. It is known that genotype 1 E2 is 6 amino acids shorter than genotype 3 E2 (section 3.3.1) but this does not agree with the observed differences in migration and the most likely explanation is differing glycosylation states. Analysis of the E2 protein sequences has shown that R3 has 10 predicted glycosylation sites, H77 has 11 (Figure 3.4) and F2 has 9 due to the mutation at position 558 which abolishes a glycosylation site (Figure 3.6). This is in agreement with the observed mobilities of the untreated E2 proteins (Figure 4.10), with F2 E2 migrating fastest and H77 E2 migrating slowest. After glycosidase treatment, the unmodified E2 proteins all co-migrate; indicating that differing glycosylation states was responsible for the different mobilities. The E2-p7 product is indistinguishable from the E2 product in the untreated samples, however, once the *N*-linked sugars are removed, the uncleaved E2p7 is visible as a slower migrating protein above the fully cleaved E2 product (Figure 4.10). This larger product was first described by Grakoui *et al.* (1993a) but was only identified as E2p7 by Lin *et al.* (1994) and Mizushima *et al.* (1994). This work confirms that as with genotype 1, the cleavage at the E2/p7 site is incomplete in genotype 3. The E1 protein has 5 predicted glycosylation sites in R3 and H77 and 4 in F2 (due to mutation at position 236). However, in genotype 1a it has been shown that the site at

position 325 is not glycosylated due to the presence of a proline immediately after the tripeptide recognition sequence (Meunier *et al.*, 1999). This proline residue is present in genotype 3 as well, so this site may also be precluded from glycosylation in genotype 3 E1. The relative mobilities of the untreated E1 proteins of R3, F2 and H77 (Figure 4.10) suggests that R3 has an additional glycosylation site or conversely, that H77 has one less. Possibly, the 5th glycosylation site is used in R3 but not in H77. Although, if this site is used in R3, it would also be used in F2 which would mean that both H77 E1and F2 E1 have 4 glycosylation sites and should co-migrate but they do not. Similar differences in mobility of the H77 and Gla E1 proteins were reported by Patel *et al.* (1999a) and are also seen in Figure 4.9 but are not explained by differing glycosylation states. There is also a discrepancy in the mobility of H77 E1 in the reduced and non-reduced samples (Figures 4.8 and 4.9), with the non-reduced E1 migrating faster than the reduced E1. This suggests that the E1 proteins in the native complex and aggregate may be processed differently or that E1 in the native complex has a more compact conformation due to reshuffling of disulphide bonds as proposed by Dubuisson and Rice (1996). This is not observed with either the genotype 1 Gla isolate or the genotype 3 isolate. Treatment with peptide *N*-glycosidase F (PNGase F) resulted in co-migration of E1 from R3, F2 and H77 but when treated with endoglycosidase H (Endo H), H77 E1 migrated marginally faster than E1 of R3 and F2 (Figure 4.10). The reason for treating the proteins with both Endo H and PNGase F is that it provides additional information on the subcellular localisation of these proteins. Endo H removes the high-mannose *N*-linked sugars but not the complex forms. In contrast, PNGase F removes both high-mannose and complex sugars. Resistance to Endo H digestion but not PNGase F digestion indicates that the protein has left the ER and is in the medial or trans Golgi, where complex sugars are added. There is no evidence of resistance to Endo H digestion in any of the constructs, which confirms that E1 and E2 are retained in the ER. These data would therefore not support any proposed surface expression of R3 E2 as suggested by the calnexin co-localisation data. There is a difference in mobility between Endo H- and PNGase F-treated H77 E1. Dubuisson *et al.* (1994) reported a similar observation with the genotype 1b HCV-BK strain but not with the genotype 1a HCV-H strain which is most similar to H77. There is no obvious explanation for this other than perhaps one of the glycans on H77 and BK E1

is resistant to PNGase F digestion. These results do suggest that the unmodified E1 protein of H77 is smaller than that of R3 or F2. This is not predicted from the sequence so may possibly be attributed to differences in disulphide bonding which creates a more compact structure, although all 8 cysteine residues are conserved in both genotype 1 and 3.

Chapter 5

Interaction of HCV E2 with PKR and PERK.

5.1. Investigation of HCV E2 interaction with Protein Kinase R (PKR).

5.1.1. Introduction

The most established functional difference between HCV genotypes 1 and 3 to date is that patients infected with genotype 3 have a much higher response rate to interferon treatment, either alone or in combination with ribavirin (McHutchison *et al.*, 1998, Zeuzem *et al.*, 2000). The mechanism behind this remains unclear but two HCV proteins, NS5A and E2, have been implicated in the evasion of the interferon (IFN) response through inhibition of protein kinase R (PKR) activity (Gale & Katze, 1998). PKR is a critical component of the IFN-induced antiviral response and several viruses have evolved mechanisms to block its activity. The high failure rate of IFN therapy in HCV infected individuals suggests that HCV has also developed such a mechanism. Most studies have focused on the role of the NS5A protein since Enomoto *et al.* (1995 &1996) identified an IFN-sensitivity determining region (ISDR) within NS5A of genotype 1b. If there were four or more amino acid "mutations" in the ISDR, it was reported that all patients responded to IFN therapy. Several groups have attempted to confirm the predictive value of the sequence within the ISDR but this has proved controversial as, at first, most studies

from outside Japan were unable to show such a correlation (Zeuzem *et al.*, 1997, Khorsi *et al.*, 1997, Hofgartner *et al.*, 1997). More recently, some European studies have reported a correlation between "mutant" type 1b isolates and sustained response (Saiz *et al.*, 1998, Sarrazin *et al.*, 1999) and suggested the discrepancy is due to the relatively low prevalence of genotype 1b isolates in Western countries compared to Japan. Genotype 1a has a high prevalence in the United States and Europe and shows the same response rates as genotype 1b, however no correlation between response and mutations in the ISDR of type 1a isolates has been demonstrated (Sarrazin *et al.*, 2000a, Zeuzem *et al.*, 1997). Likewise, no correlation has been reported for genotype 3a (Saiz *et al.*, 1998, McKechnie *et al.*, 2000, Sarrazin *et al.*, 2000b).

The biological significance of the ISDR appeared to be confirmed when Gale *et al.* (1997) demonstrated that NS5A interacted with and inhibited PKR in a yeast expression system and that deletion of the ISDR abolished this ability. In addition to the ISDR, an additional 26 amino acids downstream of the ISDR were required for PKR-binding and introduction of multiple mutations in the ISDR (as found in IFN-responsive isolates) abrogated the NS5A/PKR interaction and inhibition (Gale *et al.*, 1998). The same laboratory was able to show that NS5A from genotype 3a nonresponder was also able to inhibit PKR in the yeast assay (personal communication, C. Blakely). This was the same patient (and serum sample) as was used in this present study but is the only known information on a non-genotype 1 NS5A effect on PKR. Despite this evidence, there is still doubt over the role of the ISDR in combating the antiviral effects of IFN. Several studies have shown that expression of NS5A confers partial resistance to IFN, allowing growth of IFN-sensitive viruses such as encephalomyocarditis virus (EMCV) (Paterson *et al.*, 1999, Polyak *et al.*, 1999, Song *et al.*, 1999). However, the expression of a mutant, in which the entire ISDR was deleted, retained this ability (Polyak *et al.*, 1999) and Paterson *et al.* (1999) reported that NS5A clones derived from a responder patient were more inhibitory than those from a nonresponder, indicating that the effects of NS5A *in vitro* do not correlate with the clinical response. Following on from these studies, Francois *et al.* (2000) examined the inhibitory effect of the entire HCV genotype 1a polyprotein and reported that it conferred partial resistance to IFN as measured by the growth of EMCV.

However, they also reported that there was no evidence that PKR activity was affected by the expression of the HCV proteins.

The development of a subgenomic replicon system for HCV allowed investigation of the effect of IFN on HCV replication. The levels of both replicon RNA and HCV proteins are reported to decline after IFN treatment, in a dose-dependent manner (Frese *et al.*, 2001). Interestingly, Blight *et al.* (2000) have identified an adaptive mutant which lacks the entire ISDR, but both this mutant and the wild type replicons were highly sensitive to IFN treatment. Similarly, Gou *et al.* (2001a) reported sensitivity to IFN in replicons derived both from IFN-resistant and –sensitive cDNA clones from patients. However, HCV replicons persisted even at high concentration of IFN and long-term treatment failed to select IFN-resistant variants. Subsequently, IFN has been reported to "cure" cells from replicons and this process is dependent on cellular factors, but not PKR or RNaseL (Guo *et al.*, 2001b). To date, there have been no reports on the effect of the HCV replicon on PKR activity and none of the replicons, even if including the HCV structural genes, form virus particles (Pietschmann *et al.*, 2001), so it is still not possible to observe the effects of IFN in terms of the complete virus life cycle.

The involvement of E2 in mediating IFN resistance was first suggested by Taylor *et al.* (1999). Within E2, the authors identified a 12 amino acid sequence which shares homology with the autophosphorylation sites of PKR and the phosphorylation site of its substrate, eIF2α (Figure 5.1). This region was named the PKR- eIF2α phosphorylation homology domain (PePHD). The PePHD regions of the IFN-sensitive genotypes (2 and 3) were noted to have less homology to PKR than the more resistant genotypes (1a and 1b) and this was proposed as an explanation for the differing responses to IFN therapy. The authors suggested that NS5A may be involved in the development of resistance over the course of IFN treatment as demonstrated by the appearance of different NS5A quasispecies, whereas E2 is responsible for the intrinsic resistance or sensitivity of an HCV genotype. Genotype 1 E2 was shown to interact with PKR in an *in vitro* assay and the PePHD region was required for this interaction. PKR activity was also inhibited by genotype 1 E2 in yeast and mammalian cells (in a PePHD-dependent manner) and this

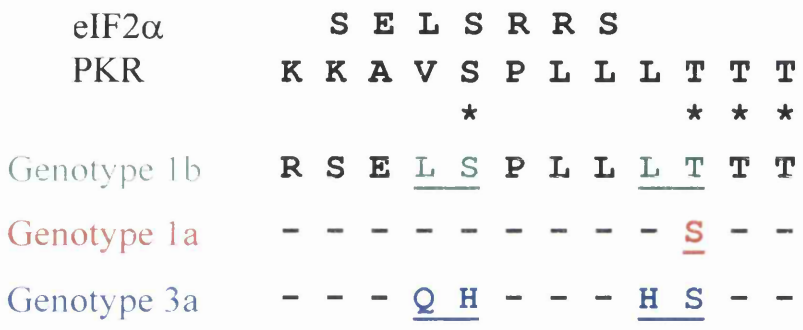


Figure 5.1 Amino acid sequence of the PePHD region for genotype 1b (HCV-J), genotype 1a (HCV-H) and genotype 3a (NZL1). The sequence surrounding the PKR phosphorylation site of eIF2α and the autophosphorylation sites of PKR are shown in comparison. The phosphorylation sites are indicated with "*". In the PePHD alignment, "-" represents an identical amino acid. Taken from Taylor *et al.* (1999).

was abolished when the PePHD was mutated to resemble that of genotypes 2 and 3. In a follow-up study, it was shown that this inhibition is not due to E2 acting as a pseudosubstrate for PKR autophosphorylation sites (Taylor *et al.*, 2001).

5.1.2. Sequence analysis of the PePHD region of genotype 3 E2 in patients with differing responses to interferon therapy.

There are extensive clinical data to show that HCV genotype 3 is more sensitive to IFNα than genotype 1. However, a proportion of genotype 3-infected individuals do not have a sustained response to treatment. In the paper describing identification of the PePHD region in E2, Taylor *et al.* (1999) indicated that the blocking of IFNα through PKR occurred only when the sequence matched that of genotype 1, and that this was abolished when the sequence was mutated to resemble that of genotype 2 or 3. Therefore, those genotype 3-infected individuals who are found to be resistant to IFNα treatment may harbour a PePHD variant similar to that of genotype 1.

A visiting worker, Alexandra Cochrane, performed the following study in collaboration with Alexander Orr and myself. RNA was extracted from serum, reverse-transcribed and the PePHD region amplified by PCR from 14 genotype 3-infected patients. The patients showed a range of treatment outcomes which included sustained responders (SR), non-responders (NR), relapsers (REL) and breakthrough (BT) patients (see Figure 5.2 for definitions). The breakthrough patients were of particular interest, as due to the quasispecies nature of HCV, it is possible that a PePHD variant exists as a minority at first and then undergoes positive selection during the course of interferon treatment. In view of this, the PePHD regions of the BT and REL patients were sequenced from serum samples collected both before treatment and during (patients 5 and 6) or after (patients 4, 8 and 9) treatment (for simplicity these are referred to as post treatment samples). Following RNA extraction, the viral RNA was reverse transcribed and the cDNA was amplified through two rounds of PCR. After purification, the PCR products were sequenced and the derived amino acid sequences (position 617-707) and the response to

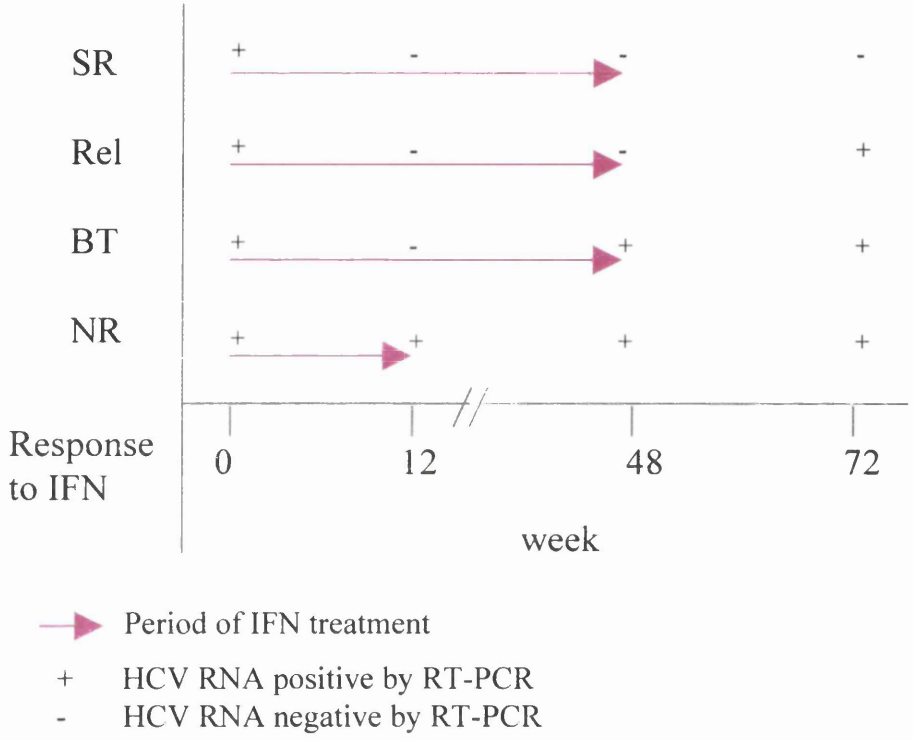


Figure 5.2 Diagrammatic representation of the response definitions for IFN treatment of chronic hepatitis C. SR = sustained responder, Rel = relapse, BT = breakthrough and NR = non-responder.

therapy for each patient is shown in Figure 5.3. The PePHD region in genotype 3 is from aa 665-677 (numbering according to NZL1) and is highlighted in yellow. Only one patient (a non-responder) had a PePHD variant which differed from the consensus genotype 3 sequence. This variant had a single amino acid change towards the genotype 1b sequence. Incidentally, the patient is the same one whose serum was used for amplification of the structural genes of genotype 3 HCV (see chapter 3). The 30 amino acids downstream of the PePHD region showed no variation. Some variation was observed in the upstream sequence but this showed no correspondence to either treatment response or the time of sample collection.

In conclusion, in this study of 14 genotype 3 infected patients there were no consistent differences between responders and non-responders and there was no evidence for positive selection of PePHD variants in the breakthrough or relapser patients.

5.1.3. Localisation of E2 and PKR within the cell

5.1.3.1. Cloning of H77 E2 into pcDNA 3.1

For the purposes of studying the effect of E2 on endogenous cellular proteins, it was necessary to express the E2 protein in a system other than the SFV system in which there is shutdown of host cell protein synthesis. The pcDNA 3.1 vector (Figure 5.4) was chosen as a suitable vector where the expression of E2 is under the control of a cytomegalovirus (CMV) promoter. A construct encoding H77 E1 and E2 in the pcDNA vector (pcH77E1E2) was available from Dr A.Patel (MRC Virology Unit, Glasgow). The strategy for subcloning E2 from this E1/E2 construct is illustrated in Figure 5.5. The plasmid was first cut with the restriction endonuclease *Nco*I which cleaved at the end of E1 but before the signal sequence for E2. This released a fragment of approximately 1.5Kb including the E2 gene and part of the vector. The 5' overhang was filled in with Klenow DNA polymerase to create a blunt end and the fragment was then cleaved with the restriction endonuclease *Xba*I which cuts in the vector portion. The resultant

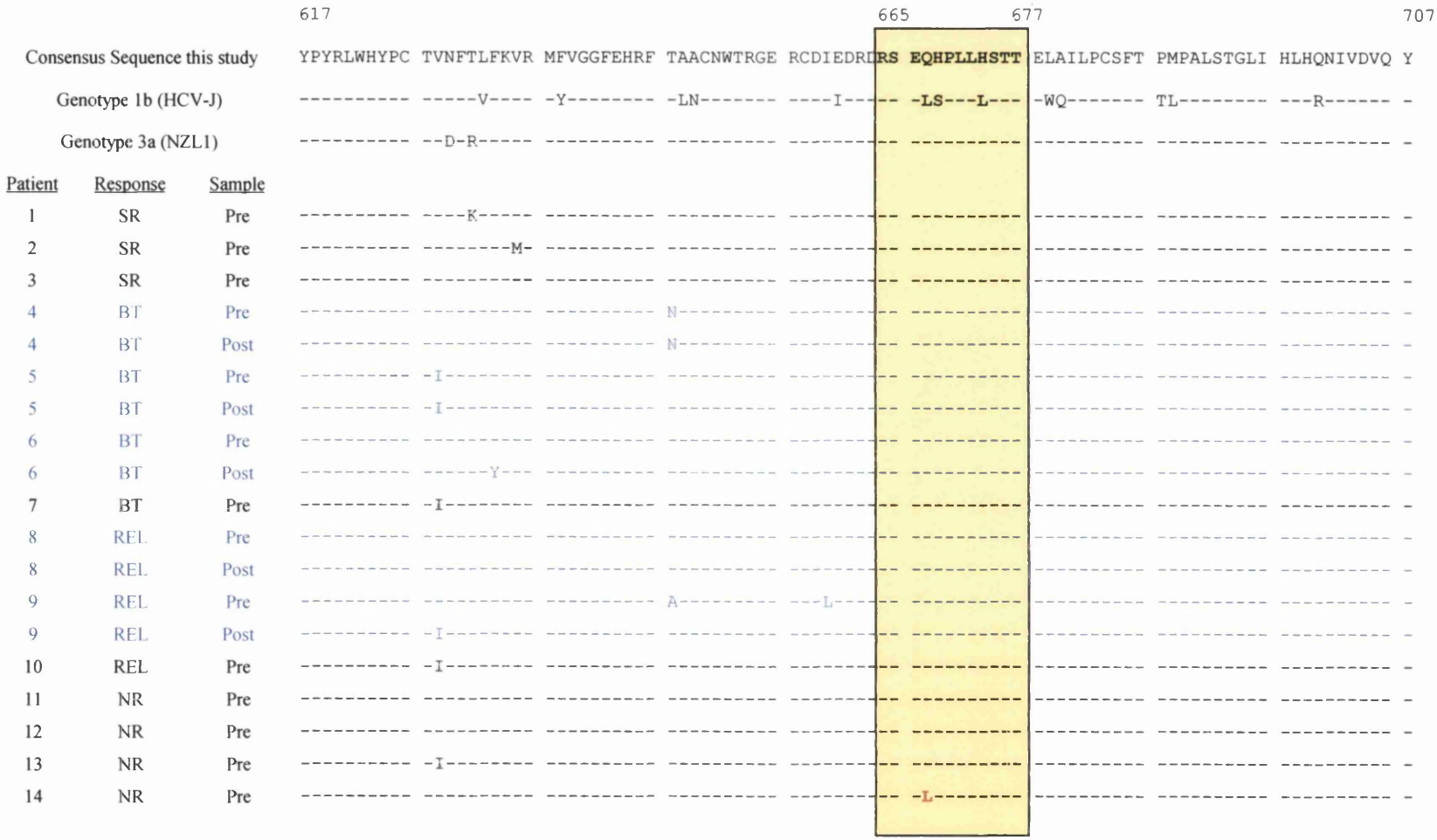

617
665
677
707
Consensus Sequence this study YPYRLWHYPC TVNFTLFKVR MFVGGFEHRF TAACNWTRGE RCDIEDRDRS EQHPLLHSTT ELAILPCSFT PMPALSTGLI HLHQNIVDVQ Y
Genotype 1b (HCV-J) ---------- -----V---- -Y-------- -LN------- ----I----- -LS---L--- -WQ------- TL-------- ---R------ -
Genotype 3a (NZL1) ---------- --D-R----- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Patient Response Sample
1 SR Pre ---------- ----K----- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
2 SR Pre ---------- --------M- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
3 SR Pre ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
4 BT Pre ---------- ---------- ---------- N--------- ---------- ---------- ---------- ---------- ---------- -
4 BT Post ---------- ---------- ---------- N--------- ---------- ---------- ---------- ---------- ---------- -
5 BT Pre ---------- -I-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
5 BT Post ---------- -I-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
6 BT Pre ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
6 BT Post ---------- ------Y--- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
7 BT Pre ---------- -I-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
8 REL Pre ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
8 REL Post ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
9 REL Pre ---------- ---------- ---------- A--------- ---L------ ---------- ---------- ---------- ---------- -
9 REL Post ---------- -I-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
10 REL Pre ---------- -I-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
11 NR Pre ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
12 NR Pre ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
13 NR Pre ---------- -I-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
14 NR Pre ---------- ---------- ---------- ---------- ---------- -L-------- ---------- ---------- ---------- -

Figure 5.3. Alignment of derived amino acid sequences of part of the E2 protein for 14 HCV genotype 3 infected patients. For five patients pre-treatment and post-treatment –failure sequences are given. The PePHD (amino acids 665 to 667), denoted in bold type, is compared with the sequences published for genotypes 1b (Kato *et al.*, 1990) and 3a (Sakamoto *et al.*, 1994). Amino acid numbering as for genotype 3a. Dashes (-) indicate amino acids identical to consensus sequence. SR = sustained responder; BT = breakthrough; REL = relapser; NR = non-responder. 'Pre' denotes pre-treatment sample, 'post' denotes post or during treatment sample.

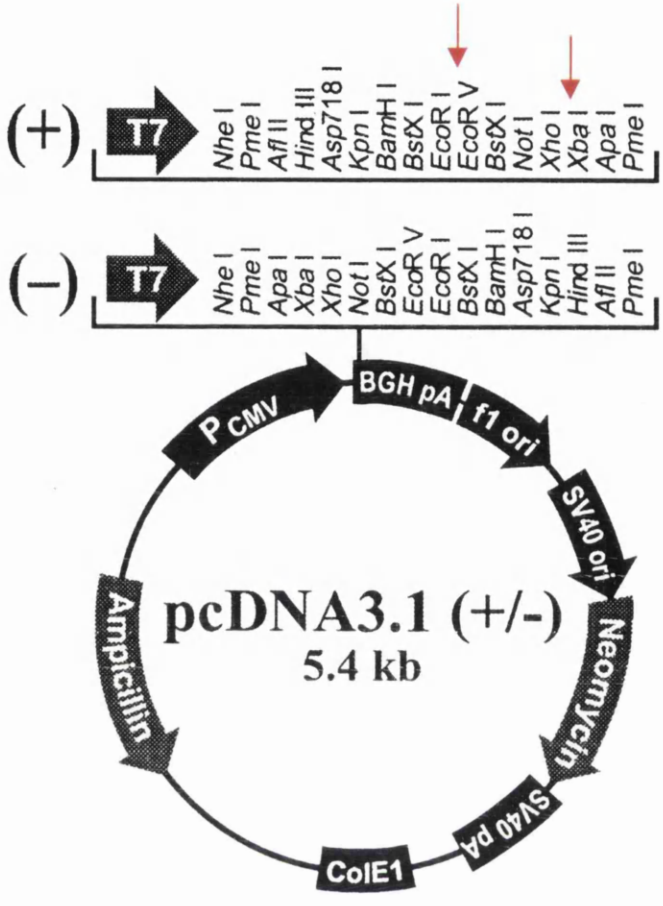


Figure 5.4. Schematic of the pcDNA3.1(+) plasmid. The *Xba*I and *EcoR*V restriction sites in the multiple cloning site, which were used for the insertion of the E2 gene, are indicated with red arrows. The ampicillin resistance gene and the CMV promoter are shown. This figure was taken from the Invitrogen catalogue.

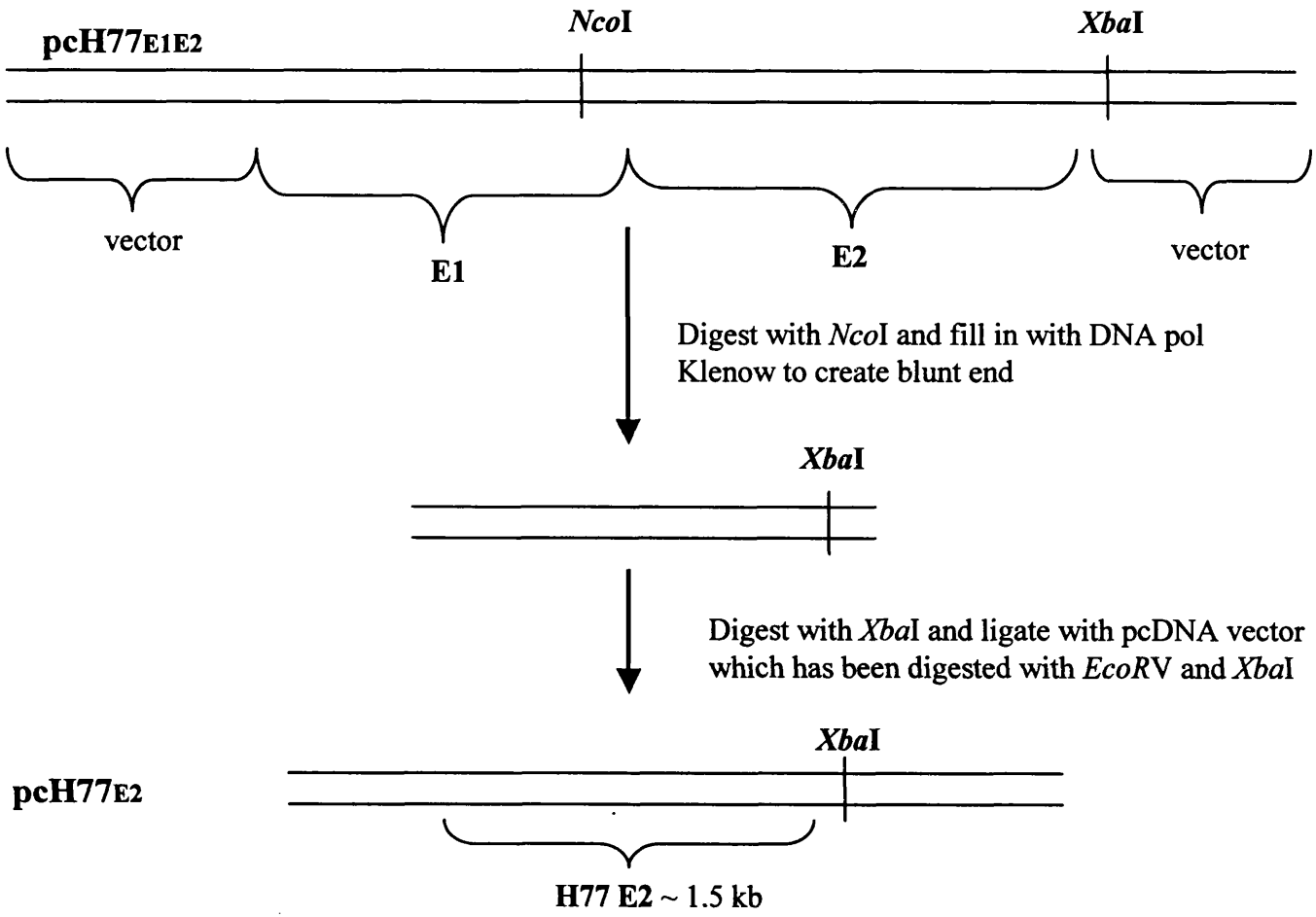


Figure 5.5. Illustration of the strategy for subcloning the E2 gene of the type 1 H77 strain into pcDNA 3.1 to form pcH77E2.

fragment was inserted into the pcDNA vector which had been cleaved with the *EcoR*V and *Xba*I restriction endonucleases. This construct was named pcH77E2.

5.1.3.2. Treatment of cells with interferon α

This work was done by Suzanne Barnard (Division of Virology, University of Glasgow), using the same batch of interferon which was used in the PKR study. A cis-reporter plasmid, containing a consensus interferon-stimulated response element (ISRE) upstream of the luciferase gene, was used to assess the ability of this particular batch of interferon-α to activate transcription of the reporter gene. The plasmid was transfected into 2fTGH cells, which have been shown to have an intact IFN-signalling pathway (Pellegrini *et al.*, 1989) and following treatment with IFNα, the cells were assayed for luciferase activity. The results, with increasing concentrations of IFNα, are shown in Figure 5.6. Luciferase activity increased with increasing concentrations of IFNα up until 2500 U/ml where it started decreasing, probably due to saturation of the system. This indicated that this batch of IFNα was active.

5.1.3.3. Effect of interferon treatment on PKR expression

It has been widely reported that exposure of cells to IFN type I results in an increase in levels of PKR as a result of transcriptional activation of the gene (Meurs *et al.*, 1990). Two monoclonal antibodies directed against PKR were obtained for the purpose of studying the putative interaction of PKR and E2. One of these PKR antibodies (71/10) has been well characterised (Laurent *et al.*, 1985) and was a gift from A.G. Hovanessian (Institut Pasteur) but was only available in limited quantities, while the other (B-10) is available commercially from Santa Cruz but has not been used extensively in published PKR studies. No positive control for PKR expression was readily available and therefore the use of a previously characterised PKR antibody would assist in positive identification of the PKR protein. The ability of these antibodies to recognise PKR in both untreated and IFNα treated Huh7 cells was analysed by indirect immunofluorescence and confocal

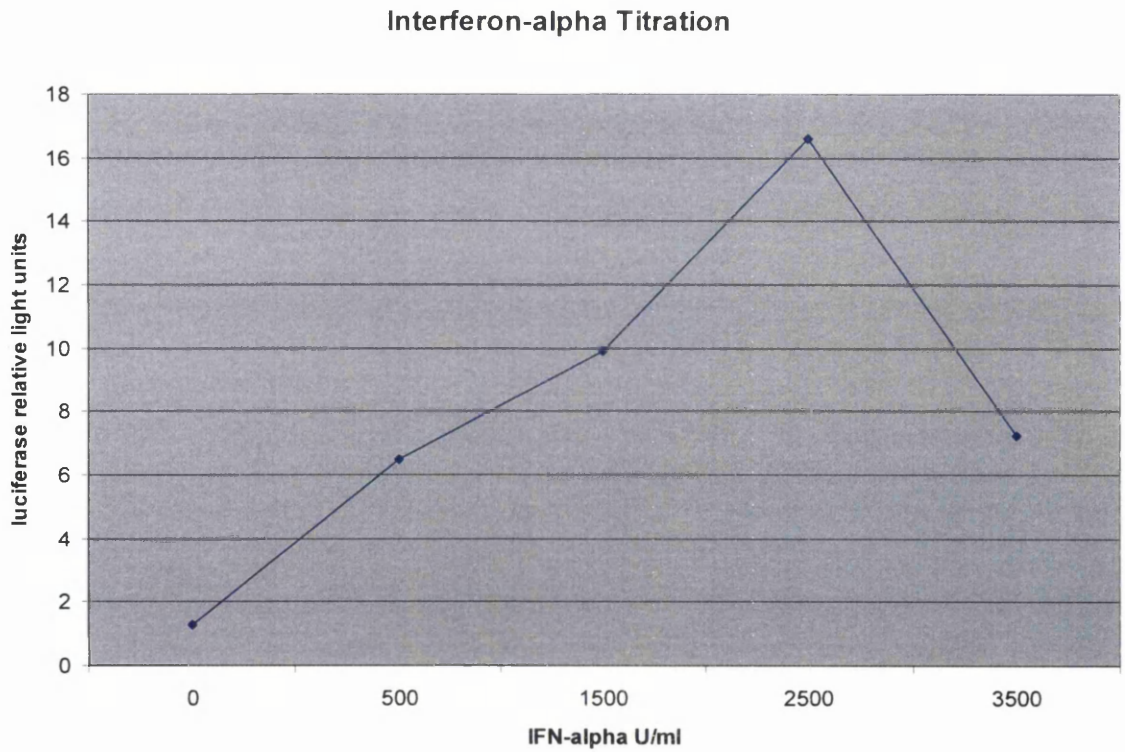


Figure 5.6. Stimulation of the interferon sensitivity response element (ISRE) by increasing concentrations of IFN-α as measured by luciferase activity. 2fTGH cells were transfected with a cis-reporter plasmid (containing a ISRE) and treated with increasing concentrations of IFN. The cells were then assayed for luciferase activity. Data kindly supplied by S. Barnard.

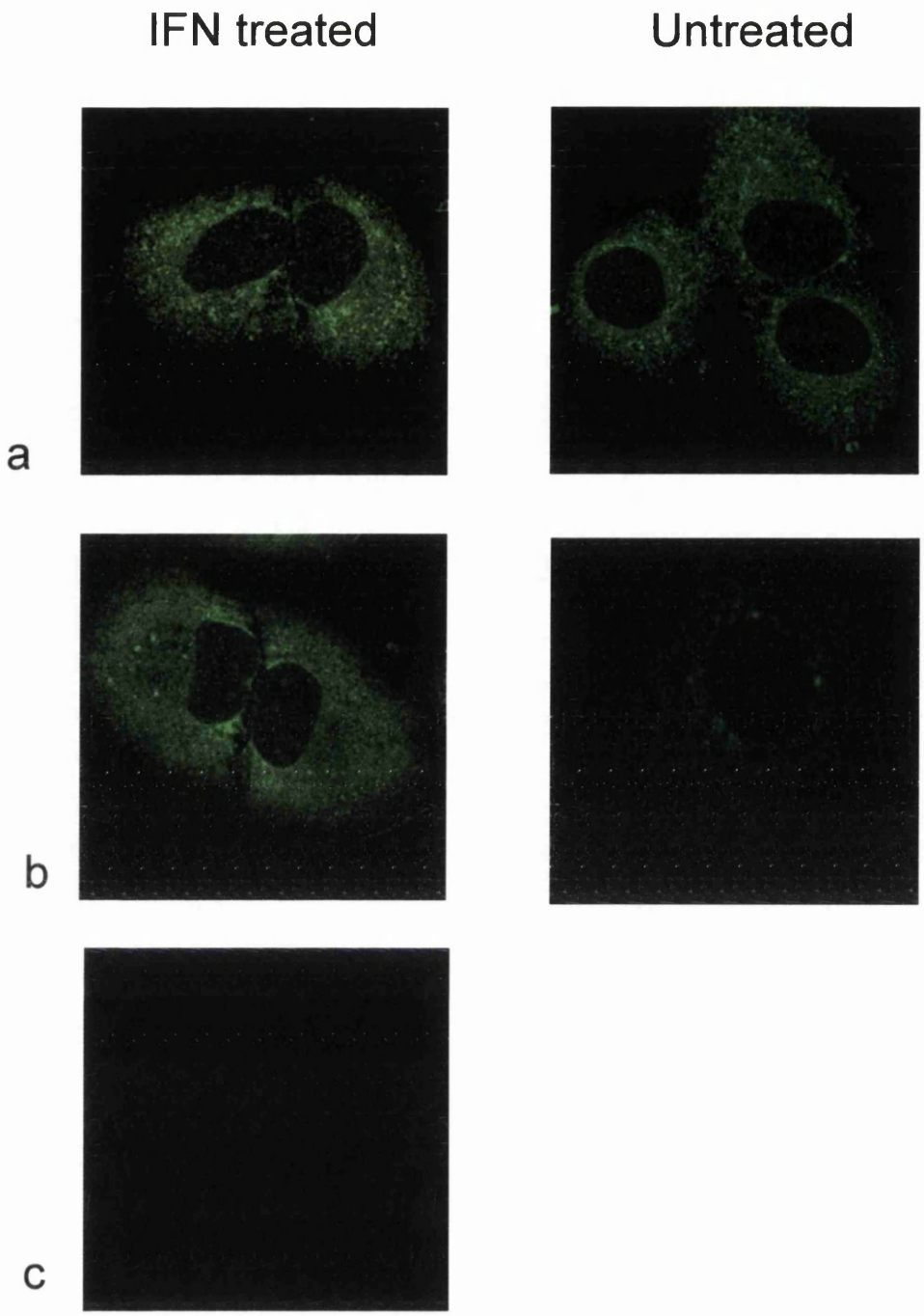


Figure 5.7. Localisation and levels of endogenous PKR expression in interferon-treated and untreated Huh7 cells. Cells were treated with 1000 U/ml of interferon-alpha for 18 hrs before being fixed in methanol and processed for indirect immunofluorescence. PKR was detected by two antibodies a) mouse monoclonal 71/10 (Dr A. Hovenessian) and b) mouse monoclonal B-10 (Santa Cruz). Panel (c) represents cells probed with goat anti-mouse FITC alone.

microscopy (Figure 5.7). There was no clear difference in levels of PKR expression between IFN-treated and untreated cells with the 71/10 antibody (a). However, with the B-10 antibody, there was a noticeable increase in PKR expression in cells treated with IFNα (b).

5.1.3.4. Localisation of E2 and PKR

The reported interaction of genotype 1 E2 with PKR (Taylor *et al.*, 1999) can only be significant if these two proteins are found in the same compartment within a cell. Colocalisation studies using immunofluorescence and confocal microscopy were performed to determine if this is the case. It is important to look at this in an appropriate cell type. Huh7 cells were transfected with pcH77E1E2 and treated with IFNα before probing for both E2 and endogenous PKR with antibodies R646 and 71/10 respectively (Figure 5.8.). The cellular localisation of PKR and E2 can be seen in panels (a) and (b) respectively and panel (c) shows the merged images of (a) and (b), where areas of colocalisation are depicted in yellow. PKR showed cytoplasmic staining as well as faint nuclear staining. E2 displayed typical endoplasmic reticulum (ER) staining as shown previously in Figure 4.7. The merged image showed small regions of possible colocalisation around the nuclear membrane but there was no evidence for definitive colocalisation between genotype 1 E2 and PKR. Unfortunately it was not possible to repeat this with genotype 3 E2 due to the lack of a polyclonal antibody to recognise type 3 E2, or conversely, a polyclonal PKR antibody.

5.1.4. Interaction of PKR and E2

5.1.4.1. Lack of evidence for interaction of endogenous PKR and E2.

I further investigated the possibility of a direct interaction of E2 with PKR using immunoprecipitation in which an antibody was used to isolate a specific protein and any other protein which physically interacted with it. Proteins were then identified by

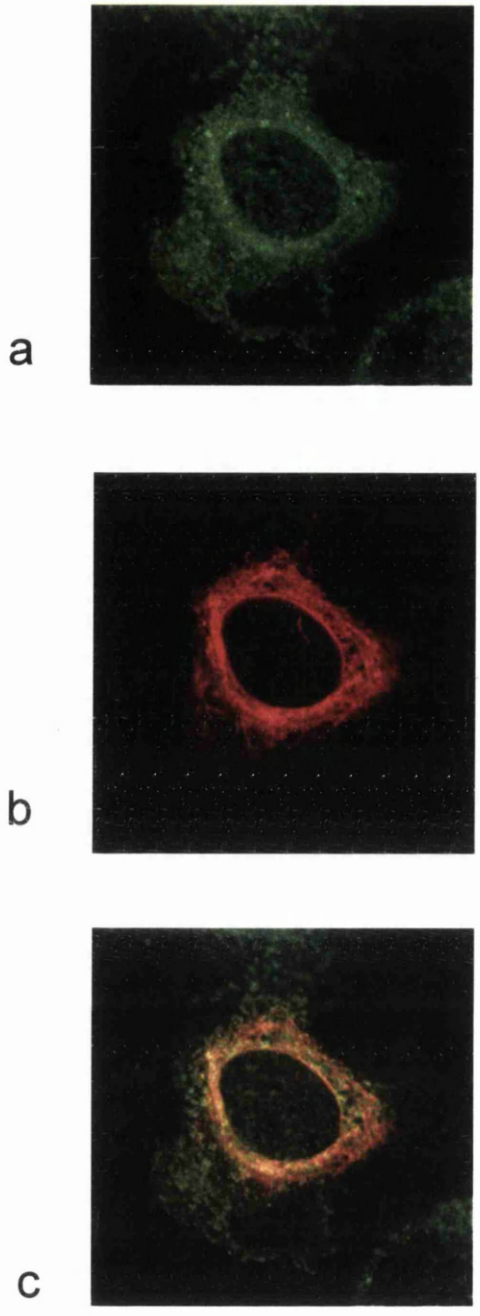


Figure 5.8. Dual staining for type 1 E2 and endogenous PKR in interferon-treated Huh7 cells. Cells were transfected with pcH77$_{E1E2}$ and treated with interferon as described in Figure 5.7. They were then fixed and processed for indirect immunofluorescence. a) endogenous PKR detected with mouse monoclonal 71/10. b) type 1 E2 detected with rabbit polyclonal R646. c) merged image of a and b. Bound antibodies were detected using goat anti-mouse FITC and goat anti-rabbit TRITC.

Western blot using an appropriate antibody. Huh7 cells were transfected with pcH77E2 and were either treated with IFNα or left untreated. Anti-E2 antibodies (ALP98 or AP33) and anti-PKR antibody (B-10), were used to immunoprecipitate either E2 or PKR from the cell extracts. Precipitated proteins were separated by SDS-PAGE and then analysed by Western blot with anti-E2 or anti-PKR antibodies (Figure 5.9). As the epitope recognised by ALP98 lies close to the putative sight of interaction (the PePHD region); AP33 was also used for immunoprecipitation of E2. The E2-expressing cell extracts where anti-E2 was used for both immunoprecipitation and Western blot analysis are visible in lanes 1 and 4 of (a) and (c). The E2 protein is indicated. In contrast, E2 was not visible when anti-PKR was used for immunoprecipitation and anti-E2 for Western blot analysis (lanes 7 and 10 of (a) and (c)). Likewise, endogenous PKR was visible in lanes 7-12 of (b) and (d) when anti-PKR was used for both immunoprecipitation and Western blot analysis. In contrast, it was not visible when anti-E2 was used for immunoprecipitation (lanes 1 and 4 of (b) and (d)). This showed that the anti-E2 antibodies could precipitate E2 and that anti-PKR could precipitate PKR but not vice versa, which indicated that genotype 1 E2 did not interact with PKR. No differences were observed between the IFNα treated cells (lanes 4-6 and 10-12) and the untreated cells (lanes 1-3 and 7-9), except for a slight increase in PKR expression in the IFN-treated cells (compare lanes 7,8 and 9 to 10, 11 and 12 in (b) and (c)). The only difference observed between the anti-E2 antibodies, ALP98 and AP33, was an extra band at approximately 22K (indicated with star) in lanes 1 and 4 of (c) when using ALP98. This band was also visible in Western blots using ALP98 (see Figure 4.2). It is too small to represent the non-glycosylated form of E2 but must represent an E2-associated protein, as it was visible only in E2-expressing cells.

5.1.4.2. Effect of interferon on E2 expression in Huh7 cells

A noticeable increase in PKR levels in IFN-treated cells was evident in the previous experiment. This was not as evident with E2 due to the high expression levels. To assess whether IFN has an effect on E2 expression levels, Huh7 cells were transfected with pcH77E2 and the cells either treated with IFNα or left untreated. Two fold dilutions of the E2-expressing cell extracts were analysed by Western blot using anti-E2, ALP98

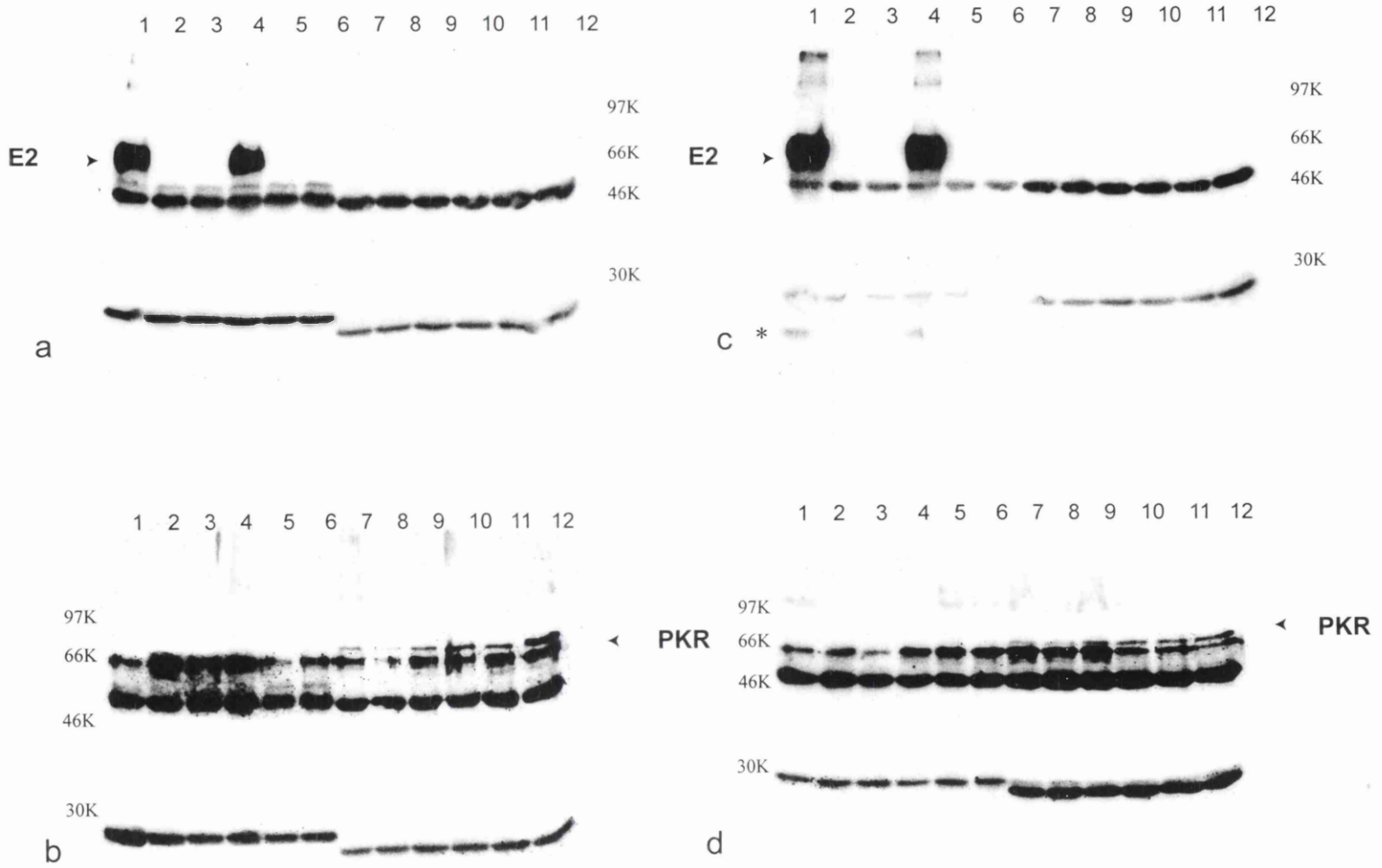

1 2 3 4 5 6 7 8 9 10 11 12
97K
66K
46K
30K
E2
a
1 2 3 4 5 6 7 8 9 10 11 12
97K
66K
46K
30K
E2
c
*
1 2 3 4 5 6 7 8 9 10 11 12
97K
66K
46K
30K
PKR
b
1 2 3 4 5 6 7 8 9 10 11 12
97K
66K
46K
30K
PKR
d

Figure 5.9. Immunoprecipitation of genotype 1 E2 with PKR in interferon-treated and untreated cells. Huh7 cells were transfected with pcH77E2 and treated with interferon as described before. Cell extracts were subjected to immunoprecipitation with either anti-E2 (lanes 1-6) or anti-PKR (B-10) (lanes 7-12). Anti-E2, AP33 was used in (a) and (b) and ALP98 was used in (c) and (d). Precipitated proteins were separated on a 10% polyacrylamide gel and transferred to membrane. Western blot analysis was performed using anti-E2 (AP33 for (a) and ALP98 for (c)) or anti-PKR (B-10) (b and d). Lanes 1-3 and 7-9 show pcH77E2-, pcDNA- and mock-transfected cells which are untreated. Lanes 4-6 and 10-12 show pcH77E2-, pcDNA- and mock-transfected cells which have been treated with interferon. The positions of E2 and PKR are indicated.

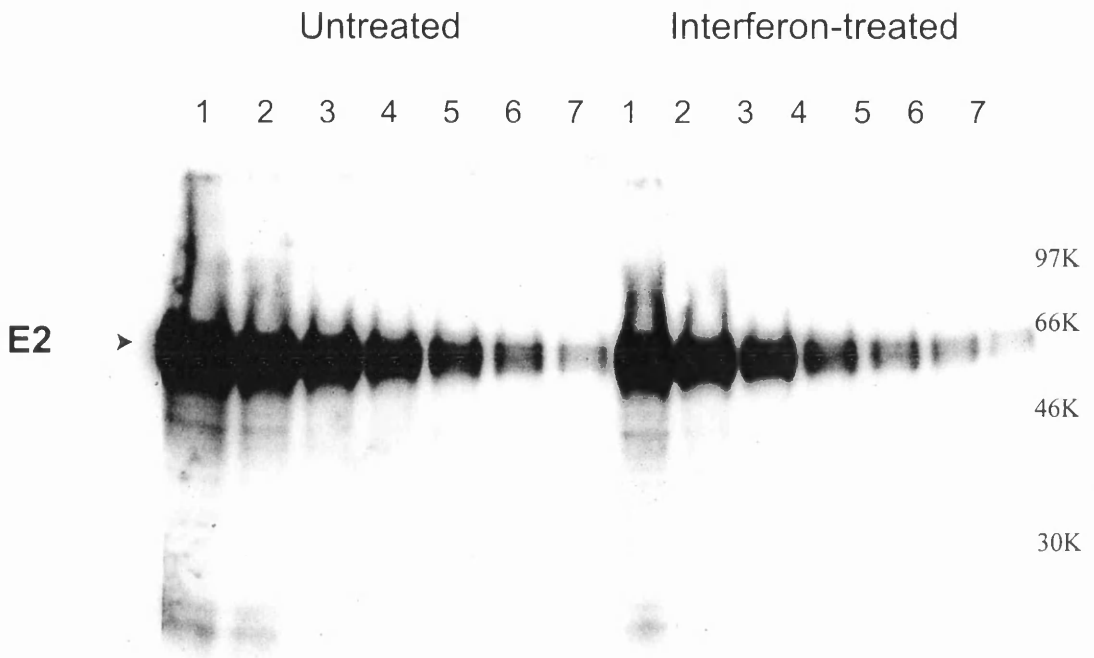


Figure 5.10. Effect of interferon treatment on E2 expression levels. Huh7 cells were transfected with pcH77E2 and treated with 1000 U/ml of interferon for 18 hrs. Two-fold dilutions of the cell extracts were analysed by Western blot using mouse monoclonal anti-E2, ALP98. Lanes 1-7 show two-fold dilutions of cell extract from neat to 1/64.

(Figure 5.10). Comparison of E2 levels in untreated cells and treated cells revealed that there was an approximate two fold decrease in E2 expression in IFN-treated cells (compare untreated lane 7 to treated lane 6). Such a small difference is of doubtful biological significance.

5.1.5. Discussion

Sequencing of the PePHD region from 14 genotype 3 infected patients revealed only one patient where the sequence differed from that of the consensus for type 3. This variant contained mutation Q668L, which corresponds to the genotype 1 (or IFN resistant) sequence at that position. This patient is a nonresponder, which agrees with the hypothesis of Taylor *et al.* (1999), that variation towards the genotype 1 sequence, leads to IFN resistance. However, the PePHD regions were identical in all other patients who displayed a range of responses to treatment. Similar results have been reported by other studies looking at genotype 3 PePHD sequences (Sarrazin *et al.*, 2000b, Puig-Basagoiti *et al.*, 2001 and Abid *et al.*, 2000). Sarrazin *et al.* (2000b) looked at the PePHD regions and the flanking sequences (amino acids 638-704) for 33 type 3 patients and concluded that the number of mutations in this region of E2 did not correlate with treatment response. The authors also did phylogenetic and conformational analyses on these sequences but again, were unable to demonstrate any link between the mutations and sensitivity or resistance to IFN treatment. Of the 33 patients, 10 were found to have PePHD variants and the authors noted that hydrophobic mutations within the hydrophilic region of the PePHD, were exclusively found in 4 patients with sustained response. Three of these mutations were Q668L, which is identical to the mutation seen in a nonresponder in this study and in a responder in the study by Abid *et al.* (2000). The term "responder" in this last study was not defined, nor was the time of sample collection in terms of IFN treatment, but these conflicting reports indicate that this mutation does not correlate with treatment response. As in this study, Sarrazin *et al.* (2000b) also looked at the PePHD and flanking regions of relapser patients and nonresponders, both before and after treatment and also reported no changes over the course of treatment. However, ours is the first study to report on the "breakthrough" patients and the lack of PePHD change in

patients 5 and 6 is particularly important as the "post-treatment" samples were taken while IFN therapy was ongoing and showed that there was no selection for IFN-resistant variants within this region. In conclusion, there is no evidence that IFN exerts selection pressure in the PePHD region of genotype 3 and pre-treatment sequencing of this region is unlikely to provide useful information to clinicians as an indicator of treatment response. Similar results have been reported for genotype 1 (Berg *et al.*, 2000, Polyak *et al.*, 2000, Chayama *et al.*, 2000, Sarrazin *et al.*, 2000b, Puig-Basagoiti *et al.*, 2001 and Abid *et al.*, 2000).

Although the lack of selection of a PePHD variant during IFN treatment suggests that it is not the primary resistance mechanism, it does not preclude a role for E2 in IFN resistance, possibly through an interaction with PKR. In the study by Taylor *et al.* (1999), the major evidence for a PKR interaction with genotype 1 E2 came from an *in vitro* assay using bacterially expressed, His-tagged PKR. To prove that this reflects the situation *in vivo*, it was important to investigate the interaction with endogenous PKR in mammalian cells. Interferon-α (which was used as part of a treatment regime for HCV-infected patients) was used to induce PKR expression. It was first shown that this IFN could activate transcription of a reporter gene from an ISRE in cells that are known to have an intact IFN signalling pathway (Figure 5.6). It has previously been shown that Huh7 cells (which were used in this study) also have an intact IFN signalling system (Melen *et al.*, 2000), although Keskinen *et al.* (1999) have reported that Huh7 cells have a poor ability to produce and respond to type I IFN during viral infection (with influenza A, Sendai and vesicular stomatitis viruses) *in vitro*. I investigated the ability of IFN to induce the expression of PKR in our Huh7 cells and with the commercial PKR antibody, there was an increase in PKR expression in IFN-treated cells, but another antibody failed to make such a clear distinction (Figure 5.7). This latter PKR antibody has been well characterised and similar immunofluorescence results with interferon treated cells have been reported (Jimenez-Garcia *et al.*, 1993), indicating that this technique is possibly not optimal for detecting changes in expression levels.

Having established that endogenous PKR could be detected by indirect immunofluorescence in IFN-treated Huh7 cells, this technique was used to detect areas of colocalisation for PKR and E2 (Figure 5.8). As described by Taylor *et al.* (1999), small areas of colocalisation were visible in the perinuclear region but this was not convincing evidence of colocalisation overall and is most likely due to overlapping staining of E2 in the ER and PKR in the cytoplasm. This is in agreement with the report from Francois *et al.* (2000) where no colocalisation was shown between PKR and the entire HCV polyprotein. Further investigation of the putative interaction of endogenous PKR and E2 by immunoprecipitation, showed no evidence for a direct interaction of genotype 1 E2 with PKR (Figure 5.9). The anti-E2 antibody, ALP98, recognises an epitope close to the PePHD region and therefore could possibly block any interaction between E2 and PKR. For this reason, another anti-E2 (AP33) was also used but this antibody also failed to co-precipitate PKR. Therefore, it has not been possible to repeat the *in vitro* observations of Taylor *et al.* (1999) with endogenous PKR. Interestingly, Podevin *et al.* (2001) have recently shown that NS5A and endogenous PKR do not colocalise or co-immunoprecipitate and that neither wild-type nor mutant NS5A has an effect on PKR activity. These results contradict those of Gale *et al.* (1997&1999), however in the latter studies, PKR inhibition by NS5A was examined in yeast *Saccharomyses cerevisiae* and NIH 3T3 cells, whereas Podevin *et al.* used a liver cell line (Huh7).

According to the hypothesis, genotype 3 E2 does not interact with PKR due to the PePHD region having less homology to PKR and eIF2α. Unfortunately time constraints meant it was not possible to confirm or deny this by repeating this study with genotype 3 E2. In addition, all available antibodies that recognise genotype 3 E2, recognise epitopes adjacent to the PePHD region which may interfere with any interaction. However, the fact that no interaction was observed between genotype 1 E2 and PKR, as proposed, suggests that an interaction with genotype 3 E2 is unlikely. A recent report at the 8th International HCV conference (Pavio *et al.*, 2001) described an unglycosylated form of E2 (termed p38) which is found in the cytoplasm. The authors stated that it was this form which was found to interact with PKR and not the fully glycosylated form which is ER-bound (Taylor, 2001). Firstly, it should be noted that the E2 construct used in their study did not

contain the full E2 signal sequence (aa 363-377 of polyprotein) and therefore it is not surprising to find a cytoplasmic form of E2 which presumably has not been properly processed in the ER. Secondly, this finding contradicts that of the original report (Taylor *et al.*, 1999), where it was shown that the E2 signal sequence was necessary for efficient inhibition of PKR, implying that it is the ER-bound and fully glycosylated form of E2 which interacts with PKR. No evidence was seen in this present study for an unglycosylated form of E2.

5.2. Investigation of E2 interaction with PKR-like endoplasmic reticulum resident kinase (PERK).

5.2.1. Introduction

One of the anomalies with the proposed E2 interaction with PKR is that E2 (and specifically the PePHD region) is located on the lumenal side of the ER membrane, while PKR is most likely associated with ribosomes on the cytoplasmic side of the membrane (Besse *et al.*, 1998, Zhu *et al.*, 1997, Jiminez-Garcia *et al.*, 1993). Therefore, the identification of an ER-resident eIF2α kinase (PERK) which showed homology to PKR initiated interest in investigation of a possible interaction with E2. Although the domain showing homology with PKR is located on the cytoplasmic side of the ER membrane (and does not include a PePHD region), the lumenal domain of PERK is responsible for sensing stress in the ER due to accumulation of misfolded protein. ER stress activates the kinase domain of PERK, which phosphorylates serine residue 51 of eIF2α and inhibits protein translation (Harding *et al*, 1999). With the knowledge that expression of the HCV glycoproteins leads to accumulation of misfolded protein aggregates (at least *in vitro*), it seemed plausible that an interaction between E2 and the lumenal domain of PERK (either directly or indirectly) may have an effect on the downstream signalling of PERK.

5.2.2. Localisation of E2 and PERK

As with the PKR study, to prove any interaction between E2 and PERK, it was important to show that the two proteins are found in the same region of the cell by immunofluorescence and confocal microscopy. Huh7 cells were transfected with pcH77E2 and probed for both E2 and PERK (Figure 5.11). E2 staining is shown in panel (a) and PERK staining in panel (b). Areas of colocalisation of the two proteins are shown in yellow and seen in panel (c). Extensive colocalisation in the ER network could be seen in panel (c). However, the reason for this clear pattern of colocalisation was realised when the appropriate controls were included (Figure 5.12) and it was found that the incorrect secondary antibody had been used. E2-expressing cells were stained with ALP98 and both FITC- and Cy3-conjugated secondary antibodies (panels a and b). Panel (a) showed E2 staining however, the red staining in panel (b) was due to reaction of anti-goat Cy3 conjugate with goat anti-mouse FITC conjugate. This gave the appearance of colocalisation in panel (c) that was not authentic. Therefore the PERK staining in Figure 5.11b was not true and using an appropriate secondary antibody (generated in rabbits), this staining was repeated in Figure 5.13. Once again, Huh7 cells were transfected with pcH77E2 and probed for E2 (panels a-c) and PERK (panels d-f). The merged images are seen in panels (g), (h) and (i). Only cells in panels (c), (f) and (i) were probed for both E2 and PERK and therefore the yellow areas seen in panel (i) represented regions of colocalisation of these two proteins. Although there were distinct areas of colocalisation, the staining of PERK alone, in panel (e), did not resemble ER localisation as previously reported (Harding *et al.*, 1999).

5.2.3. Colocalisation of PERK and calnexin

To confirm that PERK is localised to the ER, Huh7 cells were probed for PERK and calnexin, an ER marker (Figure 5.14). Cells in panels (a), (d) and (g) were probed with anti-calnexin and those in panels (b), (e) and (h) were probed with anti-PERK. Panels (g), (h) and (i) show the merged images. Only cells in panels (c), (f) and (i) were probed for

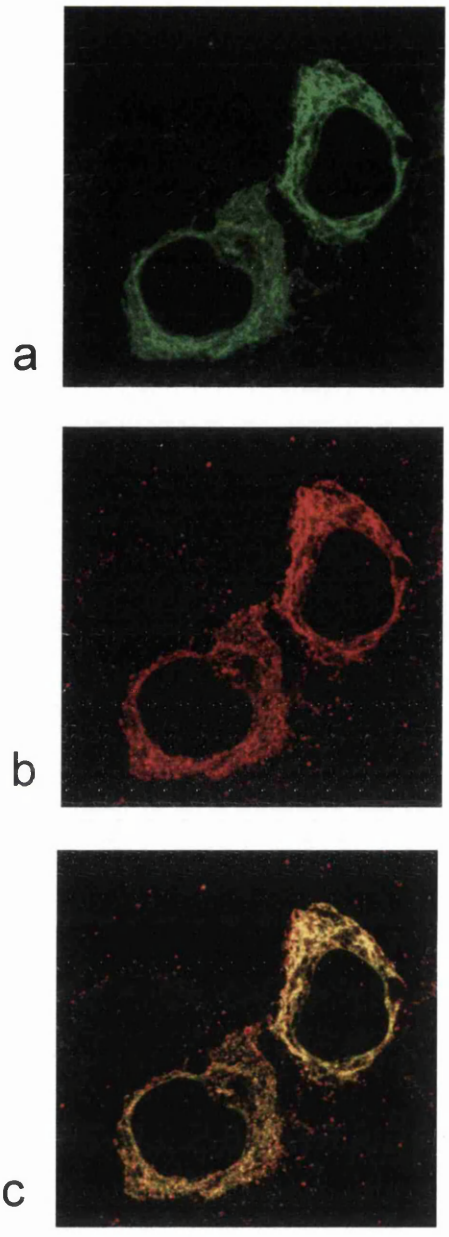


Figure 5.11. Apparent co-localisation of E2 and PERK which was subsequently found to be due to cross-reaction of the secondary antibodies. Huh7 cells were transfected with pcH77$_{E2}$ and fixed in methanol before being processed for indirect immunofluorescence using mouse anti-E2 (ALP98) (a) and goat anti-PERK (N-18) (b). Panel (c) represents the merged image of a and b. Bound antibodies were detected using goat anti-mouse FITC and rabbit anti-goat Cy3.

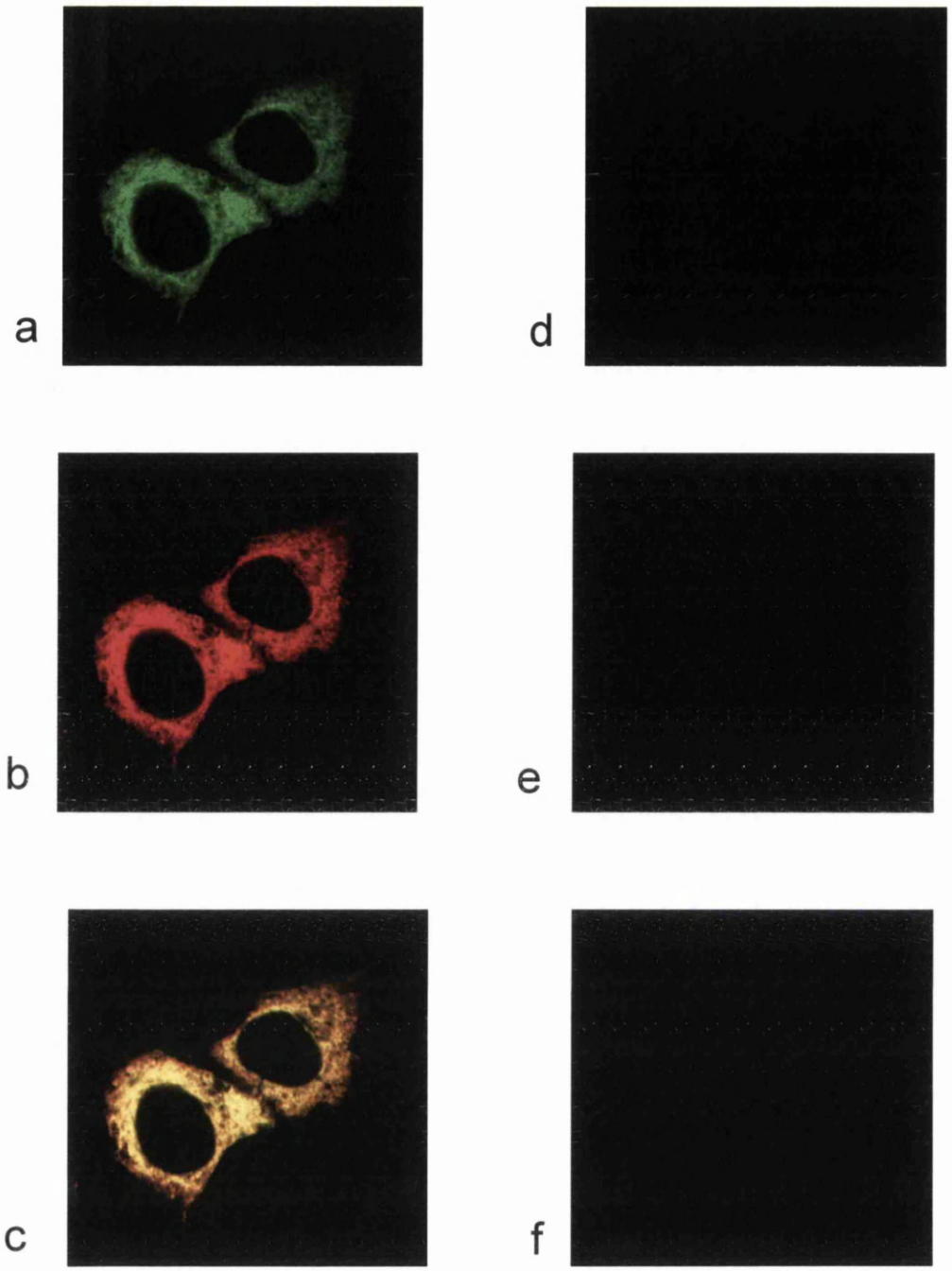


Figure 5.12. Controls showing evidence of cross-reaction between the secondary antibodies used in Fig 5.11. Huh7 cells were transfected with pcH77E2 and processed for indirect immunofluorescence. In panels (a) and (b), cells were probed with mouse monoclonal anti-E2 (ALP98) and bound antibody was detected with goat anti-mouse FITC conjugate and rabbit anti-goat Cy3 conjugate. In panels (d) and (e), the cells were stained with the two secondary antibodies alone. Panels (c) and (f) represent the merged images of panels (a) and (b), and (d) and (e) respectively.

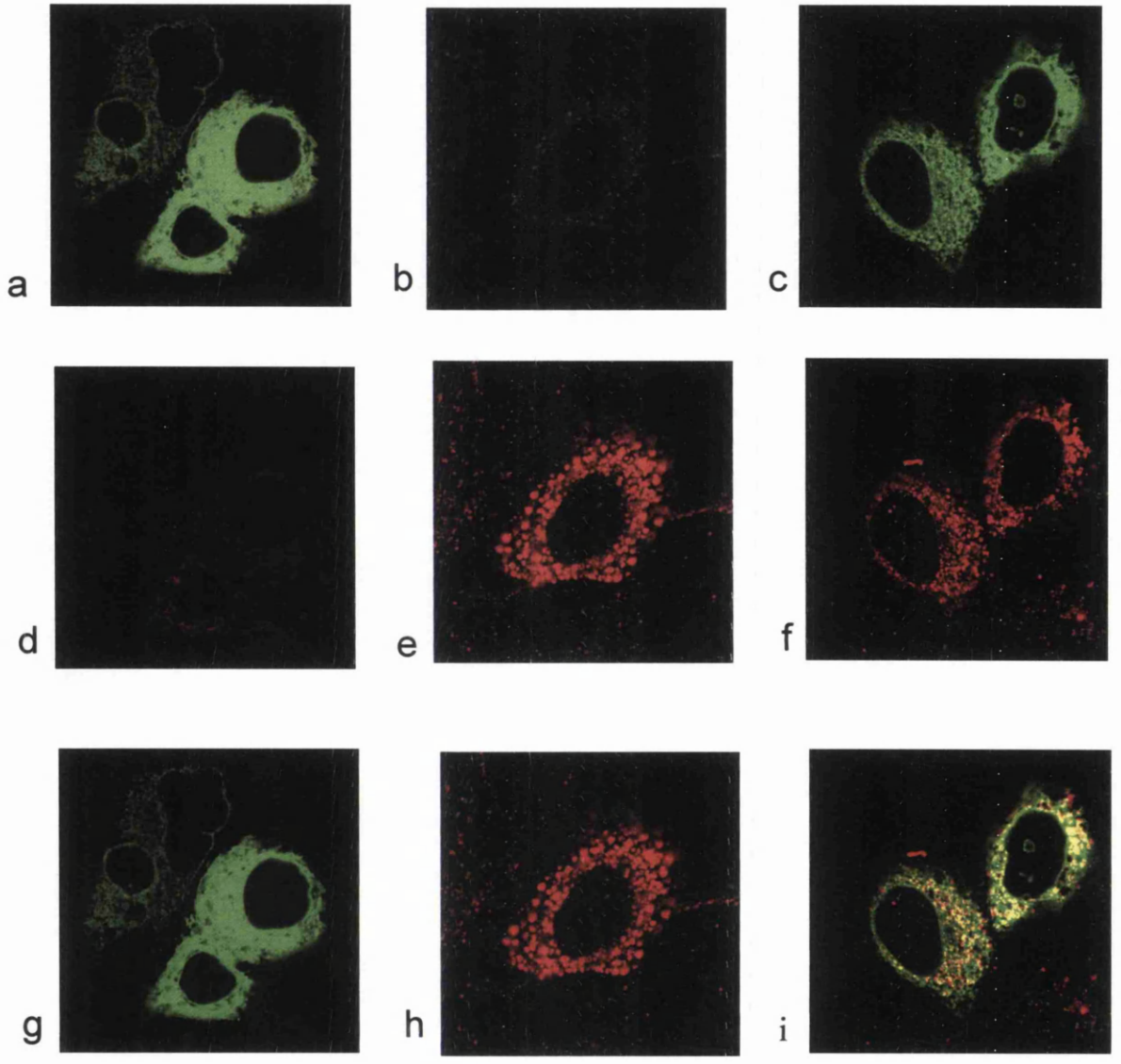

a
b
c
d
e
f
g
h
i

Figure 5.13. Localisation of E2 and PERK using the correct secondary antibodies. Huh7 cells were transfected with pcH77E2 and processed for indirect immunofluorescence. Cells in panels (a), (d), (g) and cells in panels (b), (e), (h) were probed singly with anti-E2 (ALP98) and anti-PERK (N-18), respectively. Cells in panels (c), (f) and (i) were doubly probed with anti-E2 (ALP98) and anti-PERK (N-18). In all panels, bound antibodies were detected using rabbit anti-mouse IgG FITC and rabbit anti-goat IgG Cy3. The merged images are shown in panels (g), (h) and (i).

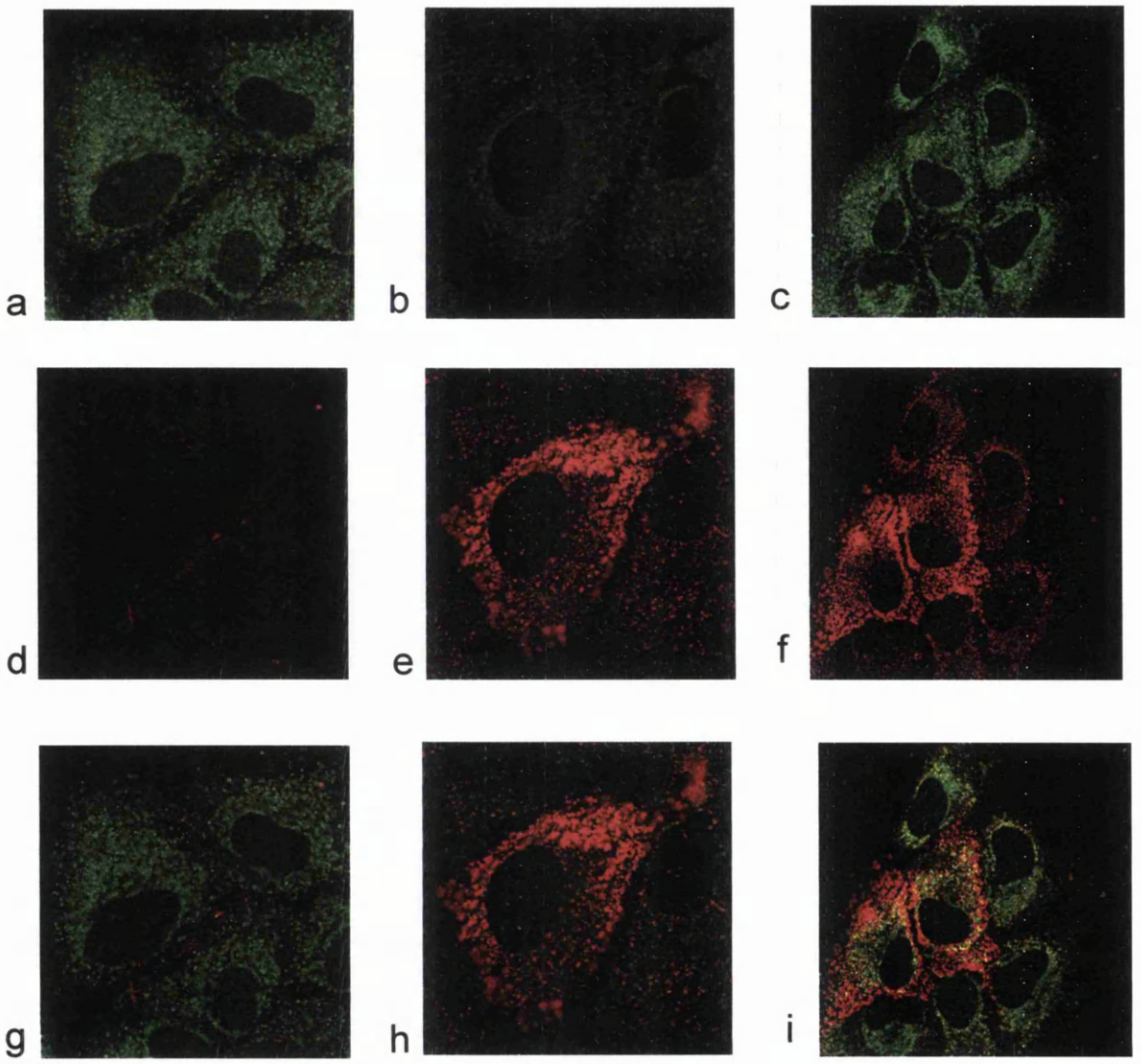

a
b
c
d
e
f
g
h
i

Figure 5.14. Localisation of PERK and calnexin. Huh7 cells were fixed in methanol and processed for indirect immunofluorescence. Cells in panels (a), (d), (g) and cells in panels (b), (e), (h) were probed singly with mouse anti-calnexin (SPA-866) and goat anti-PERK (N-18), respectively. Cells in panels (c), (f) and (i) were doubly probed with anti-calnexin (SPA-866) and anti-PERK (N-18). In all panels, bound antibodies were detected using rabbit anti-mouse IgG FITC and rabbit anti-goat IgG Cy3. The merged images are shown in panels (g), (h) and (i).

both PERK and calnexin and therefore panel (i) showed areas of colocalisation in yellow. This dual staining for PERK and calnexin did not provide convincing evidence of colocalisation and therefore casts doubt on the specificity of the anti-PERK antibody. An attempt at confirming the specificity of the antibody by Western blot analysis proved negative, as the expected 124K protein representing PERK was not seen (data not shown).

5.2.4. Discussion

Initially, dual staining for PERK and E2 showed evidence of extensive colocalisation (Figure 5.11). This was expected as PERK is known to localise to the ER (Harding *et al.*, 1999). However the appropriate controls indicated that this was not a true result as even when the PERK antibody was omitted, there was evidence of PERK staining (Figure 5.12). In fact this was due to cross-reaction of the two secondary antibodies. This was corrected by using a secondary antibody generated in a different species and the dual staining repeated (Figure 5.13). This time, the extent of colocalisation was not as comprehensive, although distinct regions of colocalisation were visible in the perinuclear area of the cell. However, the localisation of PERK did not resemble that of an ER-localised protein and unfortunately no positive control was available to test the specificity of the antibody. To confirm that PERK was not showing ER localisation, cells were probed for both PERK and calnexin, which is an ER-marker (Figure 5.14). This showed no evidence for colocalisation of PERK and calnexin, implying that PERK does not localise to the ER. This would contradict all previously reported data, so the specificity of this antibody has to be questioned. The antibody also failed to detect a protein of the correct predicted size by Western blot (data not shown). Another phenomenon noted in the immunofluorescence studies and visible in Figure 5.14, was the increased levels of expression of PERK (or the protein recognised by the antibody) in some cells. These were untransfected cells, so cannot be attributed to an overly expressed protein. One possible explanation is that the expression of this protein is affected by the cell cycle. As the specificity of the antibody could not be assured, this was not investigated further.

This particular commercial antibody was raised against an N-terminal peptide. Discussion with the manufacturer revealed that it had been tested against an expressing cell line (of unknown type) and they expressed doubt that the level of cellular PERK would be high enough for detection. Over-expression of a protein can sometimes lead to defects in processing, so it is possible that although the peptide used to raise the antibody is accessible in the expressing cell-line, it is hidden in the correctly processed form of PERK. In addition, it is possible that the different cell types could lead to differences in reactivity, although there is no evidence to suggest this. Another anti-PERK antibody, raised against a C-terminal peptide, is available but the high expense of these commercial antibodies prevented this from being pursued further.

The failure to investigate a link between E2 and PERK was disappointing and if time had permitted, it would have been worthwhile examining the relative levels of phosphorylated and unphosphorylated forms of eIF2α in E2-expressing and non-expressing cells. Of course it would not be possible to attribute a decrease in the level of phosphorylated eIF2α to either PKR or PERK directly, or for that matter to the other two known mammalian eIF2α-kinases, HRI and GCN2. However it would give insight as to the effect of HCV E2 on downstream signalling events from these kinases. As mentioned earlier, due to the intra-lumenal location of E2, it is unlikely to be in close proximity to the cytoplasmic (and kinase) domain of PERK. However, as part of the unfolded protein response (UPR), the misfolded forms of E2 should be targeted for degradation. This involves them being directed to the translocon (in the ER membrane) and retranslocated into the cytosol. Here, they are ubiquitinated and degraded by the proteosome (Plemper & Wolf, 1999). Therefore, during this process, E2 may be in close proximity to the kinase domain of PERK. In the absence of a cell-culture system, the high proportion of misfolded HCV glycoproteins has previously been attributed to over-expression of these proteins. The possibility that these misfolded proteins are part of a viral strategy aimed at preventing the cell from initiating a general shutdown in translation (and ultimately cell death), cannot be overlooked. If true, it is likely that PERK would play a significant role in this process.

Chapter 6

Binding of genotypes 1 and 3 E2 to human CD81

6.1. Introduction

The putative position of E1 and E2 in the viral envelope suggests that they are directly involved in the attachment of the virus to the host cell. It also implies that they would be targets for neutralising antibodies. The lack of a cell culture system for HCV made it difficult to pursue this line of questioning until the development of a surrogate test for binding of virus to the cell, called the neutralisation of binding (NOB) assay (Rosa *et al.*, 1996). This assay was based on the ability of antibodies or sera to prevent binding of E2 to the human MOLT-4 T-cell line. E2 bound to these cells with high affinity and binding was inhibited by sera from E1/E2-vaccinated chimpanzees, which were partially protected from challenge with infectious sera. "Neutralising antibodies" were also found in the sera of chronically-infected HCV patients and, in those who resolved the infection naturally, the emergence of high NOB titres coincided with viral clearance (Ishii *et al.*, 1998). This indicated that these antibodies might be responsible for blocking HCV infection *in vivo* and implicated E2 in binding of HCV to target cells. The cellular molecule responsible for binding E2 was identified by screening a cDNA library prepared from the MOLT-4 cell line and identifying a clone which conferred E2-binding to mouse fibroblasts upon transient transfection (Pileri *et al.*, 1998). The clone encoded the human CD81 protein, leading to its identification as a putative HCV receptor. CD81 is a member of the tetraspanin family of proteins. It traverses the membrane four times and has two extracellular loops. The E2-binding site was mapped to the large extracellular loop (LEL)

(Pileri *et al.*, 1998), which is the most variable region of the protein but significantly, is highly conserved in humans and chimpanzees, the only known species susceptible to HCV infection. However, E2 has since also been shown to bind CD81 of tamarins (Meola *et al.*, 2000, Allander *et al.*, 2000a), which are not susceptible to HCV infection, so binding of E2 to CD81 alone is apparently insufficient to predict species permissiveness to infection. The wide expression of CD81 throughout human tissue (except red blood cells and platelets) does not explain the cellular tropism of HCV either, although importantly, E2 has been shown to bind to human hepatocytes (Flint *et al.*, 1999a).

Following the identification of the human cell surface molecule (CD81) which interacts with HCV E2, there have been several reports aimed at elucidating the structural and functional characteristics of E2 that are important for this interaction. In the NOB assay, Rosa *et al.* (1996) demonstrated that sera from patients infected with genotypes 1, 2 and 3 were all able to neutralise binding of genotype 1 E2. This implies that binding of E2 involves a region or regions that are conserved between genotypes. However, almost all studies on CD81-binding have concentrated exclusively on E2 of genotype 1a. This bias was first addressed by Allander *et al.* (2000b) with the isolation of human monoclonal antibodies which were selected for reactivity against genotype 1a E2 from a genotype 2b-infected patient. Some of these antibodies also recognised genotype 1b E2 but recombinant E2 protein for the other genotypes was not available for further analysis. Four of these conformation-dependent antibodies tested positive in the NOB assay with genotype 1a. This was interpreted as inhibition of E2 binding to CD81 but this cannot be verified from these data, as the CD81/E2 interaction was not examined directly. In a similar study, Hadlock *et al.* (2000) isolated monoclonal antibodies from a genotype 1b-infected patient by selection for reactivity against genotype 1a E2. Some of these antibodies were also cross-reactive with genotypes 2a and 2b E2 and, as with the previous study, almost all recognised conformational epitopes. Six antibodies tested positive in the NOB assay with genotype 1a E2 and all six were capable of inhibiting the interaction of genotypes 1a, 1b, 2a and 2b E2 with CD81. Surprisingly these antibodies were also able to recognise the pre-formed E2/CD81 complex, although with some antibodies this was genotype specific. This suggests that the antibodies recognise epitopes that are not

directly involved in binding CD81 and possibly they inhibit CD81 binding by preventing a conformational change in E2.

6.2. Lack of binding of genotype 3 E2 to human CD81.

6.2.1. E2 capture by human CD81.

For evaluating the ability of E2 to bind human CD81, E2 was expressed from the SFV constructs described in Table 4.1 and crude cell extract expressing E2 was prepared for use in ELISA. BHK cells were used, as human cells would provide endogenous CD81 which would reduce the amount of E2 available for capture. Pileri *et al.* (1998) have mapped the binding of genotype 1 E2 to the major extracellular loop of human CD81 and this portion of the protein was expressed as a GST fusion protein and was kindly provided by Arvind Patel. For the sake of simplicity, this will be referred to as CD81. The design of the capture ELISA is shown in Figure 6.1. It was essentially the same as that described in chapter 4 for testing antibody reactivity apart from the use of CD81 as the capture protein and the use of two-fold dilutions of crude cell extract expressing E2 to evaluate the effect of E2 concentration on binding. As there was no guarantee that the E2 proteins would be expressed at equivalent levels from all constructs, before evaluating CD81 binding, two-fold dilutions of E2 cell extract were captured in a GNA ELISA and bound E2 detected with anti-E2 antibody. The dilution of E2 extract that gave equivalent absorbance readings was determined for each construct and this was taken into account when calculating the starting dilution of E2 extract for the binding ELISA. This ensured that the relative amounts of E2 were equivalent in terms of reactivity with the anti-E2 antibody that would be used for detecting bound E2. Each CD81 capture ELISA was performed in parallel with a GNA capture ELISA as a control for all the remaining interactions required for a positive reaction.

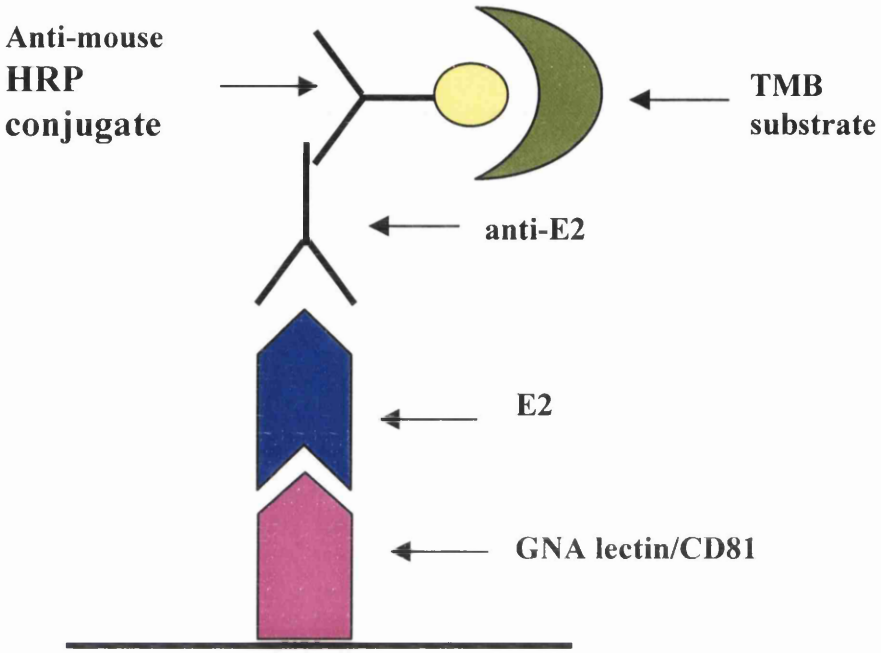


Figure 6.1. Diagrammatic representation of the *Galanthus nivalis* agglutinin (GNA) lectin or human CD81 capture ELISA. ELISA plates were coated with GNA lectin or CD81 overnight at room temperature. Following blocking, two fold dilutions of crude cell extract expressing E2 was added. Bound E2 was detected by the addition of mouse monoclonal anti-E2, anti-mouse peroxidase-conjugated antibody and tetramethyl benzidine (TMB) substrate.

Figure 6.2 shows the results of a GNA capture ELISA (a) and a CD81 capture ELISA (b) where ALP98 was used for detecting bound E2. Both genotype 3 (R3) and genotype 1 (H77 and Gla) E2 were analysed. Constructs expressing both E2 alone and in the context of core and E1 were included for R3 and H77 to see if this had an effect on CD81 binding. Figure 6.2 (a) shows binding to GNA for all E2-expressing constructs. The degree of binding was approximately equivalent for all constructs and decreased as the E2 concentration decreased. In the CD81 capture ELISA (b) only the H77 constructs showed evidence of binding. The level of CD81 binding was similar for H77 E2 expressed alone and in addition to core and E1. It should be noted that there was less binding of H77 E2 to CD81 than to GNA. For example, with the H77 E1/E2 construct, compare an absorbance reading of 1.5 with CD81 to a reading of over 2.5 with GNA at a ¼ dilution (2 on the X-axis). The same experiment is shown in Figure 6.3 except that AP266 was used for detecting bound E2. As with the previous experiment, only H77 E2 was capable of binding to CD81, although this time it appeared that E2 expressed in conjunction with core and E1 bound CD81 better than E2 expressed alone. Again, the level of CD81 binding was much reduced compared to GNA binding of E2 for the same constructs. The genotype 3 F2 construct was also analysed for CD81 binding in Figure 6.3 but as with R3, it did not bind CD81.

One of the complicating factors in interpreting the binding data in Figures 6.2 and 6.3 was that crude cell extract was used and therefore the amount of E2 protein was unknown. As described earlier, this was controlled by predetermining the dilution of extract that gave equivalent reactivity with a particular antibody in a GNA capture ELISA for different constructs. In these assays, genotype 1 E2 consistently showed increased reactivity with ALP98 compared to the same dilution of the genotype 3 E2 extract (data not shown). This suggested that either the expression level of genotype 3 E2 was not as high as that of genotype 1 E2 or that there was a difference in reactivity of ALP98 with types 1 and 3 E2. To investigate these possibilities, the same dilutions of E2 extract that showed this phenomenon were analysed by Western blot for reactivity with ALP98 (Figure 6.4). ALP98 recognised genotype 3 (R3) E2 down to a dilution of 1/512 (a) and genotype 1 (H77) E2 down to a dilution of 1/256 (b). Therefore, this data contradicted what was

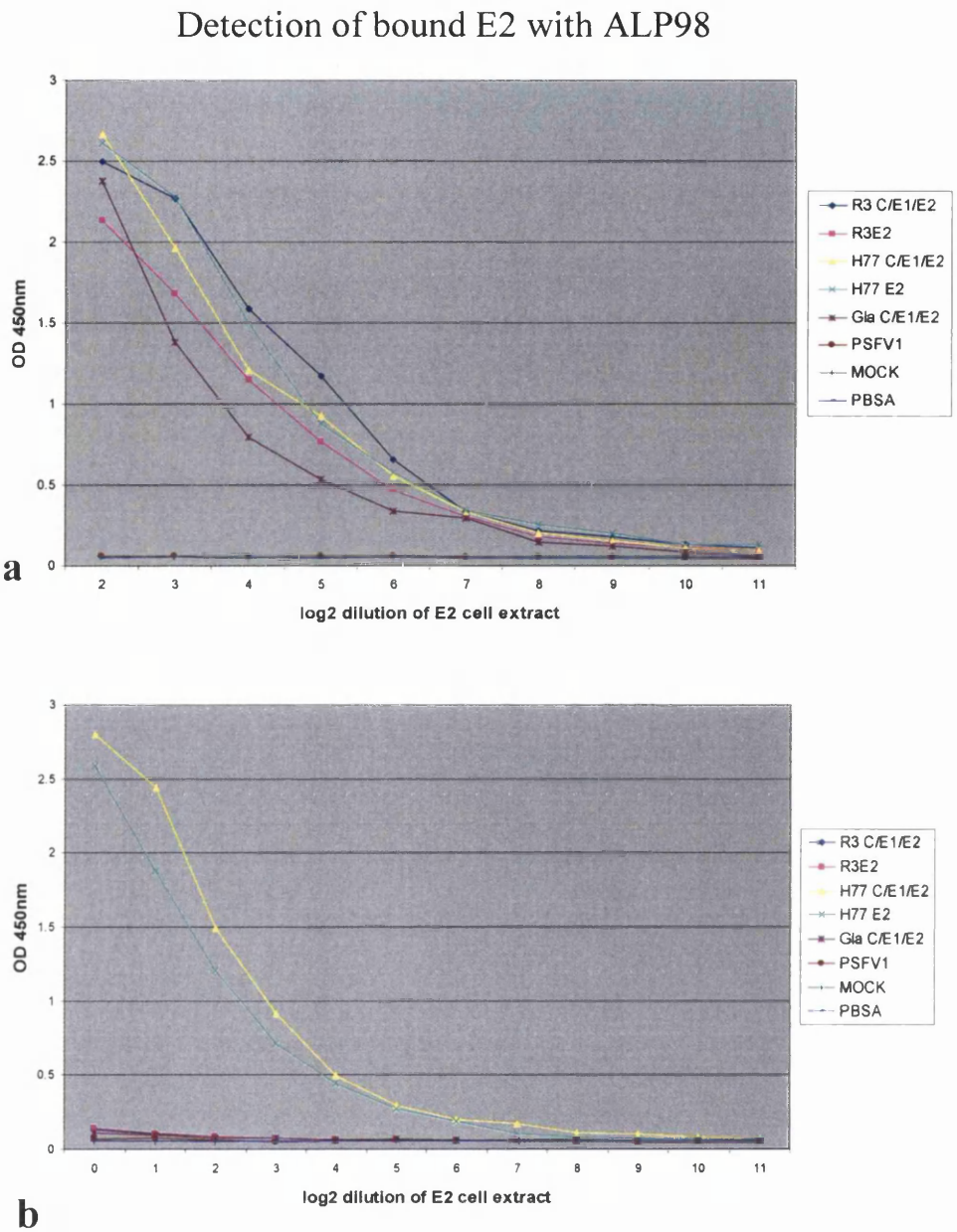


Figure 6.2. Capture ELISA of two-fold dilutions of genotype 1 and 3 E2 binding to GNA (a) and human CD81 (b) as measured by light absorbance at 450nm. The ELISA was set up as described in Fig. 6.1. and mouse monoclonal anti-E2, ALP98, was used for detecting bound E2. E2-expressing extracts were prepared by electroporating BHK cells with *in vitro* transcribed RNA from genotype 3 constructs (pSFVR3CE1E2, pSFVR3E2), genotype 1 constructs (pSFVH77CE1E2, pSFVH77E2, pSFVGlaCE1E2) and pSFV1.

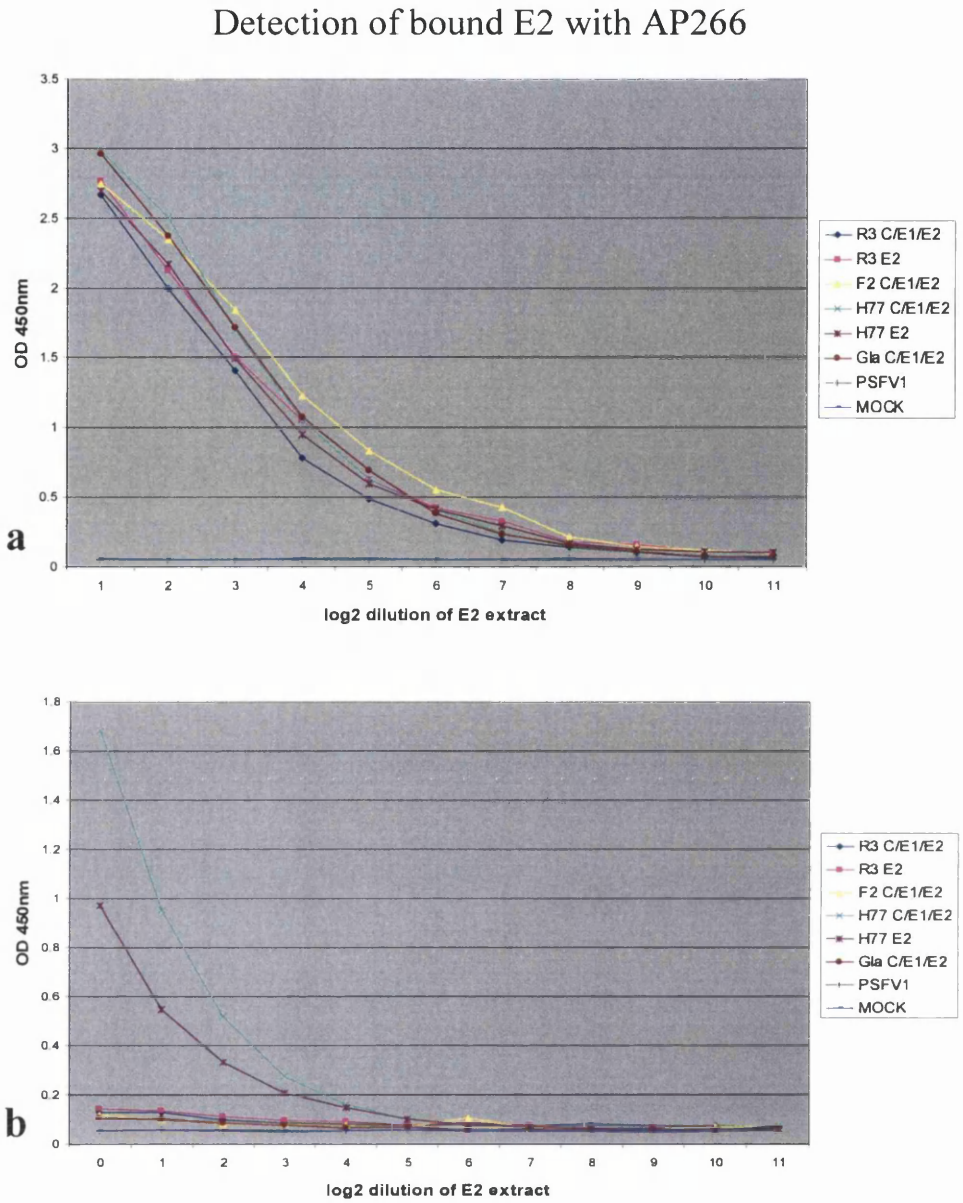


Figure 6.3. Capture ELISA of two-fold dilutions of genotype 1 and 3 E2 binding to GNA (a) and human CD81 (b) as measured by light absorbance at 450nm. The ELISA was set up as described in Fig. 6.1. and mouse monoclonal anti-E2, AP266 was used for detecting bound E2. E2-expressing extracts were prepared by electroporating BHK cells with *in vitro* transcribed RNA from genotype 3 constructs (pSFVR3CE1E2, pSFVR3E2, pSFVF2CE1E2), genotype 1 constructs (pSFVH77CE1E2, pSFVH77E2, pSFVGlaCE1E2) and pSFV1.

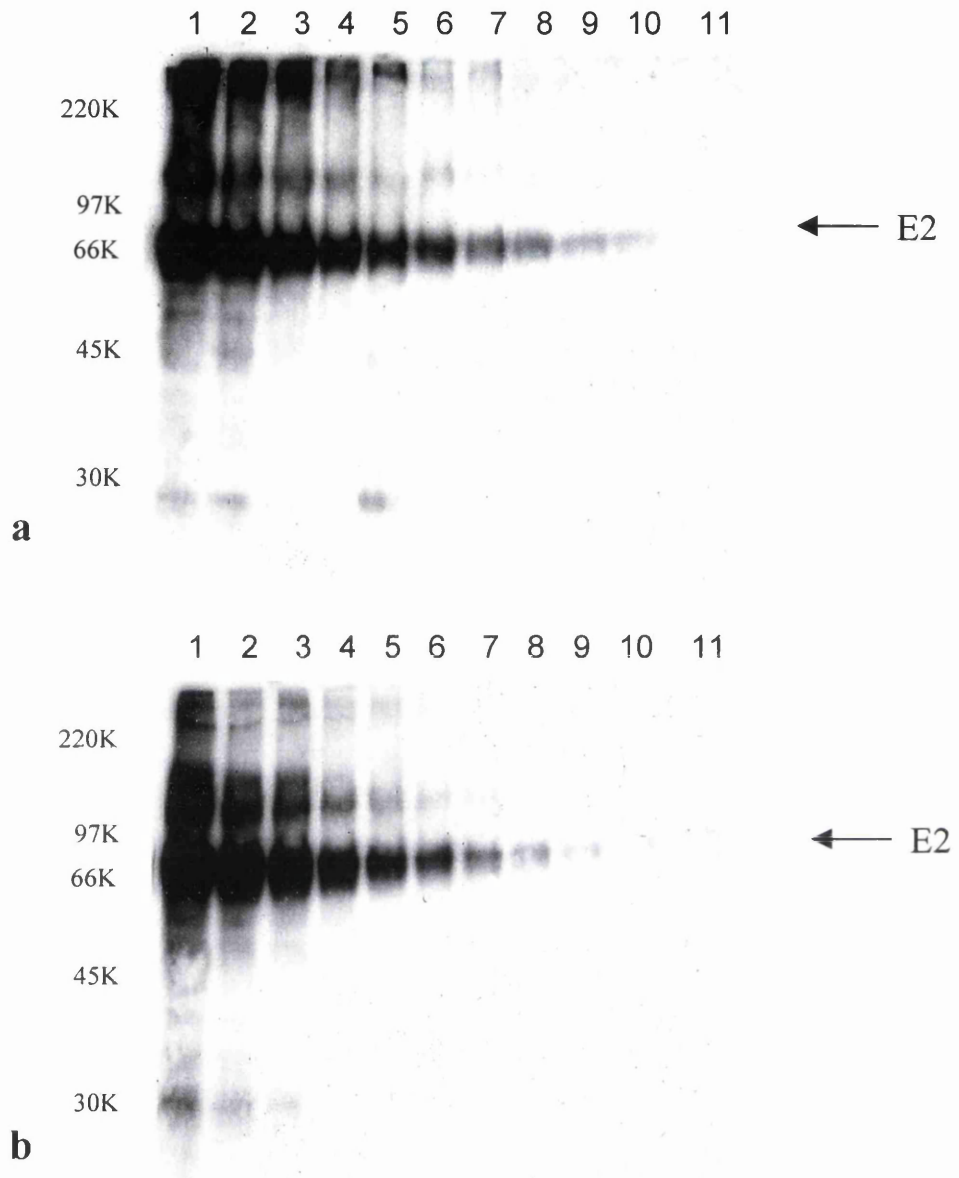


Figure 6.4. Comparison of genotype 3 (a) and genotype 1 (b) E2 expression levels. BHK cells were electroporated with *in vitro* transcribed RNA from pSFVR3CE1E2 (a) or pSFVH77CE1E2 (b) and crude cells extracts were separated by SDS-PAGE and analysed by Western blot using anti-E2, ALP98. Lanes 1-11 represent two fold dilutions of E2-expressing cell extract from neat to 1/1024.

shown in ELISA analysis. A possible explanation is that GNA had more affinity for H77 E2 due to the predicted extra glycan.

The amount of protein detected by ALP98 was always greater when bound to GNA than to CD81. This may be because CD81 bound only native E2, whereas GNA also bound aggregated protein. To assess the ability of both native and misfolded E2 to bind CD81, the binding of H77 E2 to GNA and CD81 was analysed by detection with conformation-dependent antibodies (Figure 6.5). If only native E2 binds CD81, binding to GNA and CD81 should be equivalent when E2 is detected with a conformation-dependent antibody. When detected with a linear-epitope antibody, the binding of E2 to GNA should be higher than to CD81 as GNA will bind both aggregated and native E2. Three different conformation-dependent antibodies (H50, H53 and H60) and a linear epitope antibody (ALP98) were used to detect E2 bound to both GNA and CD81. As previously described, E2 detected with ALP98 showed increased binding to GNA compared to CD81 and in agreement with the theory, E2 detected with H60 showed equivalent binding to GNA and CD81. However, the reactivities of neither H50 nor H53 concurred with that of H60. H50 displayed a reactivity pattern similar to ALP98. H53 detected equivalent amounts of E2 binding to CD81 as did H60 but in comparison, it detected increased binding to GNA.

6.2.2. Genotype 3 E2 does not compete with genotype 1 E2 for CD81 binding.

In Figures 6.2 and 6.3, slight increases in CD81 binding above non-specific levels were observed for genotype 3 (and genotype 1 Gla strain) E2-expressing cell extracts at the highest concentrations. This may represent binding on a much smaller scale than demonstrated for H77 E2. A procedure for determining if this is the case, is to assess whether genotype 3 E2 has the ability to compete with genotype 1 E2 for CD81 binding. By using an antibody that recognises only type 1 E2, it is possible to compare the CD81-binding ability of type 1 E2 in the presence and absence of genotype 3 E2. If genotype 3 E2 is able to compete for CD81 binding, the binding of genotype 1 E2 is expected to

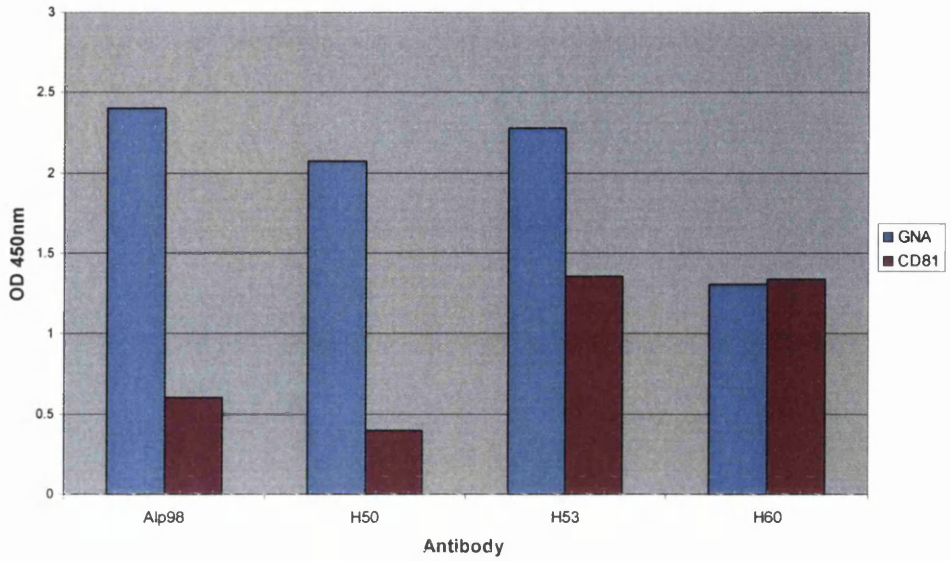


Figure 6.5. Comparison of genotype 1 E2 binding to GNA and human CD81 when detected with conformational antibodies. BHK cells were electroporated with *in vitro* transcribed RNA from pSFVH77E2. E2-expressing cell extract was added to both GNA and CD81 coated ELISA plates and bound E2 was detected with a linear epitope antibody (ALP98) and conformational epitope antibodies (H50, H53 and H60).

decrease in the presence of type 3 E2. In Figure 6.6 (a) this was analysed by keeping the amount of type 1 E2 constant and increasing the amount of type 3 E2 (blue line) and vice versa (pink line). As the concentration of E2 was unknown, the relative amounts of types 1 and 3 E2 were expressed as percentages ranging from 0% to 50%. The total amount of protein was made up with cell extract expressing the SFV proteins from the pSFV1 vector. E2 bound to CD81 was detected with AP320, which only recognised type 1 E2. There was no indication of a decrease in binding of genotype 1 E2 as the amount of type 3 E2 was increased. To determine whether genotype 3 E2 has an effect on the ability of genotype 1 E2 to bind CD81 in a concentration-dependent manner, the binding of increasing amounts of type 1 E2 was assessed both in the presence (blue line) and absence (pink line) of genotype 3 E2 (Figure 6.6 (b)). If type 3 E2 were able to compete for CD81 binding with type 1 E2, the blue curve would be expected to shift to the right when compared to the pink curve. However, no difference was observed, hence type 3 E2 did not interfere with the CD81-binding ability of type 1 E2.

6.2.3. Lack of CD81 binding to E2 when ELISA is performed in reverse.

Several studies have used the ELISA format described in Figure 6.1 to demonstrate binding of genotype 1 E2 to human CD81. However this form of sandwich ELISA has not been used to prove the reverse binding i.e. binding of CD81 to E2. Demonstration that an interaction between two proteins is valid for both orientations in an assay provides more convincing evidence that the interaction is biologically significant. The format for analysing the binding of CD81 to E2 is illustrated in Figure 6.7. E2 was captured by anti-E2, ALP98. As with the previous ELISA, E2 was provided in the form of crude cell extract from E2-expressing cells and was used in two-fold dilutions. The CD81 protein, expressed as a GST fusion protein, was added and bound CD81 was detected with rabbit anti-GST antibody, anti-rabbit HRP-conjugated antibody and addition of TMB substrate. Both genotype 3 (R3) and genotype 1 (H77 and Gla) E2 were analysed, even though only H77 E2 had previously been shown to bind CD81. However, in this format, CD81 did

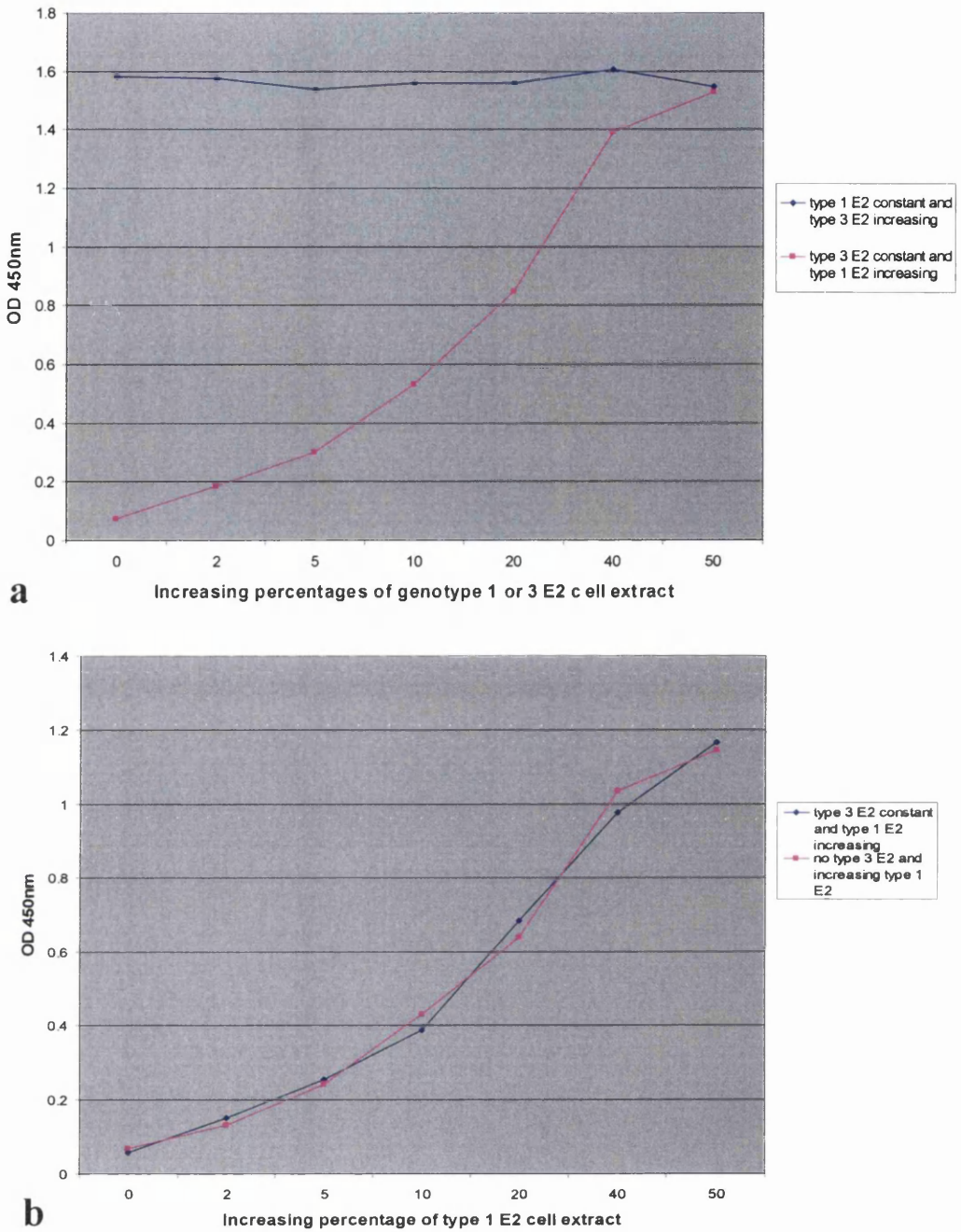


Figure 6.6. Analysis of the ability of type 3 E2 to compete with type 1 E2 for CD81 binding. E2-expressing cell extracts were made as described before from the relevant type 1 and type 3 constructs, pSFVH77CE1E2 and pSFVR3CE1E2 respectively. E2 extracts were added to CD81-coated wells and bound E2 was detected by anti-E2 mouse monoclonal, AP320 which is specific for type 1 E2. In panel (a), the amount of type 1 (blue line) or type 3 (pink line) E2 extract remained constant at 50% of the total E2 content while the amount of type 3 or type 1 E2 respectively, was increased from 0-50%. In panel (b) the amount of type 3 E2 either remained constant (blue line) or was absent (pink line), while the amount of type 1 E2 was increased.

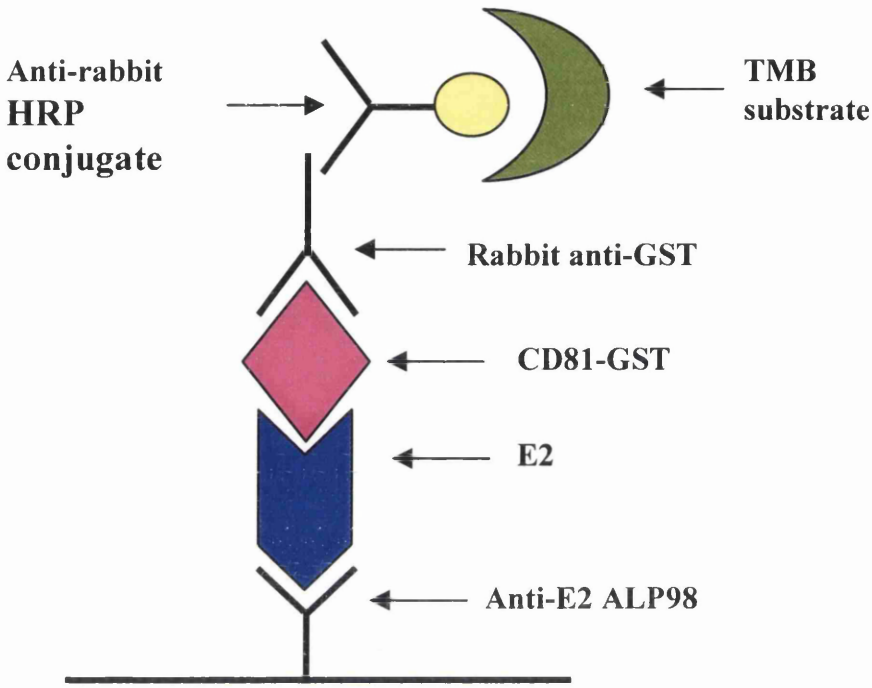


Figure 6.7. Diagrammatic representation of the reverse human CD81 ELISA. ELISA plates were coated with mouse monoclonal anti-E2, ALP98 overnight at room temperature. Following blocking, two fold dilutions of crude cell extract expressing E2 was added. The human CD81-GST protein was added and bound CD81 was detected by the addition of rabbit polyclonal anti-GST antibody, anti-rabbit peroxidase-conjugate and tetramethyl benzidine (TMB) substrate.

not bind to any of these E2 proteins (Figure 6.8). The absorbance values were lower than those observed for E2 binding to CD81 in Figures 6.2 and 6.3 and were independent of E2 concentration with fluctuating absorbance values seen for both E2-expressing and non-E2 expressing cell extracts. There was a decrease in absorbance between dilutions of 1/32 and 1/128 but this applied to non-E2 expressing cells as well, so may be due to the decrease in concentration of a cellular protein. This was supported by the observation that PBSA did not follow this pattern. As this ELISA relied on several individual reactions, it was necessary to confirm these before stating that CD81 did not bind E2. Firstly, E2 expression from the R3, H77 and Gla constructs was verified by GNA-capture ELISA with detection by ALP98 (Figure 6.9(a)). This reverse ELISA relied on E2 capture by ALP98, which was demonstrated in Figure 6.9 (b) by detection with anti-E2 polyclonal antibody, R646. The level of non-specific reactivity was higher here than previously observed. Evidence that the anti-GST antibody recognised the CD81-GST fusion protein is provided in Figure 6.9 (c) where antibody was added to CD81-coated wells and reactivity detected by anti-rabbit HRP-conjugated antibody. There was no evidence of non-specific reactivity between the conjugated antibody and CD81 when the GST antibody was absent. Finally, CD81 was added to ALP98-coated wells and any binding detected with anti-GST (Figure 6.9(d)). There was no evidence of non-specific binding of CD81 to ALP98 (blue line) and when CD81 was removed from the reaction (pink line), the same pattern was observed, indicating that there was no cross-reactivity between ALP98 and anti-GST. Therefore the lack of observed binding of CD81 to E2 in Figure 6.8 could not be attributed to any of the other required interactions in this ELISA.

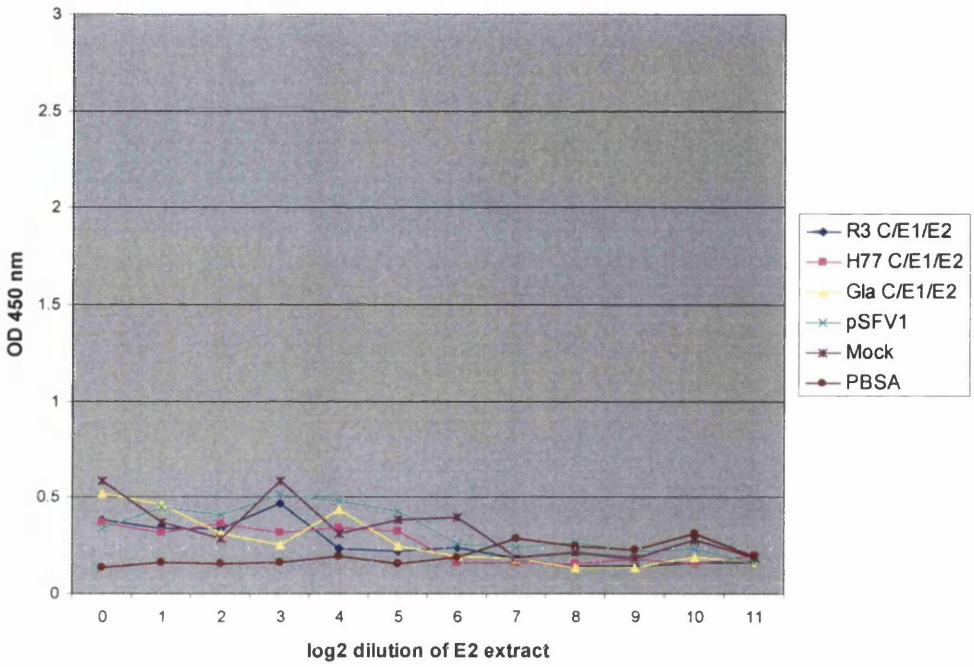
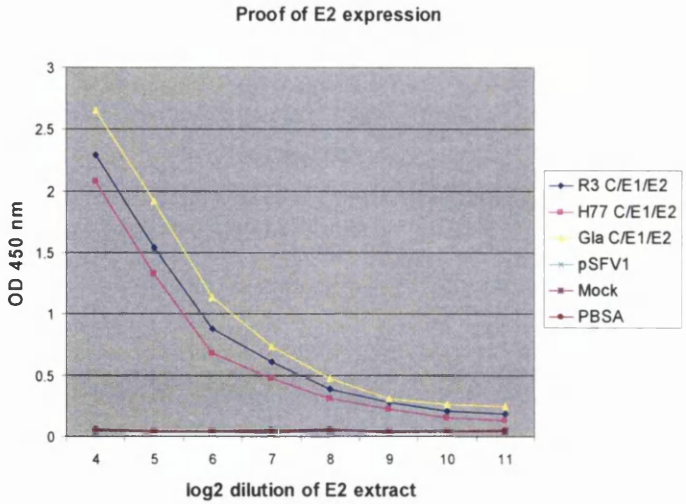
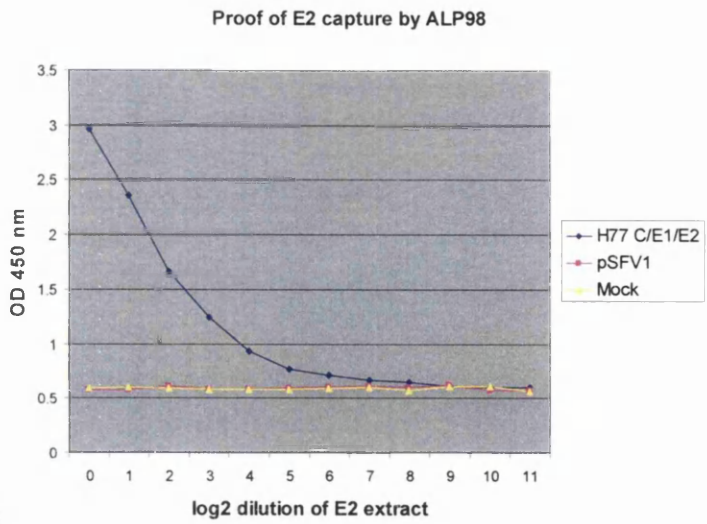


Figure 6.8 ELISA demonstrating failure of CD81 to bind E2 when E2 is captured by anti-E2, as measured by light absorbance at 450 nm. The ELISA was set up as described in Fig. 6.7 using mouse monoclonal, ALP98 to capture E2 and rabbit anti-GST to detect bound CD81-GST protein. E2-expressing extracts were prepared by electroporating BHK cells with *in vitro* transcribed RNA from a genotype 3 construct (pSFVR3CE1E2), genotype 1 constructs (pSFVH77CE1E2,pSFVGlaCE1E2) and pSFV1.

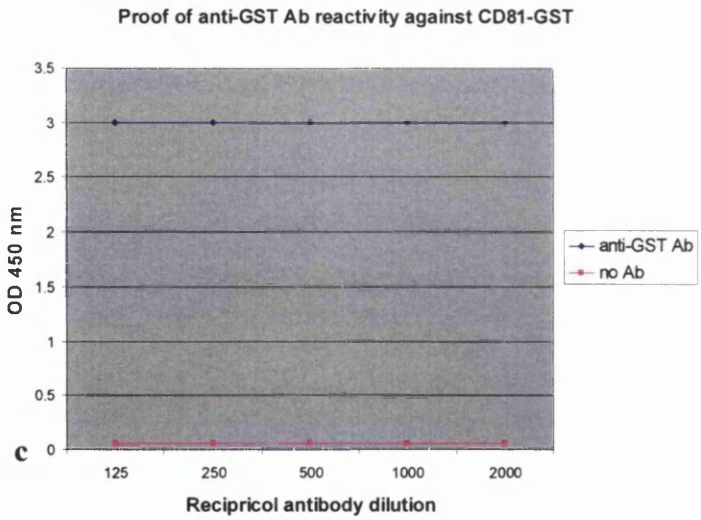
Figure 6.9. Controls for the reverse CD81 ELISA, showing evidence of all the individual interactions apart from that of E2 and CD81. Evidence of E2 expression is shown by a GNA capture ELISA with detection by ALP98 and anti-mouse HRP (a). E2-capture by ALP98 and detection with R646 and anti-rabbit HRP is shown in (b). Reactivity of the anti-GST antibody against CD81-GST when the wells are coated with CD81 and antibody binding detected with anti-rabbit HRP, is shown in (c). Figure (d) shows lack of CD81 binding to ALP98-coated wells when CD81 was detected with anti-GST antibody and anti-rabbit HRP (blue line) and no cross reactivity between anti-GST and ALP98 when CD81 was omitted (pink line). All binding was measured in terms of light absorbance at 450 nm following reaction of horseradish peroxidase (HRP) with TMB substrate and addition of H_2SO_4.


Proof of E2 expression
OD 450 nm
log2 dilution of E2 extract
R3 C/E1/E2
H77 C/E1/E2
Gla C/E1/E2
pSFV1
Mock
PBSA

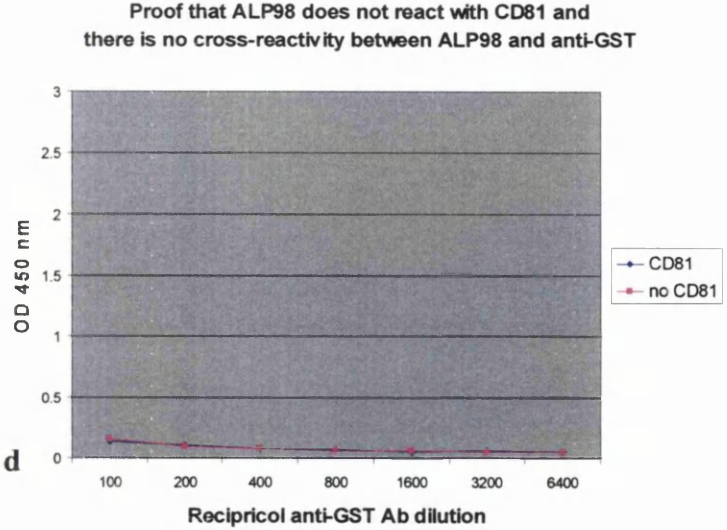
a


Proof of E2 capture by ALP98
OD 450 nm
log2 dilution of E2 extract
H77 C/E1/E2
pSFV1
Mock

b


Proof of anti-GST Ab reactivity against CD81-GST
OD 450 nm
Recipricol antibody dilution
anti-GST Ab
no Ab

c

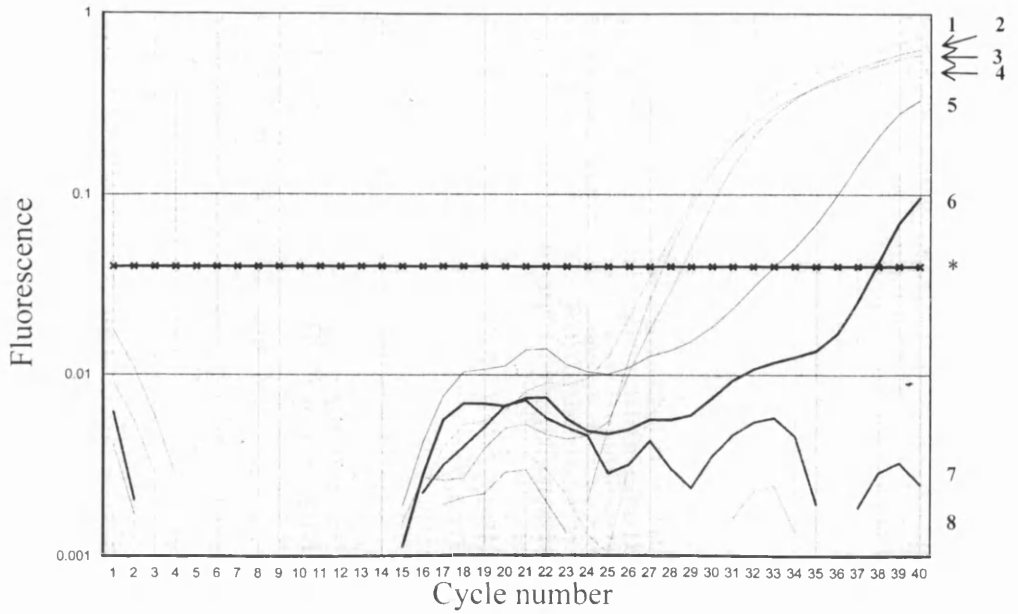

Proof that ALP98 does not react with CD81 and
there is no cross-reactivity between ALP98 and anti-GST
OD 450 nm
Recipricol anti-GST Ab dilution
CD81
no CD81

d

6.3. Binding of virus particles from serum to human CD81

The potential role of an E2 interaction or non-interaction with CD81 cannot be interpreted fully by investigating E2 in isolation, as it may not reflect the structure of E2 in the virion. CD81 has been proposed as a candidate HCV receptor based on the finding that free virions in chimpanzee infectious sera (genotype 1a) bind human CD81 (Pileri *et al.*, 1998). To add to the ELISA evidence that genotype 3 E2 does not bind CD81, I attempted to show that genotype 3 virions do not bind CD81 under conditions where genotype 1 virions do bind.

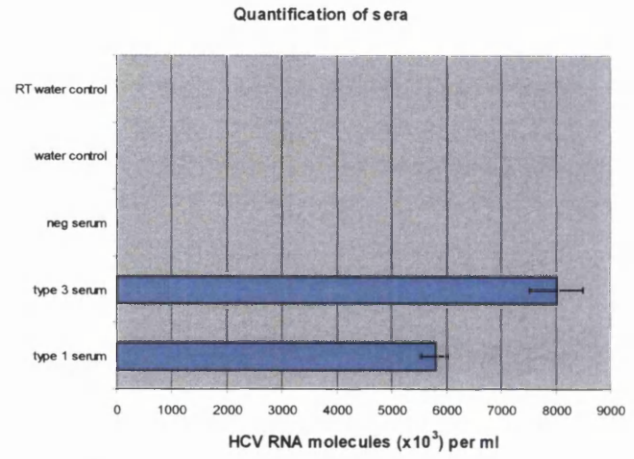
Chimpanzee infectious plasma, as used by Pileri *et al.* (1998) and Hadlock *et al.* (2000), was not obtainable for this study. However both genotypes 1 and 3 HCV positive human sera were available. A quantitative PCR detection system was used to assess the binding of virion-associated HCV RNA to human CD81. This involved coating agarose beads with the CD81-GST fusion protein and incubating these beads with the human sera. The number of virions remaining bound to the beads were determined by RT-PCR using the GeneAmp 5700 sequence detection system (Applied Biosystems). A dilution series of a serum of known titre, calibrated against the international standard, was included for generation of a standard curve. Figure 6.10 shows an example of the PCR result for the unknown genotypes 1 and 3 sera and the dilution series of the standard. The amount of PCR product generated was measured by emission of fluorescence and plotted against PCR cycle number. A standard curve of RNA copies per ml versus cycle number was drawn for the known serum and this was used to determine the titre of the unknown sera. All PCR reactions were performed in triplicate but only one of each is shown in Figure 6.10 so as to not complicate the figure unnecessarily. The calculated average titres for the unknown genotype 1 and genotype 3 sera are shown in Figure 6.11(a). These sera were incubated with either CD81-GST-coated beads (CD81) or GST-coated beads (beads) and the average number of bound virus particles are shown in Figure 6.11(b). With genotype 3 there was more binding to CD81 than beads alone but this was reversed for genotype 1.



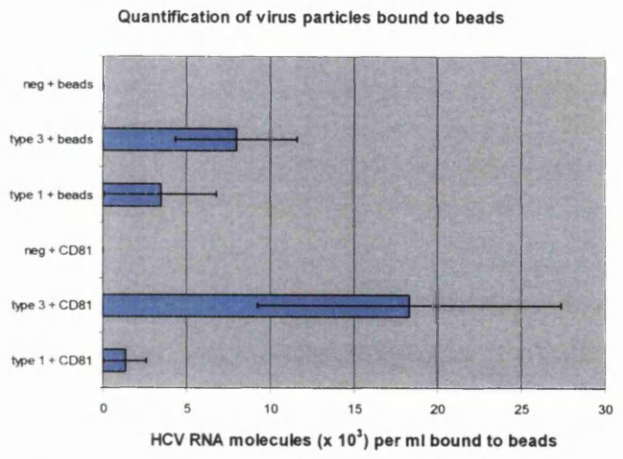
1	unknown type 3 serum
2	undiluted standard $3x10^6$ copies /ml
3	unknown type 1 serum
4	Standard $1x10^6$ copies /ml
5	Standard $1x10^5$ copies /ml
6	Standard $1x10^4$ copies /ml
7	RT control
8	No template control

Figure 6.10. Depiction of the amount of PCR product at each PCR cycle as determined by the amount of fluorescence measured by the GeneAmp 5700 sequence detection system. Everything above the threshold (*) is considered positive. A dilution series of a serum of known titre is included and a standard curve of RNA copies /ml vs Cycle number allows the titre of an unknown serum to be determined. The result of this calculation for the unknown types 1 and 3 sera are shown in Figure 6.11(a).

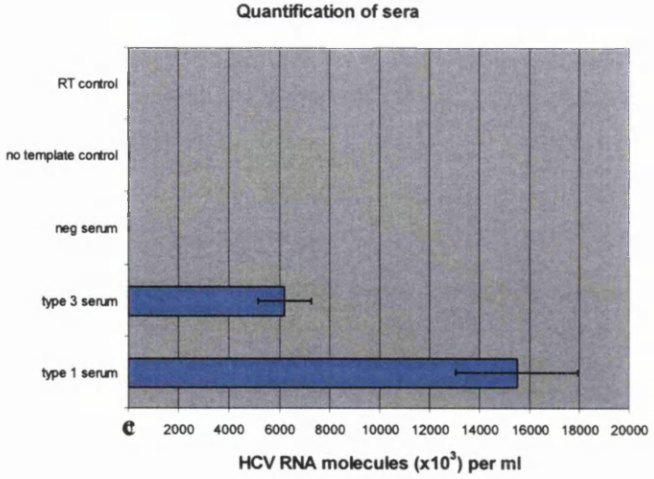
Figure 6.11. Quantification of genotype 1 and genotype 3 HCV virus particles bound to CD81-GST coated beads and GST-coated beads. Two independent experiments are shown (a&b and c&d). The HCV positive sera used in the binding experiments were quantified by RT-PCR against a standard (a and c). In figure (a) the type 1 serum was diluted so that the viral titre was approximately equivalent to that of the type 3 serum. In figure (b) both sera were undiluted. The same type 1 serum was used for both experiments but a different type 3 serum was used for each experiment. Human CD81 expressed as a GST fusion protein was bound to glutathione agarose beads as was GST alone. The beads were blocked in a 2% BSA solution and washed in TEN buffer. A 200 µl aliquot of either type 1, type 3 or HCV-negative sera was added and the beads incubated at 37°C. After thorough washing in TEN buffer, bound virus was eluted in lysis buffer. The amount of bound virus was determined by quantification of HCV RNA by RT-PCR against a standard (b and d). All PCR reactions were performed in triplicate and the standard deviation is indicated.


Quantification of sera
RT water control
water control
neg serum
type 3 serum
type 1 serum
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
HCV RNA molecules (x10³) per ml

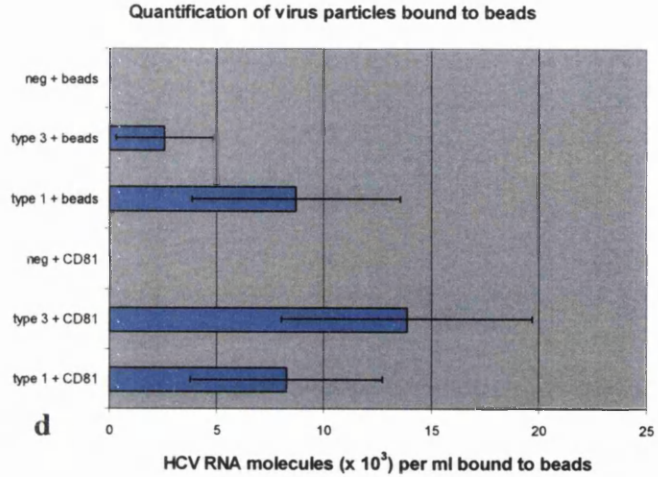
a


Quantification of virus particles bound to beads
neg + beads
type 3 + beads
type 1 + beads
neg + CD81
type 3 + CD81
type 1 + CD81
0
5
10
15
20
25
30
HCV RNA molecules (x 10³) per ml bound to beads

b


Quantification of sera
RT control
no template control
neg serum
type 3 serum
type 1 serum
2000
4000
6000
8000
10000
12000
14000
16000
18000
20000
HCV RNA molecules (x10³) per ml

c


Quantification of virus particles bound to beads
neg + beads
type 3 + beads
type 1 + beads
neg + CD81
type 3 + CD81
type 1 + CD81
0
5
10
15
20
25
HCV RNA molecules (x 10³) per ml bound to beads

d

In this experiment the genotype 1 serum was diluted so that the viral titre was equivalent to that of the genotype 3 serum (a). To verify that this did not affect the binding to CD81, in the following binding experiment both sera were used undiluted (Figure 6.11(c)&(d)). Unfortunately not enough of the original genotype 3 serum was available, so an alternative was used in (c) and (d). As expected, the undiluted genotype 1 serum showed increased binding in Figure 6.11(d) compared to (b) but there was no notable difference between binding to CD81 and binding to beads alone. With the new genotype 3 serum there was increased binding with CD81 than with the beads only (d). Time limitations did not allow further experimentation.

6.4. Discussion

The only report of CD81 binding to E2 other than genotype 1 is that from Hadlock *et al.* (2000) which demonstrated binding of genotype 2 E2 to CD81. However, when recombinant genotype 3 E2 protein was analysed here for CD81 binding there was no reactivity (Figures 6.2b and 6.3b). This was not due to the antibody not recognising the E2/CD81 complex as strong reactivity was shown with the E2/CD81 complex of genotype 1a (H77). This dose-dependent binding curve for H77 E1/E2 and E2 is equivalent to that reported for the truncated H77 E2 protein (Patel *et al.*, 2000) and more significantly, H77 virus-like particles (VLPs) generated in insect cells (Owsianka *et al.*, 2001). A panel of anti-E2 antibodies was analysed for ability to inhibit these various forms of H77 E2 from binding CD81 (Owsianka *et al.*, 2001). Neither ALP98 nor AP266 were shown to inhibit the E2-CD81 interaction, so the epitopes recognised by these antibodies are not involved in the CD81 interaction, which is supported by the binding curves in Figures 6.2b and 6.3b. Interestingly, AP33 was shown to inhibit CD81 binding of truncated E2, full-length E1/E2 and VLPs, suggesting that the epitope recognised by AP33 (aa 412-423) may be in direct contact with CD81. In chapter 4 it was shown that AP33 does not recognise genotype 3 E2 and sequence comparison of the antibody epitope indicated only one amino acid difference, a lysine residue in R3 compared to an asparagine in H77 (Figure 4.6). The lysine side chain is larger than that of asparagine and is also positively charged whereas asparagine is uncharged, so this amino acid change could conceivably alter the

protein structure in this region, which may interfere with the E2/CD81 interaction. Comparison of the sequence variation at this position in the other four published genotype 3 isolates reveals a lysine at position 415 in HCVCENS1 but an asparagine in NZL1, AF046866 and HPCHK6 (Figure 3.4). Therefore the lysine at position 415 is not unique to Gla-3a but unfortunately there are no data available on the ability of E2 from these other isolates to bind CD81. To confirm that the lysine residue is responsible for the loss of CD81 binding it would be necessary to mutate this residue to an asparagine and determine if this confers CD81 binding on genotype 3 (R3) E2. Apart from a variable amino acid at position 414 (either valine or isoleucine, as in R3 and H77), the remainder of the epitope is fully conserved between all 5 genotype 3 isolates and H77, consistent with this being a critical region. Owsianka *et al.* (2001) also identified a downstream region (aa 528-535) that is possibly involved in the E2/CD81 interaction. It is not known if the antibody used for identifying this region recognises genotype 3 E2 but amino acid sequence comparison of genotypes 1 and 3 E2 shows that 4 out of the 8 residues in this epitope vary between genotype 1 and either 2 or more genotype 3 isolates (Figure 3.4). It is therefore unlikely that the antibody would recognise type 3 E2. Two different regions (aa 480-493 and 544-551) were identified as potential sites of CD81 interaction by Flint *et al.* (1999a). However, the antibodies used for the identification of those regions were shown not to inhibit CD81 binding in the study by Owsianka *et al.* (2001).

Flint *et al.* (2000) demonstrated that the aggregated form of E2 did not bind CD81 and the intracellular forms of E2 bound CD81 with greater affinity than the extracellular secreted forms. They suggested that only the correctly folded and processed form of E2 bound CD81. This could explain the observed decrease in binding of H77 E2 to CD81 compared to GNA (Figures 6.2 and 6.3). Both aggregated and native forms of E2 would be expected to bind to GNA but if only the native form binds CD81, there would be a relative fall in observed binding. This theory could be tested with the use of conformation-dependent antibodies that should recognise only native forms of E2. Therefore these antibodies should detect equivalent amounts of E2 binding to both GNA and CD81. Figure 6.5 indicates that this was true for H60 but H53 and H50 both detected more binding to GNA than CD81. This could be explained if H53 and H50 were able to

inhibit CD81 binding but for H53, it has been shown that this is not the case (Flint *et al.*, 1999a, Owsianka *et al.*, 2001). H50 tested positive in a NOB assay (Patel *et al.* 2000), so could conceivably inhibit the E2/CD81 interaction, which would explain the lack of reactivity against CD81-bound E2 in Figure 6.5. H53 actually recognises an equivalent amount of CD81-bound E2 to H60, so possibly the increased reactivity with GNA-bound E2 is due to recognition of an epitope that is also present on the aggregated form of E2.

This is not the first report of an E2 protein from an HCV isolate that does not bind to CD81. Patel *et al.* (2000) demonstrated that the truncated E2 protein of genotype 1a Glasgow strain did not bind CD81 and this was confirmed with the full-length E2 protein in Figures 6.2b and 6.3b. These authors also found that Gla E2 formed mostly high molecular weight aggregates and was not recognised by conformation-dependent antibodies. Chimeras of Gla and H77 E2 were constructed in an attempt to identify a region in H77 E2 that conferred CD81 binding on Gla E2 or recognition by conformation-dependent antibodies or both (Patel *et al.*, 2000). Three chimeras (all containing the C-terminal portion of H77 E2) showed recognition by conformation-dependent antibodies and improved formation of native E2 as analysed under non-reducing conditions. However, only one of these three chimeras was able to bind CD81. This chimera contained a large portion of H77 E2 (aa 406-660), indicating that there may be multiple regions in the Gla E2 protein which are responsible for the lack of CD81 binding. It also indicated that CD81-binding may not be exclusively related to the ability of E2 to fold correctly. Yagnik *et al.* (2000) compared the CD81 binding affinities of genotype 1a H77 E2 and genotype 1b N2 E2 and found that E2 of N2 strain showed reduced binding efficiency compared to H77 E2. Another genotype 1b E2 protein (amplified from patient serum) was shown to bind CD81 with no indication of reduced efficiency compared with genotype 1a E2 (also amplified from patient serum) (Hadlock *et al.*, 2000). As different isolates were used in these two studies, it is impossible to compare them directly but there is a suggestion that the strength of the CD81 interaction with E2 may be isolate specific.

In Figures 6.2b and 6.3b there is a suggestion of increased binding above non-specific levels when concentrated E2 extracts of R3 and Gla are used. This could point to very

low levels of CD81 binding with these isolates. To investigate this further, the ability of genotype 3 (R3) E2 to compete with genotype 1 (H77) E2 for CD81 binding was examined (Figure 6.6.). However there was no evidence that the presence of genotype 3 E2 had any effect on the binding capabilities of genotype 1 E2, suggesting that R3 E2 does not bind CD81.

The capture ELISA has been used extensively to demonstrate the binding of HCV E2 to human CD81. However, this technique has demonstrated the binding of E2 to CD81 only and not the reverse of CD81 binding to E2, which should be equally strong. Surprisingly, binding of CD81 to H77 E2 could not be shown (Figure 6.8). All the other required interactions in this ELISA were verified (Figure 6.9). It is possible that CD81 in solution and CD81 bound to the plastic wells, have different conformations and this may interfere with its interaction with E2. Similarly, E2 bound by ALP98 may have a different conformation from E2 in solution. Neither CD81 nor E2 are present in a soluble form *in vivo*, as the proposed interaction is presumed to occur between cell-bound CD81 and the HCV virion. In an attempt to replicate the *in vivo* situation, Pileri *et al.* (1998) and Hadlock *et al.* (2000) showed that genotype 1a HCV particles in infectious chimpanzee plasma bound to CD81-coated polystyrene beads, as determined by quantitative RT-PCR. I performed a similar assay to assess the ability of genotype 1 and genotype 3 virions in human sera to bind CD81-coated agarose beads (Figure 6.11). Two independent experiments showed that genotype 3 virus particles (from two separate sera) bound preferentially to CD81, whereas the genotype 1 virus particles showed only non-specific binding to the agarose beads. These results contradict those of the CD81 capture ELISA, however, time constraints meant that the procedure was not fully optimised and many more sera need to be analysed before these results can be interpreted with confidence. Only a minor proportion of the HCV input was captured by CD81 (for the genotype 3 sera) and this created difficulties for quantification by PCR as product was only detected at the end of the 40-cycle PCR program, where the system is inherently less accurate. This explains the large standard deviations for detection of virus bound to the agarose beads. The level of non-specific binding to GST-coated agarose beads was high and different blocking procedures need to be investigated to improve this. Another

complicating factor is that antibodies present in the sera may interfere with binding. Ideally, one would want to use sera from acutely infected patients when they are PCR positive but antibody negative but these sera are very rare, as patients are usually asymptomatic during this stage. In addition to repeating this experiment with more sera of both genotypes, it would be necessary to show that the observed binding is due to CD81 by demonstrating its inhibition by anti-CD81 antibody.

These conflicting reports of strong binding, reduced binding and no binding to CD81, depending on the isolate, support the view that CD81 is not the primary receptor for HCV. Clearly, the expression of CD81 on a wide range of cell types does not explain the liver tropism of HCV. CD81 possibly acts as a co-receptor but different virus strains may use related but different co-receptors as with HIV and the chemokine receptors (Feng *et al.*, 1996, Alkhatib *et al.*, 1996). Along the same vein, Flint *et al.* (1999b) demonstrated that a cell surface-expressed form of E2 was able to interact with soluble human CD81 but was unable to induce cell fusion of CD81-positive HEK cells, which would be necessary for infection. HCV particles in plasma have been shown to be associated with lipoproteins, particularly low-density lipoproteins (LDLs) (Prince *et al.*, 1996). This led to the suggestion that the LDL receptor may be a candidate for the HCV receptor and it has since been shown that expression of the LDL receptor confers HCV binding on COS-7 cells (Monazahian *et al.*, 1999). Furthermore, Agnello *et al.* (1999) demonstrated that HCV and the other Flaviviridae viruses, bovine viral diarrhoea virus (BVDV) and hepatitis G virus, are endocytosed via the LDL receptor. Endocytosis correlated with LDL receptor activity and was inhibited in the presence of anti-LDL receptor antibody. However, there was also evidence of small amounts of intracellular HCV in cells that could not be inhibited by anti-LDL receptor antibody. Therefore the LDL receptor may not be the only means of viral entry. It may be a general mechanism employed by viruses that associate with lipoproteins. The HCV genotype of the sera used in this study was not specified. Wünschmann *et al.* (2000) investigated the two proposed HCV receptors, CD81 and LDL receptor, in terms of binding both E2 and HCV. The specific binding of E2 (unspecified isolate) to soluble human CD81 and to MOLT-4 T cells was confirmed. HCV was also shown to bind to MOLT-4 cells but the binding was strongly dependent on

the level of LDL receptor expression and could be inhibited by increased concentrations of LDL, whereas binding of E2 was unaffected. E2 was shown not to interact with either LDL or LDL receptor, so association of HCV with LDL is not mediated by E2. These data suggest that binding of HCV to MOLT-4 cells is due to a LDL interaction with LDL receptor and not an E2 interaction with CD81. However, similarly to Agnello *et al.* (1999), HCV binding to LDL receptor could not be blocked completely by LDL, so the involvement of another cellular molecule in HCV attachment cannot be ruled out. Again, the HCV genotype of the sera used in this study was unspecified but this is not expected to have a bearing on HCV binding to the LDL receptor via LDL, if the interaction with LDL is non-specific. In conclusion, the role of CD81 in attachment of HCV to cells is far from clear but differences in binding efficiency between isolates do not support the role of CD81 as the primary HCV receptor.

Chapter 7

General Discussion

7.1. H77 and R3 as representatives of genotypes 1 and 3

This study involved the direct comparison of the HCV genotype 3 glycoproteins with those of genotype 1. For these purposes, R3 and H77 were used as representatives of genotype 3 and genotype 1, respectively. The question is how representative of these two genotypes are these two clones? The full-length H77 clone is widely accepted as being not only representative of genotype 1a but also of the HCV genome in general, as RNA transcripts from this clone are infectious when injected into the liver of a chimpanzee (Yanagi *et al.*, 1997, Kolykhalov *et al.*, 1997). This infectious clone has proved invaluable in furthering the understanding of HCV biology, especially in the absence of an *in vitro* propagation system. However, it has meant that many biological features of HCV have been defined in terms of H77 only. Another genotype 1a infectious clone has recently been described (Lanford *et al.*, 2001) and it will be interesting to see what differences there are between these two clones. The only non-genotype 1 infectious clone described so far is that of genotype 2a and interestingly, a chimeric genome consisting of the genotype 2a structural genes in the context of a genotype 1a (H77) backbone, was found not to be infectious (Yanagi *et al.*, 1999a). During the construction of the infectious clones, it had been shown that minor deviations from the consensus amino acid sequence rendered the cDNA clones non-infectious (Yanagi *et al.*, 1997). In the absence of a genotype 3 infectious clone, or the facilities to produce one, the optimum method for obtaining a clone that would encode fully functional structural proteins, was to construct a

consensus sequence for this region. In this regard, every effort was made to ensure that the amino acid sequence of the R3 clone matched that of the PCR majority sequence (see section 3.5). Comparison with four published genotype 3 sequences confirmed that the R3 clone contained a genotype 3 sequence. It is more difficult to determine how representative this clone is of "typical" genotype 3 structural proteins as all isolates have amino acid differences. Those amino acids which are unique to R3 may translate to a structural or functional difference when compared to the published protein sequences. There were 13 such positions in E2 where R3 differed from the consensus genotype 3 sequence but the three positions that involve a proline residue would be expected to be most important in terms of protein structure. To determine if these residues do translate into a structural difference, they would have to be mutated to the genotype 3 consensus residue at that position. Work is currently underway to construct a second genotype 3 clone with which to compare the characteristics of the R3 proteins. This should help to distinguish between features which are truly characteristic of genotype 3 and those which are characteristic of R3.

7.2. Generalised differences between genotypes 1 and 3

In terms of disease progression there is no conclusive evidence to suggest that genotypes 1 and 3 differ in any respects. However steatosis, a histopathological feature of chronic hepatitis C infection characterised by lipid accumulation in hepatocytes, has been shown to be more frequent in genotype 3 infections (Rubbia-Brandt *et al.*, 2001). The core protein has been implicated in the development of steatosis due to its attachment to lipid droplets (Barba *et al.*, 1997). However, the exact mechanism by which lipids accumulate in the liver and the differences between the highly conserved genotypes 1 and 3 core which may affect this process remain to be determined.

The most established difference between genotypes 1 and 3 is the better response of patients infected with HCV genotype 3 to treatment with interferon-alpha (IFNα), either alone or in combination with ribavirin (McHutchinson *et al.*, 1998, Poynard *et al.*, 1998).

Much attention has been focussed on the NS5A protein and its potential role in IFNα resistance via inhibition of PKR. Most studies have concentrated on genotype 1 NS5A only. When provided in trans, genotype 1 NS5A has been shown to confer partial resistance to IFN, allowing rescue of IFN-sensitive viruses from the antiviral effects of IFN (Paterson *et al.*, 1999, Polyak *et al.*, 1999). However, there have been no reports on the effect of genotype 3 (or 2) NS5A expression on these IFN-sensitive viruses. If they too confer resistance to IFN, it may suggest that NS5A is responsible for mediating a general level of IFN resistance for HCV and additional pathways possibly exist for conferring an additional level of IFN resistance on genotype 1 HCV. This is not surprising as the IFN response involves a multitude of effects which are mediated by the protein products of IFN-induced genes. Therefore a virus has to contend with this multi-pronged attack and often evolves more than one mechanism to counteract the effects of IFN. For example, at least two IFN-resistance mechanisms have been identified so far in influenza virus, vaccinia virus, herpes simplex virus and HIV (Goodbourn *et al.*, 2000, Gale & Katze, 1998). The identification of the PePHD region in the E2 protein, whose sequence differed between genotypes 1 and 2/3 was proposed as a predictor of responsiveness to IFNα treatment (Taylor *et al.*, 1999). However this study (section 5.1.2.) and others showed that the PePHD sequence was highly conserved within genotypes and when mutations did occur they were not predictive of treatment outcome (Polyak *et al.*, 2000, Sarrazin *et al.*, 2000a). The mechanism by which genotype 1 E2 mediates resistance to IFN was shown to be via interaction with and inhibition of PKR, which was dependent on the PePHD (Taylor *et al.*, 1999). However, I could not confirm the interaction of genotype 1 (H77) E2 with PKR or observe co-localisation of these two proteins in the cell (sections 5.1.3. and 5.1.4.). Limitations on antibody choice, meant that the ability of genotype 3 E2 to interact with PKR could not be analysed but this aspect should be investigated further. Therefore the role of PKR and the PePHD in E2 in mediating IFN resistance remains under question. Examination of the levels of phosphorylated eIF2α in the presence and absence of both genotype 1 and 3 E2 would verify whether E2 expression could affect protein translation. This could be interpreted as a downstream effect of PKR or other mammalian eIF2α kinases. There have also been no reports on the ability of E2 (of either genotype 1 or 3) to confer resistance to IFN-

sensitive viruses, which would give an indication of whether E2 is involved in mediating HCV resistance to the antiviral effect of IFN.

7.3. Differences between the H77 and R3 glycoproteins

7.3.1. Antibody recognition

Monoclonal antibody reactivity provides structural information on a protein. Interestingly, there were several stretches of conserved amino acids in E2 between H77 and R3 and yet, no antibodies were specific for these regions. The only genotype 1 antibodies that were reactive against both H77 and R3 recognised overlapping epitopes over a highly conserved 16 amino acid region. As this region is conserved between two major genotypes, it may indicate that it is structurally or functionally important. This may be because it contains two conserved cysteine residues, which are involved in disulphide bonding as well a conserved glycosylation site. As yet, there have been no reports of an interaction with another protein via this region. It is clear there are few anti-E2 antibodies that are cross-reactive between genotypes and even between subtypes there is often no observed cross-reactivity (Patel *et al.*, 2000, Allander *et al.*, 2000b). This will be a major obstacle in the design of an HCV vaccine. For future work with genotype 3 E2 it would be beneficial to generate genotype 3 specific E2 antibodies, as these would give more insight into the structural features of genotype 3 E2 and how it differs from genotype 1 E2. The generation of genotype 3 conformation-dependent antibodies would be especially useful, as none of the genotype 1 conformation-dependent E2 antibodies tested in this study were able to recognise R3 E2. These antibodies have been important for distinguishing between the aggregate and native E1E2 complexes of genotype 1 and for determining which of these E2 forms are involved in interactions with other proteins (e.g. CD81) (Cocquerel *et al.*, 1998, Flint *et al.*, 2000).

7.3.2. Protein size and glycosylation states

Both the E1 and E2 proteins of R3 and H77 showed different mobilities when separated by SDS-PAGE. For E2, this could be explained by differing glycosylation states, with H77 E2 having one extra glycosylation site compared to R3 E2. The 10 predicted glycosylation sites in R3 E2 are conserved in the four published genotype 3 sequences and in H77 E2, so are presumably important for stability or biological activity. Extensive glycosylation is characteristic of viral glycoproteins and has been suggested to serve as a protection mechanism by covering potential immunogenic epitopes (Braakman & van Anken, 2000). Therefore, is the extra glycosylation site in H77 E2 significant? Does it confer extra protection from the immune system or does it allow for interactions with other proteins that are not seen with genotype 3 E2? If glycosylation status was identified as modulating a certain biological activity of E2, mutants of genotype 3 E2 could be generated to see if the presence of an 11^{th} glycan affects this activity.

Five predicted glycosylation sites are conserved between H77, R3 and the four published genotype 3 sequences. However, it has been shown that the 5^{th} site in genotype 1 E1 is not used (Meunier *et al.*, 1999) and size comparison of H77 E1 and R3 E1 indicated that R3 E1 has an additional glycan. The 5^{th} site in genotype 1 E1 is said not to be used due to the presence of a proline residue immediately after the tripeptide recognition sequence (Meunier *et al.*, 1999) but this sequence is present in R3 as well, so why is this site glycosylated in R3 but not in H77? Possibly there is another site in H77 E1 that is not used or conversely, an additional glycosylation site in R3 E1 that would account for the differences in protein mobility. A difference was also observed between the mobility of the H77 E1 proteins when analysed under reducing and non-reducing conditions and this was not seen with R3 E1. This suggests either a difference in structure or modification between the H77 E1 protein involved in the native complex and that involved in the aggregate. This characteristic may not be attributable to all genotype 1 E1 proteins, as it was not observed for the genotype 1a Glasgow isolate. Likewise, further analysis of other genotype 3 isolates is required to determine if this feature of R3 is true of all genotype 3 E1 proteins. Mutagenesis of the predicted glycosylation sites on H77 E1 and R3 E1

would help to determine the true glycosylation status of the E1 proteins. A further anomaly was observed when the unglycosylated forms of E1 were compared in that H77 E1 appeared to be smaller than R3 E1. This is not predicted from the sequence and the only explanation is that H77 E1 is folded into a more compact structure than R3 E1. There are obvious differences between H77 E1 and R3 E1 but these features should be analysed in other isolates before they are recognised as genuine differences between genotype 1 and genotype 3.

7.3.3. CD81 interaction

One of the most interesting differences observed between H77 and R3 was the lack of binding of R3 E2 to human CD81 compared to H77 E2. Apart from the reported binding of genotype 2 E2 (Hadlock *et al.*, 2000), all other previous documentation of E2/CD81 binding has involved genotype 1 E2. A possible site of direct interaction with CD81 has been identified on E2, which spans the epitope recognised by the anti-E2 antibody, AP33 (Owsianka *et al.*, 2001). This antibody did not show any reactivity against R3 E2, which is probably due to a single amino acid difference in the AP33 epitope between the E2 proteins of R3 and H77. Does this amino acid difference also result in inhibition of a CD81 interaction with R3 E2? To answer this, the amino acid difference in R3 will be mutated to match that of H77 to see if this confers CD81-binding on R3 E2. The major question is whether the lack of CD81 binding is characteristic of all genotype 3 isolates or just R3? Sequence analysis of the other four published genotype 3 sequences showed that one other isolate contains the same amino acid difference in the AP33 epitope as R3 but the other 3 match the H77 sequence at that position. It would be interesting to know if the E2 protein of these isolates shows binding to CD81. The region on E2 involved in the interaction with CD81 is believed to be discontinuous, so the amino acid sequence of other regions, in addition to the AP33 epitope, are likely to influence CD81 binding of E2. Research is underway to delineate the precise location of the residues on E2 involved in the CD81 interaction. A large-scale comparison of the CD81-binding ability of E2 proteins from a number of isolates of different genotypes, combined with sequence

comparison, will provide a better picture of whether CD81-binding is genotype-specific or isolate-specific.

In contrast to the work with recombinant E2, assessment of the ability of virions in human sera to bind CD81-coated agarose beads, suggested that genotype 3 virions did bind CD81 but genotype 1 virions did not. Time limitations meant that the technique was not fully optimised and not enough sera were analysed to confirm this correlation, so this is a preliminary observation. What is the significance of these observations? Human CD81 has been proposed as a receptor for HCV but it is becoming increasingly clear that additional factors are required for endocytosis, with the LDL-receptor being proposed as another HCV receptor (Agnello *et al.*, 1999). There may be a number of cell-surface factors involved in binding HCV and inducing endocytosis. One of these factors may be specific for hepatocytes (and possibly PBMCs) and would explain the cellular tropism of HCV, which is not explained by CD81 expression. It is also possible that rather than having specificity for CD81, HCV can interact with other members of the tetraspanin family as well and perhaps some HCV isolates show greater affinity for these other members. A report at a recent HCV meeting described high levels of expression of tetraspanin molecules, CD9, CD81, CD82 and CD151 on the surface of primary human hepatocytes (Charrin *et al.*, 2001) but it is not yet known if HCV can bind specifically to any of these alternative molecules.

References

Abid, K., Quadri, R. & Negro, F. (2000). Hepatitis C Virus, the E2 Envelope Protein, and α-Interferon Resistance. *Science* **287**, 1555.

Aggarwal, R. & Krawczynski, K. (2000). Hepatitis E: an overview and recent advances in clinical and laboratory research. *Journal of Gastroenterology and Hepatology* **15**, 9-20.

Agnello, V., Abel, G., Elfahal, M., Knight, G. B. & Zhang, Q. X. (1999). Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proceedings of the National Academy of Sciences USA* **96**, 12766-12771.

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.

Ali, N. & Siddiqui, A. (1995). Interaction of polypyrimidine tract-binding protein with the 5'non-coding region of the hepatitis C viral RNA genome and its functional requirement in internal initiation of translation. *Journal of Virology* **69**, 6367-6375.

Ali, N. & Siddiqui, A. (1997). The La antigen binds 5' non-coding region of the hepatitis C virus RNA in the context of the initiator AUG codon and stimulates internal ribosome entry site-mediated translation. *Proceedings of the National Academy of Sciences USA* **18**, 2249-2254.

Alkhatib, G., Combadiere, C., Broder, C. C., Feng, Y., Kennedy, P. E., Murphy, P. M. & Berger, E. A. (1996). CC CKR5: A RANTES, MIP-1a, MIP-1b receptor as a fusion co-factor for macrophage-tropic HIV-1. *Science* **272**, 1955-1958.

Allander, T., Forns, X., Emerson, S. U., Purcell, R. H. & Bukh, J. (2000a). Hepatitis C virus envelope protein E2 binds to CD81 of tamarins. *Virology* **277**, 358-367.

Allander, T., Drakenberg, K., Beyene, A., Rosa, D., Abrignani, S., Houghton, M., Widell, A., Grillner, L. & Persson, M. A. (2000b). Recombinant human monoclonal antibodies against different conformational epitopes of the E2 envelope glycoprotein of hepatitis C virus that inhibit its interaction with CD81. *Journal of General Virology* **81**, 2451-2459.

Alter, H. J., Purcell, R. H., Holland, P. V. & Popper, H. (1978). Transmissible agent in non-A, non-B hepatitis. *Lancet* **1**, 459-463.

Alter, M. J., Margolis, H. S., Krawczynski, K., Judson, F. N., Mares, A., Alexander, W. J., Hu, P. Y., Miller, J. K., Gerber, M. A. & Sampliner, R. E. (1992). The natural history of community-acquired hepatitis C in the United States. *New England Journal of Medicine* **327**, 1899-1905.

Alter, M. J. (1999). Hepatitis C virus infection in the United States. *Journal of Hepatology* **31**, 88-91.

Altmeyer, R., Mordelet, E., Girard, M. & Vidal, C. (1999). Expression and detection of macrophage-tropic HIV-1 gp120 in the brain using conformation-dependent antibodies. *Virology* **259**, 314-323.

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.

Banjeree, R., Echeverri, A. & Dasgupta, A. (1997). Polio virus-encoded 2C polypeptide specifically binds to the 3'-terminal sequences of viral negative-strand RNA. *Journal of Virology* **71**, 9570-9578.

Banerjee, R. & Dasgupta, A. (2001). Specific Interaction of Hepatitis C Virus Protease/Helicase NS3 with the 3'-Terminal Sequences of Viral Positive- and Negative-Strand RNA. *Journal of Virology* **75**, 1708-2129.

Barba, G., Harper, F., Harada, T., Kohara, M., Goulinet, S., Matsuura, Y., 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., Ahlborn, L. 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., Ahlborn, L. L., Mous, J. & Jacobsen, H. (1994). Kinetic and structural analyses of hepatitis C virus polyprotein processing. *Journal of Virology* **68**, 5045-5055.

Bartenschlager, R., Ahlborn-Laake, L., Yasargil, K., Mous, J. & Jacobsen, H. (1995). Substrate determinants for cleavage in cis and in trans by the hepatitis C virus NS3 proteinase. *Journal of Virology* **69**, 198-205.

Bartenschlager, R. & Lohmann, V. (2000). Replication of hepatitis C virus. *Journal of General Virology* **81**, 1631-1648.

Bartenschlager, R. & Lohmann, V. (2001). Novel cell culture systems for the hepatitis C virus. *Antiviral Research* **52**, 1-17.

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.

Beard, M.R., Abell, G., Honda, M., Carroll, A., Gartland, M., Clarke, B., Suzuki, K., Lanford, R., Sangar, D.V. & Lemon, S.M. (1999). An infectious molecular clone of a Japanese genotype 1b hepatitis C virus. *Hepatology* **30**, 316-324.

Behrens, S.-E., Tomei, L. & De Francesco, R. (1996). Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *The EMBO journal* **15**, 12-22.

Benvegnu, L., Pontisso, P., Cavalletto, D., Noventa, F., Chemello, L. & Alberti, A. (1997). Lack of correlation between hepatitis C virus genotypes and clinical course of hepatitis C virus-related cirrhosis. *Hepatology* **25**, 211-215.

Berg, J. M. & Shi, Y. (1996). The galvanization of biology: a growing appreciation for the roles of zinc. *Science* **271**, 1081-1085.

Berg, T., Mas Marques, A., Hohne, M., Wiedenmann, B., Hopf, U. & Schreier, E. (2000). Mutations in the E2-PePHD and NS5A region of hepatitis C virus type 1 and the dynamics of hepatitis C viremia decline during interferon alfa treatment. *Hepatology* **32**, 1386-1395.

Berglund, P., Fleeton, M. N., Smerdou, C. & Liljestrom, P. (1999). Immunization with recombinant Semliki Forest virus induces protection against influenza challenge in mice. *Vaccine* **17**, 497-507.

Bertolotti, A., Zhang, Y., Hendershot, L. M., Harding, H. P. & Ron, D. (2000). Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nature Cell Biology* **2**, 326-332.

Besse, S., Rebouillat, D., Marie, I., Puvion-Dutilleul, F. & Hovanessian, A. G. (1998). Ultrastructural localization of interferon-inducible double-stranded RNA-activated enzymes in human cells. *Experimental Cell Research* **239**, 379-392.

Blatt, L. M., Mutchnick, M. G., Tong, M. J., Klion, F. M., Lebovics, E., Freilich, B., Bach, N., Smith, C., Herrera, J., Tobias, H., Conrad, A., Schmid, P. & McHutchinson, J. G. (2000). Assessment of hepatitis C virus RNA and genotype from 6807 patients with chronic hepatitis C in the United States. *Journal of Viral Hepatitis* **7**, 196-202.

Blight, K. J., Kolykhalov, A. A. & Rice, R. C. (2000). Efficient Initiation of HCV RNA Replication in Cell Culture. *Science* **290**, 1972-1974.

Booth, J. C. L., O'Grady, J. & Neuberger, J. (2001). Clinical guidelines on the management of hepatitis C. *Gut* **49**, i1-i21.

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.

Bowden, S. (2001). New hepatitis viruses: contenders and pretenders. *Journal of Gastroenterology and Hepatology* **16**, 124-131.

Braakman, I. & Van Anken, E. (2000). Folding of viral envelope glycoproteins in the endoplasmic reticulum. *Traffic* **1**, 533-539.

Bradley, D. W., McCaustland, K. A., Cook, E. H., Schable, C. A., Ebert, J. W. & Maynard, J. W. (1985). Posttransfusion non-A, non-B hepatitis in chimpanzees: physiochemical evidence that the tubule-forming agent is a small, enveloped virus. *Gastroenterology* **88**, 773-779.

Brechot, C. (1997). Hepatitis C virus 1b, cirrhosis, and hepatocellular carcinoma. *Hepatology* **25**, 772-774.

Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., De Francesco, R. & Rey, F. A. (1999). Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proceedings of the National Academy of Sciences USA* **96**, 13034-13039.

Brown, E. A., Zhang, H., Ping, L.-H. & Lemon, S. (1992). Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Research* **20**, 5041-5045.

Bruno, S., Silini, S., Crosignani, A., Borzio, F., Leandro, G., Bono, F. & Asti, M. (1997). Hepatitis C virus genotypes and risk of hepatocellular carcinoma in cirrhosis: a prospective study. *Hepatology* **25**, 754-758.

Bukh, J., Purcell, R. H. & Miller, R. H. (1992). Sequence analysis of the 5' noncoding region of hepatitis C virus. *Proceedings of the National Academy of Sciences USA* **89**, 4942-4946.

Bukh, J., Purcell, R. H. & Miller, R. H. (1993). At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. *Proceedings of the National Academy of Sciences USA* **90**, 8234-8238.

Bukh, J., Purcell, R. H. & Miller, R. H. (1994). Sequence analysis of the core gene of 14 hepatitis C virus genotypes. *Proceedings of the National Academy of Sciences USA* **91**, 8239-8243.

Cardoso, M. D. S., Siemoneit, K., Sturm, 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.

Cha, T.-A., Beall, E., Irvine, B., Kolberg, J., Chien, D., Kuo, G. & Urdea, M. S. (1992). At least five related, but distinct hepatitis C viral genotypes exist. *Proceedings of the National Academy of Sciences USA* **89**, 7144-7148.

Chan, S.-W., McOmish, F., Holmes, E. C., Dow, B., Peutherer, J. F., Follett, E., Yap, P. L. & Simmonds, P. (1992). Analysis of a new hepatitis C virus type and its phylogenetic relationship to existing variants. *Journal of General Virology* **73**, 1131-1141.

Chang, H. W., Watson, J. L. & Jacobs, B. L. (1992). The E3L gene of vaccinia virus encodes an inhibitor of PKR. *Proceedings of the National Academy of Sciences USA* **89**, 4825-4829.

Charrin, S., Rubinstein, E., Boucheix, C. & Petit, M.-A. (2001). CD81-associated molecules at the surface of human hepatocytes: hepatoma cell lines (HepG2 and Huh7) and primary cultures. In *8th International Symposium on Hepatitis C Virus & Related Viruses*, pp. 200. Paris.

Chayama, K., Suzuki, F., Tsubota, A., Kobayashi, M., Arase, Y., Saitoh, S., Suzuki, Y., Murashima, N., Ikeda, K., Takahashi, N., Kinoshita, M. & Kumada, H. (2000). Association of amino acid sequence in the PKR-eIF2 Phosphorylation Homolgy Domain and Response to Interferon Therapy. *Hepatology* **32**, 1138-1144.

Chen, C., You, L., Hwang, L. & Lee, Y. (1997). Direct interaction of hepatitis C virus core protein with the cellular lymphotoxin-beta receptor modulates the signal pathway of the lymphotoxin-beta receptor. *Journal of Virology* **71**, 9417-9426.

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., 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., Medina-Selby, 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.

Choo, Q. L., Kuo, G., Ralston, R., Weiner, A., Chien, D., Van Nest, 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.

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.

Choukhi, A., Pillez, A., Drobecq, H., Sergheraert, C., Wychowski, C. & Dubuisson, J. (1999). Characterization of aggregates of hepatitis C virus glycoproteins. *Journal of General Virology* **80**, 3099-3107.

Ciccaron, V., Jessee, J., Berglund, P. & Liljestrom, P. (1993). pSFV1 eukaryotic expression vector: a novel protein expression system. *Focus* **15**, 103-105.

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.

Cocquerel, L., Wychowski, C., Minner, F., Penin, F. & Dubuisson, J. (2000). Charged residues in the transmembrane domains of Hepatitis C virus glycoproteins play a major role in the processing, subcellular localization, and assembly of these envelope proteins. *Journal of Virology* **74**, 3623-3633.

Cocquerel, L., Meunier, J.-C., Op de Beeck, A., Bonte, D., Wychowski, C. & Dubuisson, J. (2001). Coexpression of hepatitis C virus envelope protein E1 and E2 in cis improves the stability of membrane insertion of E2. *Journal of General Virology* **82**, 1629-1635.

Colombo, M. (1999). Natural history and pathogenesis of hepatitis C virus related hepatocellular carcinoma. *Journal of Hepatology* **31**, 25-30.

Cox, J. S. & Walter, P. (1996). A novel mechanism for regulating activity of a transcriptional factor that controls the unfolded protein response. *Cell* **87**, 391-404.

Cuthbert, J. A. (2001). Hepatitis A: old and new. *Clinical Microbiology Reviews* **14**, 38-58.

Daemen, T., Pries, F., Bungener, L., Kraak, M., Regts, J. & Wilschut, J. (2000). Genetic immunization against cervical carcinoma: induction of cytotoxic T lymphocyte activity with a recombinant alphavirus vector expressing human papillomavirus type 16 E6 and E7. *Gene Therapy* **7**, 1859-1866.

Dammacco, F., Gatti, P. & Sanaonno, D. (1998). Hepatitis C virus infection, mixed cryoglobulinemia and non-Hodgkin's lymphoma: an emerging picture. *Leukemia and Lymphoma* **31**, 463-476.

Davis, G. L. & Lau, J. Y. (1997). Factors predictive of a beneficial response to therapy of hepatitis C. *Hepatology* **26**, 122S-127S.

Deleerleersnyder, 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.

Di Bisceglie, A., Conjeevaram, H., Fried, M., Sallie, R., Park, Y., Yurdaydin, C., Swain, M., Kleiner, D. E., Mahaney, K. & Hoofnagle, J. H. (1995). Ribavirin as monotherapy for chronic hepatitis C: a randomised, double-blind, placebo-controlled trial. *Annals of Internal Medicine* **123**, 897-903.

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.

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. (2000). Folding, assembly and subcellular localization of HCV glycoproteins. *Current Topics in Microbiology and Immunology* **242**, 135-148.

Dubuisson, J., Duvet, S., Meunier, J.-C., Op de Beeck, A., Cacan, R., Wychowski, C. & Cocquerel, L. (2000). Glycosylation of the hepatitis C virus envelope protein E1 is dependent on the presence of a downstream sequence on the viral protein. *Journal of Biological Chemistry* **275**, 30605-30609.

Dusheiko, G., Schmilovitz-Weiss, H., Brown, D., McOmish, F., Yap, P.-L., Sherlock, S., McIntyre, N. & Simmonds, P. (1994). Hepatitis C virus genotypes: An investigation of type-specific differences in geographic origin and disease. *Hepatology* **19**, 13-18.

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.

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.

Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Izumi, N., Marumo, F. & Sato, C. (1995). Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. *Journal of Clinical Investigation* **96**, 224-230.

Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Ogura, Y., Izumi, N., Marumo, F. & Sato, C. (1996). Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *New England Journal of Medicine* **334**, 77-81.

Failla, C., Tomei, L. & De Francesco, R. (1994). Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. *Journal of Virology* **68**, 3753-3760.

Farci, P., Alter, H. J., Wong, D., Miller, R. H., Shih, J. W., Jett, B. & Purcell, R. H. (1991). A long term study of hepatitis C virus replication in non-A, non-B hepatitis. *New England Journal of Medicine* **325**, 98-104.

Fearon, D. T. & Carter, R. H. (1995). The CD19/CR2/TAPA-1 complex of B lymphocytes: linking natural to acquired immunity. *Annual Reviews in Immunology* **13**, 127-149.

Feinstone, S. M., Mihalik, K. B., Kamimura, T., Alter, H. J., London, W. T. & Purcell, R. H. (1983). Inactivation of hepatitis B virus and non-A, non-B hepatitis by chloroform. *Infection and Immunology* **41**, 816-821.

Feng, Y., Broder, C. C., Kennedy, P. E. & Berger, E. A. (1996). HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane G protein-coupled receptor. *Science* **272**, 872-877.

Flajolet, M., Rotondo, G., Daviet, L., Bergametti, F., Inchauspe, G., Tiollais, P., Transy, C. & Legrain, P. (2000). A genomic approach of the hepatitis C virus generates a protein interaction map. *Gene* **242**, 369-379.

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., Maidens, C., Loomis-Price, L. D., Shotton, C., Dubuisson, J., Monk, P., Higginbottom, A., Levy, S. & McKeating, J. A. (1999a). Characterization of Hepatitis C Virus E2 Glycoprotein Interaction with a Putative Cellular Receptor, CD81. *Journal of Virology* **73**, 6235-6244.

Flint, M., Thomas, J. M., Maidens, C. M., Shotton, C., Levy, S., Barclay, W. S. & McKeating, J. A. (1999b). Functional analysis of cell surface-expressed hepatitis C virus E2 glycoprotein. *Journal of Virology* **73**, 6782-6790.

Flint, M. & McKeating, J. A. (2000). The role of the hepatitis C virus glycoproteins in infection. *Reviews in Medical Virology* **10**, 101-117.

Flint, M., Dubuisson, J., Maidens, C., Harrop, R., Guile, G. R., Borrow, P. & McKeating, J. A. (2000). Functional characterization of intracellular and secreted forms of a truncated hepatitis C virus E2 glycoprotein. *Journal of Virology* **74**, 702-709.

Forns, X., Bukh, J., Purcell, R. H. & Emerson, S. U. (1997). How *Escherichia coli* can bias the results of molecular cloning: preferential selection of defective genomes of hepatitis C virus during the cloning procedure. *Proceedings of the National Academy of Sciences USA* **94**, 13909-13914.

Forns, X., Emerson, S. U., Tobin, G. J., Mushahwar, I. K., Purcell, R. H. & Bukh, J. (1999). DNA immunization of mice and macaques with plasmids encoding hepatitis C virus envelope E2 protein expressed intracellularly and on the cell surface. *Vaccine* **17**, 1992-2002.

Forns, X., Thimme, R., Govindarajan, S., Emerson, S. U., Purcell, R. H., Chisari, F. V. & Bukh, J. (2000). Hepatitis C virus lacking the hypervariable region 1 of the second envelope protein is infectious and causes acute resolving or persistent infection in chimpanzees. *Proceedings of the National Academy of Sciences USA* **97**, 13318-13323.

Fournillier, A., Depla, E., Karayiannis, P., Vidalin, O., Maertens, G., Trepo, C. & Inchauspe, G. (1999). Expression of noncovalent hepatitis C virus envelope E1-E2 complexes is not required for the induction of antibodies with neutralizing properties following DNA immunization. *Journal of Virology* **73**, 7497-7504.

Francois, C., Duverlie, G., Rebouillat, D., Khorsi, H., Castelain, S., Blum, H. E., Gatignol, A., Wychowski, C., Moradpour, D. & Meurs, E. F. (2000). Expression of Hepatitis C Virus Proteins Interferes with the Antiviral Action of Interferon Independently of PKR-Mediated Control of Protein Synthesis. *Journal of Virology* **74**, 5587-5596.

Frank, C., Mohamed, M. K., Strickland, G. T., Lavanchy, D., Arthur, R. R., Magder, L. S., Khoby, T. E., Abdel-Wahab, Y., Ohn, E. S. A., Anwar, W. & Sallam, I. (2000). The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. *Lancet* **355**, 887-891.

Frese, M., Kochs, G., Feldmann, H., Hertkorn, C. & Haller, O. (1996). Inhibition of bunyaviruses, phleboviruses and hantaviruses by human MxA protein. *Journal of Virology* **69**, 915-923.

Frese, M., Pietschmann, T., Moradpour, D., Haller, O. & Bartenschlager, R. (2001). Interferon-alpha inhibits hepatitis C virus subgenomic RNA replication by an MxA-independent pathway. *Journal of General Virology* **82**, 723-733.

Gale, M. Jr., 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 nonstructural 5A protein. *Virology* **230**, 217-227.

Gale, M., Jr & Katze, M. G. (1998). Molecular mechanisms of interferon resistance mediated by viral-directed inhibition of PKR, the interferon-induced protein kinase. *Pharmacology and Therapy* **78**, 29-46.

Gale, M., Jr., Blakely, C. M., Kwleciszewski, B., Tan, S. L., Dossett, M., Tang, N. M., Korth, M. J., Polyak, S. J., Gretch, D. R. & Katze, M. G. (1998). Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: Molecular mechanisms of kinase regulation. *Molecular & Cellular Biology* **18**, 5208-5218.

Gale, M., Jr., 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.

Gallinari, P., Brennan, D., Nardi, C., Brunetti, M., Tomei, L., Steinkuhler, C. & De Francesco, R. (1998). Multiple enzymatic activities associated with recombinant NS3 protein of hepatitis C virus. *Journal of Virology* **72**, 6758-6769.

Gavel, Y. & von Heijne, G. (1990). Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Engineering* **5**, 433-442.

Gething, M.-J. & Sambrook, J. (1992). Protein folding in the cell. *Nature* **355**, 33-45.

Gontarek, R. R., Gutshall, L. L., Herold, K. M., Tsai, J., Sathe, G. M., Mao, J., Prescott, C. & Del Vecchio, A. M. (1999). hnRNP C and polypyrimidine tract-binding protein specifically interact with the pyrimidine-rich region within the 3'NTR of the HCV RNA genome. *Nucleic Acids Research* **27**, 1457-1463.

Goodbourn, S., Didcock, L. & Randall, R. E. (2000). Interferons: cell signalling, immune modulation, antiviral response and virus countermeasures. *Journal of General Virology* **81**, 2341-2364.

Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M. & Rice, C. M. (1993a). Expression and identification of hepatitis C virus polyprotein cleavage products. *Journal of Virology* **67**, 1385-1395.

Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M. & Rice, C. M. (1993b). Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polyprotein cleavage sites. *Journal of Virology* **67**, 2832-2843.

Gumber, S. C. & Chopra, S. (1995). Hepatitis C: a multifaceted disease. Review of extrahepatic manifestations. *Annals of Internal Medicine* **123**, 615-620.

Guo, J.-T., Bichko, V. V. & Seeger, C. (2001a). Effect of alpha interferon on the hepatitis C virus replicon. *Journal of Virology* **75**, 8516-8523.

Guo, J.-T., Zhu, Q. & Seeger, C. (2001b). How does interferon (IFN) alpha inhibit the replication of HCV subgenomes in Huh7 cells? In *8th International Symposium on Hepatitis C Virus & Related Viruses*, pp. 83. Paris.

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.

Hadlock, K. G., Lanford, R. E., Perkins, S., Rowe, J., Yang, Q., Levy, S., Pileri, P., Abrignani, S. & Foung, S. K. H. (2000). Human Monoclonal Antibodies That Inhibit Binding of Hepatitis C Virus E2 Protein to CD81 and Recognize Conserved Conformational Epitopes. *Journal of Virology* **74**, 10407-10416.

Hammond, C. & Helenius, A. (1994). Folding of VSV G protein: sequential interaction with BiP and calnexin. *Science* **266**, 456-458.

Harada, T., Tautz, N. & Thiel, H.-J. (2000). E2-p7 region of the bovine viral diarrhea virus polyprotien: processing and functional studies. *Journal of Virology* **74**, 9498-9506.

Harding, H. P., Zhang, Y. & Ron, D. (1999). Protein translation and folding are coupled by an endoplasmic- reticulum-resident kinase. *Nature* **397**, 271-274.

Harding, H. P., Zhang, Y., Bertolotti, A., Zeng, H. & Ron, D. (2000). Perk is essential for translational regulation and cell survival during the unfolded protein response. *Molecular Cell* **5**, 897-904.

He, L.-F., Alling, D., Popkin, T., Shapiro, M., Alter, H. J. & Purcell, R. H. (1987). Determining the size of non-A, non-B hepatitis virus by filtration. *Journal of Infectious Diseases* **156**, 636.

Heintges, T. & Wands, J. R. (1997). Hepatitis C virus: epidemiology and transmission. *Hepatology* **26**, 521-526.

Heinz, F. X. (1992). Comparative molecular biology of flaviviruses and hepatitis C virus. *Archives of Virology Supplement* **4**, 163-171.

Helenius, A., Trombetta, E. S., Herbert, 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.

Herbert, D. N., Foellmer, B. & Helenius, A. (1995). Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. *Cell* **81**, 425-433.

Herbert, 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.

Higginbottom, A., Quinn, E. R., Kuo, C. C., Flint, M., Wilson, L. H., Bianchi, E., Nicosia, A., Monk, P. N., McKeating, J. A. & Levy, S. (2000). Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *Journal of Virology* **74**, 3642-3649.

Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M. & Shimotohno, K. (1991a). Gene mapping of the putative structural region of the hepatitis C virus genome by in vitro processing analysis. *Proceedings of the National Academy of Sciences USA* **88**, 5547-5551.

Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M., Ohshoki, S. & Shimotohno, K. (1991b). Hypervariable regions in the putative glycoprotein of hepatitis C virus. *Biochemical and Biophysical Research Communications* **175**, 220-228.

Hijikata, M., Shimizu, Y. K., Kato, H., Inamoto, A., Shih, J.W., Alter, H.J., Purcell, R.H. & Yoshikura, H. (1993a). Equilibrium centrifugation studies of hepatitis C virus: evidence for circulating immune complexes. *Journal of Virology* **67**, 1953-1958.

Hijikata, M., Mizushima, H., Akagi, T., Mori, S., Kakiuchi, N., Kato, N., Tanaka, T., Kimura, K. & Shimotohno, K. (1993b). Two distinct proteinase activities required for the processing of a putative nonstructural 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. (1993c). Proteolytic processing and membrane association of putative nonstructural proteins of hepatitis C virus. *Proceedings of the National Academy of Sciences USA* **90**, 10773-10777.

Hofgartner, W. T., Polyak, S. J., Sullivan, D. G., Carithers, R. L., Jr. & Gretch, D. R. (1997). Mutations in the NS5A gene of hepatitis C virus in North American patients infected with HCV genotype 1a or 1b. *Journal of Medical Virology* **53**, 118-126.

Honda, M., Beard, M. R., Ping, L.-H. & Lemon, S. M. (1999). A phylogenetically conserved stem-loop structure at the 5' border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation. *Journal of Virology* **73**, 1165-1174.

Hoofnagle, J. H., Mullen, K. D., Jones, D. B., Rustgi, V., Di Bisceglie, A., Peters, M., Waggoner, J. G., Park, Y. & Jones, E. A. (1986). Treatment of chronic non-A, non-B hepatitis with recombinant human alpha interferon. *New England Journal of Medicine* **315**, 1575-1578.

Hoofnagle, J. H. (1997). Hepatitis C: the clinical spectrum of disease. *Hepatology* **26**, 15S-20S.

Hope, R. G. & McLauchlan, J. (2000). Sequence motifs required for lipid droplet association and protein stability are unique to the hepatitis C virus core protein. *Journal of General Virology* **81**, 1913-1925.

Houghton, M. (2000). Strategies and prospects for vaccination against the hepatitis C viruses. *Current Topics in Microbiology and Immunology* **242**, 327-339.

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.

Hugle, T., Fehrmann, F., Bieck, E., Kohara, M., Krausslich, H.-G., Rice, C.M., Blum, H.E. & Moradpour, D. (2001). The hepatitis C virus nonstructural protein 4B is an integral endoplasmic reticulum membrane protein. *Virology* **284**, 70-81.

Hussy, P., Langen, H., Mous, J. & Jacobsen, H. (1996). Hepatitis C virus core protein: carboxy-terminal boundaries of two processed species suggest cleavage by a signal peptide peptidase. *Virology* **224**, 93-104.

Hwang, S. B., Park, K.-J., Kim, Y.-S., Sung, Y. C. & Lai, M. M. C. (1997). Hepatitis C virus NS5B protein is a membrane-associated phosphoprotein with a predominantly perinuclear localization. *Virology* **227**, 439-446.

Ide, Y., Zhang, L., Chen, M., Inchauspe, G., Bahl, C., Sasaguri, Y. & Padmanabhan, R. (1996). Characterization of the nuclear localization signal and subcellular distribution of hepatitis C virus nonstructural protein NS5A. *Gene* **182**, 203-211.

Imai, T., Kakizaki, M., Nishimura, M. & Yoshie, O. (1995). Molecular analyses of the association of CD4 with two members of the transmembrane 4 superfamily, CD81 and CD82. *Journal of Immunology* **155**, 1229-1239.

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 of the United States of America* **88**, 10292-10296.

Inchauspe, G., Major, M. E., Nakano, I., Vitvitski, L. & Trepo, C. (1997). DNA vaccination for the induction of immune responses against hepatitis C virus proteins. *Vaccine* **15**, 853-856.

Ishak, K., Baptista, A., Bianchi, L., Callea, F., De Groote, J., Gudat, F., Denk, H., Desmet, V., Korb, G., MacSween, R. N. M., Phillips, M. J., Portmann, B. G., Poulsen, H., Scheuer, P. J., Schmid, M. & Thaler, H. (1995). Histological grading and staging of chronic hepatitis. *Journal of Hepatology* **22**, 696-699.

Ishido, S., Fujita, T. & Hotta, H. (1998). Complex formation of NS5B with NS3 and NS4A proteins of hepatitis C virus. *Biochemical and biophysical research communications* **244**, 35-40.

Ishii, K., Rosa, D., Watanabe, Y., Katayama, T., Harada, H., Wyatt, C., Kiyosawa, K., Aizaki, H., Matsuura, Y., Houghton, M., Abrignani, S. & Miyamura, T. (1998). High titers of antibodies inhibiting the binding of envelope to human cells correlate with natural resolution of chronic hepatitis C. *Hepatology* **28**, 1117-1120.

Ito, T., Mukaigawa, J., Zuo, J., Hirabayashi, Y., Mitamura, K. & Yasui, K. (1996). Cultivation of hepatitis C virus in primary hepatocyte culture from patients with chronic hepatitis C results in release of high titre infectious virus. *Journal of General Virology* **77**, 1043-1054.

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., Tahara, S. & Lai, M. (1998). The 3'-untranslated region of hepatitis C virus RNA enhances translation from an internal ribosomal entry site. *Journal of Virology* **72**, 8789-8796.

Iwawaki, T., Hosoda, A., Okuda, T., Kamigori, Y., Nomura-Furuwatari, C., Kimata, Y., Tsuru, A. & Kohno, K. (2001). Translational control by the ER transmembrane kinase/ribonuclease IRE1 under ER stress. *Nature Cell Biology* **3**, 158-164.

Jackson, R. J., Howell, M. T. & Kaminski, A. (1990). The novel mechanism of initiation of picornavirus RNA translation. *Trends in Biochemical Sciences* **15**, 477-483.

Jimenez-Garcia, L. F., Green, S. R., Mathews, M. B. & Spector, D. L. (1993). Organisation of the double-stranded RNA-activated protein kinase DAI and virus-associated VA RNA_I in adenovirus-2-infected HeLa cells. *Journal of Cell Science* **106**, 11-22.

Jouet, P., Roudot-Thoraval, F., Dhumeaux, D. & Metreau, J. M. (1994). Comparative efficacy of interferon alpha in cirrhotic and non-cirrhotic patients with non-A non-B hepatitis. *Gastroenterology* **106**, 686-690.

Kaba, S., Dutta, U., Byth, K., Crewe, E., Kahn, M., Coverdale, S., Lin, R., Liddle, C. & Farrell, G.C. (1998). Molecular epidemiology of hepatitis C in Australia. *Journal of Gastroenterology and Hepatology* **13**, 914-920.

Kaito, M., Watanabe, S., Tsukiyama-Kohara, 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.

Kanazawa, Y., Hayashi, N., Mita, E., Li, T., Hagiwara, H., Kasahara, A., Fusamoto, H. & Kamada, T. (1994). Influence of viral quasispecies on effectiveness of interferon therapy in chronic hepatitis C patients. *Hepatology* **20**, 1121-1130.

Kaneko, T., Tanji, Y., Satoh, S., Hijikata, M., Asabe, S., Kimura, K. & Shimotohni, K. (1994). Production of two phosphoproteins from the NS5A region of the hepatitis C viral genome. *Biochemical and Biophysical Research Communications* **205**, 320-326.

Kanistanon, D., Neelamek, M., Dharakul, T. & Songsivilai, S. (1997). Genotypic distribution of Hepatitis C virus in different regions of Thailand. *Journal of Clinical Microbiology* **35**, 1772-1776.

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., Lan, K. H., Ono-Nita, S. K., Shiratori, Y. & Omata, M. (1997). Hepatitis C virus nonstructural region 5A protein is a potent transcriptional activator. *Journal of Virology* **71**, 8856-8859.

Kaufman, R. J. (1999). Stress signalling from the lumen of the endoplasmic reticulum: co-ordination of gene transcriptional and translational controls. *Genes and Development* **13**, 1211-1233.

Kehm, E., Goksu, M. A. & Knopf, C. W. (1998). Expression analysis of recombinant herpes simplex virus type 1 DNase. *Virus Genes* **17**, 129-138.

Keskinen, P., Nyqvist, M., Sareneva, T., Pirhonen, J., Melen, K. & Julkunen, I. (1999). Impaired antiviral response in human hepatoma cells. *Virology* **263**, 364-375.

Khromykh, A.A., Sedlak, P.L. & Westaway, E.G. (2000). Cis-and trans-acting elements in flavivirus RNA replication. *Journal of Virology* **74**, 3253-3263.

Khorsi, H., Castelain, S., Wyseur, A., Izopet, J., Canva, V., Rombout, A., Capron, D., Capron, J. P., Lunel, F., Stuyver, L. & Duverlie, G. (1997). Mutations of hepatitis C virus 1b NS5A 2209-2248 amino acid sequence do not predict the response to recombinant interferon-alfa therapy in French patients. *Journal of Hepatology* **27**, 72-77.

Kim, D. W., Gwack, Y., Han, J. H. & Choe, J. (1995). C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. *Biochemical and biophysical research communications* **215**, 160-166.

Kim, J. L., Morgenstern, K. A., Lin, C., Fox, T., Dwyer, M. D., Landro, J. A., Chambers, S. P., Markland, W., Lepre, C. A., O'Malley, 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, 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.

Koch, J. O. & Bartenschlager, R. (1999). Modulation of hepatitis C virus NS5A hyperphosphorylation by nonstructural protein NS3, NS4A and NS4B. *Journal of Virology* **73**, 7138-7146.

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., Mihalik, K., Feinstone, S. M. & Rice, C. M. (2000). Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3' nontranslated region are essential for virus replication in vivo. *Journal of Virology* **74**, 2046-2051.

Kornfeld, R. & Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. *Annual Review of Biochemistry* **54**, 631-664.

Krieger, N., Lohmann, V. & Bartenschlager, R. (2001). Enhancement of Hepatitis C Virus RNA Replication by Cell Culture-Adaptive Mutations. *Journal of Virology* **75**, 4614-4624

Kruger, M., Beger, C., Li, Q. X., Welch, P. J., Tritz, R., Leavitt, M., Barber, J. R. & Wong-Staal, F. (2000). Identification of eIF2Bgamma and eIF2gamma as cofactors of hepatitis C virus internal ribosome entry site-mediated translation using a functional genomics approach. *Proceedings of the National Academy of Sciences USA* **97**, 8566-8571.

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.

Kuznetsov, G. & Nigam, S. K. (1998). Folding of secretory and membrane proteins. *New England Journal of Medicine* **339**, 1688-1695.

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., Hoft, D., Houghton, M., Belshe, R. B. & Ray, R. (1995). Immune responses to plasmid DNA encoding the hepatitis C virus core protein. *Journal of Virology* **69**, 5859-5863.

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.

Lanford, R. E., Sureau, C., Jacob, J. R., White, R. & Fuerst, T. R. (1994). Demonstration of *in vitro* infection of chimpanzee hepatocytes with hepatitis C virus using strand-specific RT/PCR. *Virology* **202**, 606-614.

Lanford, R. E., Lee, H., Chavez, D., Guerra, B. & Brasky, K. M. (2001). Infectious cDNA clone of the hepatitis C virus genotype 1 prototype sequence. *Journal of General Virology* **82**, 1291-1297.

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.

Lau, J. Y. N., Davis, G. L., Kniffen, J., Qian, K.-P., Urdea, M. S., Chan, C. S., Mizokami, M., Neuwald, P. D. & Wilber, J. C. (1993). Significance of serum hepatitis C virus RNA levels in chronic hepatitis C. *The Lancet* **341**, 1501-1504.

Lauer, G. M. & Walker, B. D. (2001). Hepatitis C virus infection. *New England Journal of Medicine* **345**, 41-52.

Laurent, A. G., Krust, B., Galabru, J., Svab, J. & Hovanessian, A. G. (1985). Monoclonal antibodies to an interferon-induced Mr 68,000 protein and their use for the detection of double-stranded RNA-dependent protein kinase in human cells. *Proceedings of the National Academy of Science USA* **82**, 4341-4345.

Lerat, H., Rumin, S., Habersetzer, F., Berby, F., Trabaud, M. A., Trepo, C. & Inchauspe, G. (1998). *In vivo* tropism of hepatitis C virus genomic sequences in hematopoietic cells: Influence of viral load, viral genotype, and cell phenotype. *Blood* **91**, 3841-3849.

Levy, S., Todd, S. C. & Maecker, H. T. (1998). CD81 (TAPA-1): A molecule involved in signal transduction and cell adhesion in the immune system. *Annual Reviews in Immunology* **16**, 89-109.

Liberman, E., Fong, Y.-L., Selby, M. J., Choo, Q.-L., Cousens, L., Houghton, M. & Benedict Yen, T. S. (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. (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, B. M., McCourt, D. W. & Rice, C. M. (1994). Processing in 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.

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**, 4373-4780.

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.

Lissen, E., Alter, H. J., Abad, M. A., Torres, Y., Perez-Romereo, M., Leal, M. & Pineda, J. A. (1993). Hepatitis C virus infection among sexually promiscuous groups and the heterosexual partners of hepatitis C virus infected index cases. *European Journal of Clinical Microbiology and Infectious Diseases* **12**, 827-831.

Liu, Q., Bhat, R. A., Prince, A. M. & Zhang, P. (1999). The hepatitis C virus NS2 protein generated by NS2-3 autocleavage is required for NS5A phosphorylation. *Biochemical and Biophysical Research Communications* **254**, 572-577.

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., 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.

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.

Lu, Y., Wambach, M., Katze, M. G. & Krug, R. M. (1995). Binding of the influenza virus NS1 protein to double-stranded RNA inhibits the activation of the protein kinase that phosphorylates the eIF-2 translation initiation factor. *Virology* **214**, 222-228.

Luo, G. (1999). Cellular proteins bind to the poly(U) tract of the 3' untranslated region of the hepatitis C virus RNA genome. *Virology* **256**, 105-118.

Major, M. E. & Feinstone, S. M. (2000). Characterization of hepatitis C virus infectious clones in chimpanzees: long-term studies. *Current Topics in Microbiology and Immunology* **242**, 279-298.

Manickan, E., Rouse, R. J., Yu, Z., Wire, W. S. & Rouse, B. T. (1995). Genetic immunization against herpes simplex virus: Protection is mediated by CD4+ T lymphocytes. *Journal of Immunology* **155**, 259-265.

Manns, M. P. & Rambusch, E. G. (1999). Autoimmunity and extrahepatic manifestations in hepatitis C virus infection. *Journal of Hepatology* **31**, 39-42.

Marcellin, P. (1999). Hepatitis C: the clinical spectrum of the disease. *Journal of Hepatology* **31**, 9-16.

Martell, M., Esteban, J. I., Quer, J., Genesca, J., Weiner, A., Esteban, R., Guardia, J. & Gomez, J. (1992). Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: Quasispecies nature of HCV genome distribution. *Journal of Virology* **66**, 3225-3229.

Martinot-Peignoux, M., Marcelli, P., Pouteau, M., Castelnau, C., Boyer, N., Poliquin, M., Degott, C., Descombes, I., Le Breton, V., Milotova, V., Benhamou, J. P. & Erlinger, S. (1995). Pretreatment serum hepatitis C virus RNA levels and hepatitis C virus genotype are the main and independent prognostic factors of sustained response to interferon alfa therapy in chronic hepatitis. *Hepatology* **22**, 1050-1056.

Martinot-Peignoux, M., Roudot-Thoraval, F., Mendel, I., Coste, J., Izopet, J., Duverlie, G., Payan, C., Pawlotsky, J. M., Defer, C., Bogard, M., Gerolami, V., Halfon, P., Buisson, Y., Fouqueray, B., Loiseau, P., Lamoril, J., Lefrere, J. J. & Marcellin, P. (1999). Hepatitis C virus genotypes in France: relationship with epidemiology, pathogenicity and response to interferon therapy. *Journal of Viral Hepatitis* **6**, 435-443.

Martire, G., Viola, A., Iodice, L., Lotti, L. V., Gradini, R. & Bonatti, S. (2001). Hepatitis C virus structural proteins reside in the endoplasmic reticulum as well as in the intermediate compartment/cis-Golgi complex region of stably transfected cells. *Virology* **280**, 176-182.

Matsumoto, M., Hwang, S. B., Jeng, K.-S., Zhu, N. & Lai, M. M. C. (1996). Homotypic interaction and multimerization of hepatitis C virus core protein. *Virology* **218**, 43-51.

Matsuura, Y., Suzuki, T., Suzuki, R., Sato, M., Aizaki, H., Saito, I. & Miyamura, T. (1994). Processing of the E1 and E2 glycoproteins of hepatitis C virus expressed in mammalian and insect cells. *Virology* **205**, 141-150.

McHutchison, J. G., Gordon, S. C., Schiff, E. R., Shiffman, M. L., Lee, W. M., Rustgi, V. K., Goodman, Z. D., Ling, M. H., Cort, S. & Albrecht, J. K. (1998). Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. *New England Journal of Medicine* **339**, 1485-1492.

McKechnie, V. M., Mills, P. R. & McCruden, E. A. (2000). The NS5a gene of hepatitis C virus in patients treated with interferon-alpha. *Journal of Medicine Virology* **60**, 367-378.

McLauchlan, J. (2000). Properties of the hepatitis C virus core protein: a structural protein that modulates cellular processes. *Journal of Viral Hepatitis* **7**, 2-14.

McOmish, F., Yap, P. L., Dow, B. C., Follett, E. A. C., Seed, C., Keller, A. J., Cobain, T. J., Krusius, T., Kolho, E., Naukkarinen, R., Lin, C., Lai, C., Leong, S., Medgysei, G. A., Hejjas, M., Kiyokawa, H., Fukada, K., Cuypers, T., Saeed, A. A., Al-Rasheed, A. M., Lin, M. & Simmonds, P. (1994). Geographical distribution of hepatitis C virus genotypes in blood donors:an international collaborative survey. *Journal of Clinical Microbiology* **32**, 884-892.

Melen, K., Keskinen, P., Lehtonen, A. & Julkunen, I. (2000). Interferon-induced gene expression and signalling in human hepatoma cell lines. *Journal of Hepatology* **33**, 764-772.

Meola, A., Sbardellati, A., Bruni Ercole, B., Cerretani, M., Pezzanera, M., Ceccacci, A., Vitelli, A., Levy, S., Nicosia, A., Traboni, C., McKeating, J. & Scarselli, E. (2000). Binding of Hepatitis C Virus E2 Glycoprotein to CD81 Does Not Correlate with Species Permissiveness to Infection. *Journal of Virology* **74**, 5933-5938.

Mercer, D. F., Schiller, D. E., Elliott, J. F., Douglas, D. N., Hao, C., Rinfret, A., Addison, W. R., Fischer, K. P., Churchill, T. A., Lakey, J. R. T., Tyrrell, D. L. J. & Kneteman, N. M. (2001). Hepatitis C virus replication in mice with chimeric human livers. *Nature Medicine* **7**, 927-933.

Merola, M., Brazzoli, M., Cocchiarella, F., Heile, J. M., Helenius, A., Weiner, A. J., Houghton, M. & Abrignani, S. (2001). Folding of hepatitis C virus E1 glycoprotein in a cell-free system. *Journal of Virology* **75**, 11205-11217.

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.

Meurs, E., Chong, K., Galabru, J., Thomas, N. S. B., Kerr, I. M., Williams, B. R. G. & Hovanessian, A. G. (1990). Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. *Cell* **62**, 379-390.

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.

Mihm, S., Fayyazi, A., Hartmann, H. & Ramadori, G. (1997). Analysis of histopathological manifestations of chronic hepatitis C virus infection with respect to virus genotype. *Hepatology* **25**, 735-739.

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.

Miyamura, T. & Matsuura, Y. (1993). Structural proteins of hepatitis C virus. *Trends in Microbiology* **1**, 229-231.

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.

Molinari, M. & Helenius, A. (2000). Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. *Science* **288**, 331-333.

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.

Moradpour, D., Englert, C., Wakita, T. & Wands, J. R. (1996). Characterization of cell lines allowing tightly regulated expression of hepatitis C virus core protein. *Virology* **222**, 51-63.

Mori, K. (2000). Tripartite management of unfolded proteins in the endoplasmic reticulum. *Cell* **101**, 451-454.

Moriya, K., Yotsuyanagi, H., Ishibashi, K., Shintani, Y., Fujie, H., Matsuura, Y., Miyamura, T. & Koike, K. (1997). Hepatitis C virus core protein induces steatosis in transgenic mice. *Journal of General Virology* **78**, 1527-1531.

Moriya, K., Fujie, H., Yotsuyanagi, H., Shintani, Y., Tsutsumi, T., Ishibashi, K., Matsuura, Y., Kimura, S., Miyamura, T. & Koike, K. (1998). The core protein of hepatitis C virus induces hepatocellular carcinoma in transgenic mice. *Nature Medicine* **4**, 1065-1067.

Nakano, I., Maertens, G., Major, M. E., Vitvitski, L., Dubuisson, J., Fournillier, A., De Martynoff, G., Trepo, C. & Inchauspe, G. (1997). Immunization with plasmid DNA encoding hepatitis C virus envelope E2 antigenic domains induces antibodies whose immune reactivity is linked to the injection mode. *Journal of Virology* **71**, 7101-7109.

Naoumov, N. V. (1999). Hepatitis C virus infection in Eastern Europe. *Journal of Hepatology* **31**, 84-87.

Ogata, N., Alter, H. J., Miller, R. H. & Purcell, R. H. (1991). Nucleotide sequence and mutation 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.

Ohno, T., Mizokami, M., Ohba, K., Suzuki, K., Wu, R. R. & Nouriaria, K. T. (1994). New genotype of hepatitis C virus in South Africa. *Journal of Medical Virology* **42**, 409-413.

Ohto, H., Terazawa, S., Nobuhiko, S., Sasaki, N., Hino, K. Ishiwata, C., Kako, M., Ujiie, N., Endo, C., Matsui, A., Okamoto, H. & Mishiro, S. (1994). Transmission of hepatitis C virus from mothers to infants. *New England Journal of Infectious Diseases* **330**, 744-750.

Okuda, M., Hino, K., Korenaga, M., Yamaguchi, Y., Katoh, Y. & Okita, K. (1999). Differences in hypervariable region 1 quasispecies of hepatitis C virus in human serum, peripheral blood mononuclear cells, and liver. *Hepatology* **29**, 217-222.

Owsianka, A. M. & Patel, A. H. (1999). Hepatitis C virus core protein interacts with a human DEAD box protein DDX3. *Virology* **257**, 330-340.

Owsianka, A., Clayton, R. F., Loomis-Price, L. D., McKeating, J. A. & Patel, A. H. (2001). Functional analysis of hepatitis C virus E2 glycoproteins and virus-like particles reveals structural dissimilarities between different forms of E2. *Journal of General Virology* **82**, 1877-1883.

Parodi, A. J. (1999). Reglucosylation of glycoproteins and quality control of glycoprotein folding in the endoplasmic reticulum of yeast cells. *Biochemica et Biophysica Acta* **1426**, 287-295.

Patel, A. H., Wood, J., Penin, F., Dubuisson, J. & McKeating, J. A. (2000). Construction and characterization of chimeric hepatitis C virus E2 glycoproteins: analysis of regions critical for glycoprotein aggregation and CD81 binding. *Journal of General Virology* **81**, 2873-2883.

Patel, J., Patel, A. & McLauchlan, J. (1999a). Covalent interactions are not required to permit or stabilize the non-covalent association of hepatitis C virus glycoproteins E1 and E2. *Journal of General Virology* **80**, 1681-2883.

Patel, J. (1999b). An investigation of the complexes formed between the hepatitis C virus E1 and E2 glycoproteins. In *Institute of Virology*. Glasgow: University of Glasgow.

Patel, J., Patel, A. & McLauchlan, J. (2001). The Transmembrane Domain of the Hepatitis C Virus E2 Glycoprotein Is Required for Correct Folding of the E1 Glycoprotein and Native Complex Formation. *Virology* **279**, 58-68.

Paterson, M., Laxton, C. D., Thomas, H. C., Ackrill, A. M. & Foster, G. R. (1999). Hepatitis C virus NS5A protein inhibits interferon antiviral activity, but the effects do not correlate with clinical response. *Gastroenterology* **117**, 1187-1197.

Patil, C. & Walter, P. (2001). Intracellular signalling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals. *Current Opinion in Cell Biology* **13**, 349-356.

Pavio, N., Taylor, D. R. & Lai, M. M. C. (2001). Characterization of an unglycosylated form of HCV E2 protein that is located in the cytosol and interacts with PKR. In *8th International Symposium on Hepatitis C Virus & Related Viruses*, pp. 142. Paris.

Pavlovic, J., Zurcher, T., Haller, O. & Staeheli, P. (1990). Resistance to influenza virus and vesicular stomatitis virus conferred by expression of human MxA protein. *Journal of Virology* **64**, 3370-3375.

Pawlotsky, J. M., Tsakiris, L., Roudot-Thoraval, F., Pellet, C., Stuyver, L., Duval, J. & Dhumeaux, D. (1995). Relationship between hepatitis C virus genotypes and sources of infection in patients with chronic hepatitis C. *Journal of Infectious Diseases* **171**, 1607-1610.

Pellegrini, S., John, J., Shearer, M., Kerr, I. M. & Stark, G. R. (1989). Use of a selectable marker regulated by alpha interferon to obtain mutations in the signalling pathway. *Molecular Cell Biology* **9**, 4605-4612.

Pestova, T. V., Shatsky, I. N., Fletcher, S. P., Jackson, R. J. & Hellen, C. U. (1998). A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes & Development* **12**, 67-83.

Petracca, R., Falugi, F., Galli, G., Norais, N., Rosa, D., Campagnoli, S., Burgio, V., Di Stasio, E., Giardina, B., Houghton, M., Abrignani, S. & Grandi, G. (2000). Structure-function analysis of hepatitis C virus envelope-CD81 binding. *Journal of Virology* **74**, 4824-4830.

Petrik, J., Parker, H. & Alexander, G. J. M. (1999). Human hepatic glyceraldehyde-3-phosphate dehydrogenase binds to the poly(U) tract of the 3' non-coding region of hepatitis C virus genomic RNA. *Journal of General Virology* **80**, 3109-3113.

Pietschmann, T., Lohmann, V., Rutter, G., Krieger, N. & Bartenschlager, R. (2001). Persistent replication of a selectable hepatitis C virus full length genome in cell culture. In *8th International Symposium on Hepatitis C Virus & Related Viruses*, pp. 60. Paris.

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.

Plemper, R. K. & Wolf, D. H. (1999). Retrograde protein translocation: ERADicadtion of secretory proteins in health and disease. *Trends in Biochemical Science* **24**, 266-270.

Podevin, P., Sabile, A., Gajardo, R., Delhem, N., Abadie, A., Lozach, P. Y., Beretta, L. & Brechot, C. (2001). Expression of hepatitis C virus NS5A natural mutants in a hepatocytic cell line inhibits the antiviral effect of interferon in a PKR-independent manner. *Hepatology* **33**, 1503-1511.

Polyak, S. J., Tang, N., Wambach, M., Barber, G. N. & Katze, M. G. (1996). The P58 cellular inhibitor complexes with the interferon-induced, double-stranded RNA-dependent protein kinase, PKR, to regulate its autophosphorylation and activity. *Journal of Biological Chemistry* **271**, 1702-1707.

Polyak, S. J., Paschal, D. M., McArdle, S., Gale, M. J., Jr., Moradpour, D. & Gretch, D. R. (1999). Characterization of the effects of hepatitis C virus nonstructural 5A protein expression in human cell lines and on interferon-sensitive virus replication. *Hepatology* **29**, 1262-1271.

Polyak, S. J., Nousbaum, J.-B., Larson, A. M., Cotler, S., Carithers, R. L. J. & Gretch, D. R. (2000). The protein kinase-interacting domain in the Hepatitis C Virus envelope glycoprotein-2 gene is highly conserved in genotype-1 infected patients treated with interferon. *Journal of Infectious Diseases* **182**, 397-404.

Polyak, S. J., Khabar, K. S. A., Paschal, D. M., Ezelle, H. J., Duverlie, G., Barber, G. N., Levy, D. E., Mukaida, N. & Gretch, D. R. (2001a). Hepatitis C virus nonstructural 5A protein induces interleukin-8, leading to partial inhibition of the interferon-induced antiviral response. *Journal of Virology* **75**, 6095-6106.

Polyak, S. J., Khabar, K. S., Rezeiq, M. & Gretch, D. R. (2001b). Elevated levels of interleukin-8 in serum are associated with hepatitis C virus infection and resistance to interferon therapy. *Journal of Virology* **75**, 6209-6211.

Poynard, T., Leroy, V., Cohard, M., Thevenot, T., Mathurin, P., Opolon, P. & Zarski, J. P. (1996). Meta-analysis of interferon randomised trials in the treatment of viral hepatitis C: effects of dose and duration. *Hepatology* **24**, 778-89.

Poynard, T., Marcellin, P., Lee, S., Niederau, C., Minuk, G. S., Ideo, G., Bain, V., Heathcote, J., Zeuzem, S., Trepo, C. & Albrecht, J. (1998). Randomised trial of interferon alpha 2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha 2b plus placebo for 48 weeks treatment of chronic infection with hepatitis C virus. *The Lancet* **352**, 1426-1432.

Prince, A. M., Huima-Byron, T., Parker, T. S. & Levine, M. M. (1996). Visualization of hepatitis C virions and putative defective interfering particles isolated from low-density lipoproteins. *Journal of Viral Hepatology* **3**, 11-17.

Puig-Basagoiti, F., Saiz, J.-C., Forns, X., Ampurdanes, S., Gimenez-Barcons, M., Franco, S., Sanchez-Fueyo, A., Costa, J., Sanchez-Tapias, J.-M. & Rodes, J. (2001). Influence of the Genetic Heterogeneity of the ISDR and PePHD Regions of Hepatitis C Virus on the Response to Interferon Therapy in Chronic Hepatitis C. *Journal of Medical Virology* **65**, 35-44.

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.

Ray, R. B., Lagging, L. M., Meyer, K. & Ray, R. (1996). Hepatitis C virus core protein co-operates with ras and transforms rat embryo fibroblasts to tumorigenic phenotype. *Journal of Virology* **70**, 4438-4443.

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.

Reed, K. E., Xu, J. & Rice, C. M. (1997). Phosphorylation of the hepatitis C virus NS5A protein in vitro and in vivo: properties of the NS5A-associated kinase. *Journal of Virology* **71**, 7187-7197.

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 translation. *The EMBO Journal* **14**, 6010-6020.

Rice, C. M. (1996). Flaviviridae: the viruses and their replication. In *Fields Virology*, Third edn, pp. 931-959. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott-Raven publishers.

Robertson, B., Myers, G., Howard, C., Brettin, T., Bukh, J., Gaschen, B., Gojobori, T., Maertens, G., Mizokami, M., Nainan, O., Netesov, S., Nishioka, K., Shini, T., Simmonds, P., Smith, D., Stuyver, L. & Weiner, A. (1998). Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. *Archives of Virology* **143**, 2493-2503.

Rosa, D., Campagnoli, S., Moretto, C., Guenzi, E., Cousens, L., Chin, M., Dong, C., Weiner, A. J., Lau, J. Y., 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.

Rubbia-Brandt, L., Quadri, R., Abid, K., Giostra, E., Male, P.-J., Mentha, G., Spahr, L., Zarski, J.-P., Borisch, B., Hadengue, A. & Negro, F. (2000). Hepatocyte steatosis is a cytopathic effect of hepatitis C virus genotype 3. *Journal of Hepatology* **33**, 106-115.

Rubbia-Brandt, L., Leandro, G., Spahr, L., Giostra, E., Quadri, R., Male, P. J. & Negro, F. (2001). Liver steatosis in chronic hepatitis C: a morphological sign suggesting infection with HCV genotype 3. *Histopathology* **39**, 119-124.

Ruggieri, A., Argentini, C., Kourama, F., Chionne, P., Dugo, E., Spada, E., Sabbatani, S. & Rapicetta, M. (1996). Heterogeneity of hepatitis C virus genotype 2 variants in West Central Africa (Guinea Conakry). *Journal of General Virology* **77**, 2073-2076.

Ruggieri, A., Harada, T., Matsuura, Y. & Miyamura, T. (1997). Sensitization to Fas-mediated apoptosis by hepatitis C virus core protein. *Virology* **229**, 68-76.

Rumin, S., Berthillon, P., Tanaka, E., Kiyosawa, K., Trabaud, M. A., Bizollen, T., Gouillat, C., Gripon, P., Guguen, G. C., Inchauspe, G. & Trepo, C. (1999). Dynamic analysis of hepatitis C virus replication and quasispecies selection in long-term cultures of adult human hepatocytes infected in vitro. *Journal of General Virology* **80**, 3007-3018.

Ryan, M. D., Monaghan, S. & Flint, M. (1998). Virus-encoded proteinases of the Flaviviridae. *Journal of General Virology* **79**, 947-959.

Saiz, J. C., Lopez-Labrador, F. X., Ampurdanes, S., Dopazo, J., Forns, X., Sanchez-Tapias, J. M. & Rodes, J. (1998). The prognostic relevance of the nonstructural 5A gene

interferon sensitivity determining region is different in infections with genotype 1b and 3a isolates of hepatitis C virus. *Journal of Infectious Diseases* **177**, 839-847.

Sakamoto, M., Akahane, Y., Tsuda, F., Tanaka, T., Woodfield, D. G. & Okamoto, H. (1994). Entire nucleotide sequence and characterization of a hepatitis C virus of genotype V/3a. *Journal of General Virology* **75**, 1761-1768.

Sakamuro, D., Furukawa, T. & Takegami, T. (1995). Hepatitis C virus nonstructural protein NS3 transforms NIH 3T3 cells. *Journal of Virology* **69**, 3893-3896.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). Molecular cloning: A laboratory manual, Second Edition edn. Edited by N. Ford, C. Nolan & M. Ferguson: Cold Spring Harbor Laboratory Press, USA.

Santolini, E., Migliaccio, G. & La Monica, 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. & La Monica, N. (1995). The NS2 protein of hepatitis C virus is a transmembrane polypeptide. *Journal of Virology* **69**, 7461-7471.

Sarrazin, C., Berg, T., Lee, J. H., Teuber, G., Dietrich, C. F., Roth, W. K. & Zeuzem, S. (1999). Improved correlation of between multiple mutations within the NS5A region and virological response in European patients chronically infected with hepatitis C virus type 1b undergoing combination therapy. *Journal of Hepatology* **30**, 1004-1013.

Sarrazin, C., Kornetzky, I., Ruster, B., Lee, J.-H., Kronenberger, B., Bruch, K., Roth, W. K. & Zeuzem, S. (2000a). Mutations Within the E2 and NS5A Protein in Patients Infected With Hepatitis C Virus Type 3a and Correlation With Treatment Response. *Hepatology* **31**, 1360-1350.

Sarrazin, C., Berg, T., Lee, J.-H., Ruster, B., Kronenberger, B., Roth, W. K. & Zeuzem, S. (2000b). Mutations in the protein kinase-binding domain of the NS5A protein in patients infected with Hepatitis C virus type 1a are associated with treatment response. *Journal of Infectious Diseases* **181**, 432-441.

Satoh, S., Tanji, Y., Hijikata, M., Kimura, K. & Shimotohno, K. (1995). The N-terminal region of hepatitis C virus nonstructural protein 3 (NS3) is essential for stable complex formation with NS4A. *Journal of Virology* **69**, 4255-4260.

Schalm, S., Hansen, B., Chemello, L., Bellobuono, A., Brouwer, J. T., Weiland, O., Cavalletto, L., Schvarcz, R., Ideo, G. & Alberti, A. (1997). Ribavirin enhances the efficacy but not the adverse effects of interferon in chronic hepatitis C. Meta-analysis of individual patient data from European centers. *Journal of Hepatology* **26**, 961-966.

Selby, M. J., Choo, Q. L., Berger, K., Kuo, G., Glazer, E., Eckart, M., Lee, C., Chien, D., Kuo, C. & Houghton, M. (1993). Expression, identification and subcellular localization of the proteins encoded by the hepatitis C viral genome. *Journal of General Virology* **74**, 1103-1113.

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.

Seong, Y. R., Lee, C.-H. & Im, D.-S. (1998). Characterization of the structural proteins of hepatitis C virus expressed by an adenovirus recombinant. *Virus Research* **55**, 177-185.

Shi, Y., An, J., Liang, J., Hayes, S. E., Sandusky, G. E., Stramm, L. E. & Yang, N. N. (1999). Characterization of a mutant pancreatic eIF2-a kinase, PEK, and co-localization with somatostatin in islet delta cells. *Journal of Biological Chemistry* **274**, 5723-5730.

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.

Shimoike, T., Mimori, S., Tani, H., Matsuura, Y. & Miyamura, T. (1999). Interaction of Hepatitis C Virus Core Protein with Viral Sense RNA and Suppression of Its Translation. *Journal of Virology* **73**, 9718-9725.

Sidrauski, C., Cox, J. S. & Walter, P. (1996). tRNA ligase is required for regulated mRNA splicing in the unfolded protein response. *Cell* **87**, 405-413.

Sidrauski, C. & Walter, P. (1997). The transmembrane kinase Ire1p is the site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. *Cell* **90**, 1031-1039.

Siegel, V. (1995). A second signal recognition event required for translocation into the endoplasmic reticulum. *Cell* **82**, 167-170.

Simmonds, P., Holmes, E. C., Cha, T.-A., Chan, S.-W., McOmish, F., Irvine, B., Beall, E., Yap, P. L., Kolberg, J. & Urdea, M. S. (1993). Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *Journal of General Virology* **74**, 2391-2399.

Simmonds, P., Smith, D. B., McOmish, F., Yap, P. L., Kolberg, J., Urdea, M. S. & Holmes, E. C. (1994). Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS-5 regions. *Journal of General Virology* **75**, 1053-1061.

Simmonds, P. (1995). Variability of hepatitis C virus. *Hepatology* **21**, 570-583.

Simmonds, P. (2001). The origin and evolution of hepatitis viruses in humans. *Journal General Virology* **82**, 693-712.

Sjoberg, E. M., Suomalainen, M. & Garoff, H. (1994). A significantly improved semliki forest virus expression system based on the translation enhancer segments from the viral capsid gene. *Biotechnology* **12**, 1127-1131.

Smedile, A. & Verme, G. (1999). Hepatitis D Virus: biology, pathogenesis, epidemiology, clinical description, and therapy. In *Viral Hepatitis: Diagnosis, Therapy, and Prevention*, pp. 129-150. Edited by S. Specter. Totowa, NJ: Humana Press Inc.

Smith, D. B., Pathirana, S., Davidson, F., Lawlor, E., Power, J., Yap, P. L. & Simmonds, P. (1997a). The origin of hepatitis C virus genotypes. *Journal of General Virology* **78**, 321-328.

Smith, D. B., McAllister, J., Casino, C. & Simmonds, P. (1997b). Virus 'quasispecies': making a mountain out of a molehill? *Journal of General Virology* **78**, 1511-1519.

Smith, D. B., Cuceanu, N., Davidson, F., Jarvis, L. M., Mokili, J. L. K., Hamid, S., Ludlam, C. A. & Simmonds, P. (1997c). Discrimination of hepatitis G virus/GBV-C geographical variants by analysis of the 5' non-coding region. *Journal of General Virology* **78**, 1533-1542.

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.

Spangberg, K., Goobar-Larsson, L., Wahren-Herlenius, M. & Schwartz, S. (1999). The La protein from human liver cells interacts specifically with the U-rich region in the hepatitis C virus 3' untranslated region. *Journal of Human Virology* **2**, 296-307.

Specter, S. (1999). Viral Hepatitis: diagnosis, therapy and prevention, pp. 1-402. Edited by S. Specter. Totowa, NJ: Humana Press Inc.

Steinkuhler, C., Urbani, A., Tomei, L., Biasiol, G., Sardana, M., Bianchi, E., Pessi, A. & De Francesco, R. (1996). Activity of purified hepatitis C virus protease NS3 on peptide substrates. *Journal of Virology* **70**, 6694-6700.

Stempniak, M., Hostomska, Z., Nodes, B. R. & Hostomsky, Z. (1997). The NS3 proteinase domain of hepatitis C virus is a zinc-containing enzyme. *Journal of Virology* **71**, 2881-2886.

Suzich, J. A., Tamura, J. K., Palmer-Hill, F., Warrener, P., Grakoui, A., Rice, C. M., Feinstone, S. M. & Collett, M. S. (1993). Hepatitis C virus NS3 protein polynucleotide-stimulated nucleoside triphosphate and comparison with the related pestivirus and flavivirus enzymes. *Journal of Virology* **67**, 6152-6158.

Tai, C.-L., Chi, W.-K., Chen, D.-S. & Hwang, L.-H. (1996). The helicase activity associated with hepatitis C virus nonstructural protein 3 (NS3). *Journal of Virology* **70**, 8477-8484.

Takada, N., Takase, S., Takada, A. & Date, T. (1993). Differences in the hepatitis C virus genotypes in different countries. *Journal of Hepatology* **17**, 277-283.

Takikawa, S., Ishii, K., Aizaki, H., Suzuki, T., Asakura, H., Matsuura, Y. & Miyamura, T. (2000). Cell Fusion Activity of Hepatitis C Virus Envelope Proteins. *Journal of Virology* **74**, 5066-5074.

Tanji, Y., Kaneko, T., Satoh, S. & Shimotohno, K. (1995). Phosphorylation of hepatitis C virus-encoded nonstructural protein NS5A. *Journal of Virology* **69**, 3980-3986.

Taylor, D. R., Shi, S. T., Romano, P. R., Barber, G. N. & Lai, M. M. (1999). Inhibition of the interferon- inducible protein kinase PKR by HCV E2 protein. *Science* **285**, 107-10.

Taylor, D. R. (2001). Hepatitis C virus and interferon resistance: It's more than just PKR. *Hepatology* **33**, 1547-1549.

Taylor, D. R., Tian, B., Romano, P. R., Hinnebusch, A. G., Lai, M. M. C. & Mathews, M. B. (2001). Hepatitis C Virus Envelope Protein E2 Does Not Inhibit PKR by Simple Competition with Autophosphorylation Sites in the RNA-Binding Domain. *Journal of Virology* **75**, 1265-1273.

Theodore, D. & Fried, M. W. (2000). Natural history and disease manifestations of hepatitis C infection. *Current Topics in Microbiology and Immunology* **242**, 43-54.

Thomas, D. L. (2000). Hepatitis C epidemiology. *Current Topics in Microbiology and Immunology* **242**, 25-41.

Tokita, H., Shrestha, S. M., Okamoto, H., Sakamoto, M., Horikita, M., Iizuka, H., Shrestha, S., Miyakawa, Y. & Mayumi, M. (1994). Hepatitis C virus variants from Nepal with novel genotypes and their classification into the third major group. *Journal of General Virology* **75**, 931-936.

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.

Trepo, C. & Pradat, P. (1999). Hepatitis C virus infection in Western Europe. *Journal of Hepatology Supplement* **31**, 80-83.

Trepo, C. (2000). Genotype and viral load as prognostic indicators in the treatment of hepatitis C. *Journal of Viral Hepatitis* **7**, 250-257.

Trepo, C., Lindsay, K., Niederau, C., Shiffman, M., Gordon, S., Hoefs, J., Schiff, E., Marcellin, P., Bacon, B., Fang, J., Garaud, J. & Albrecht, J. (2000). Pegylated interferon alfa-2b (Peg Intron) monotherapy is superior to interferon alfa-2b (Intron A) for the treatment of chronic hepatitis C. *Journal of Hepatology* **32**, 29.

Tsubota, A., Chayama, K., Ikeda, K., Yasuji, A., Koida, I., Saitoh, S., Hashimoto, M., Iwasaki, S., Kobayashi, M. & Hiromitsu, K. (1994). Factors predictive of response to interferon - alpha therapy in hepatitis C virus infection. *Hepatology* **19**, 1088-1094.

Tsuchihara, K., Tanaka, T., Hijikata, M., Kuge, S., Toyoda, H., Nomoto, A., Yamamoto, N. & Shimotohno, K. (1997). Specific interaction of polypyrimidine tract-binding protein with the extreme 3'-terminal structure of the hepatitis C virus genome, the 3'X. *Journal of Virology* **71**, 6720-6726.

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. & Nomoto, A. (1992). Internal Ribosome Entry Site within Hepatitis C Virus RNA. *Journal of Virology* **66**, 1476-1483.

Ulmer, J. B., Donnelly, J. J., Parker, S. E., Rhodes, G. H., Felgner, P. L., Dwarki, V. J., Gromkowski, S. H., Deck, R. R., DeWitt, C. M., Friedman, A., Hawe, L. A., Leander, K. R., Martinez, D., Perry, H. C., Shiver, J. W., Montgomery, D. L. & Liu, M. A. (1993). Heterologous protection against influenza by injection of DNA encoding a viral protein. *Science* **259**, 1745-1749.

Van Doorn, L.-J., Capriles, I., Maertens, G., DeLeys, R., Murray, K., Kos, T., Schellekens, H. & Quint, W. (1995). Sequence evolution of the hypervariable region in the putative envelope region E2/NS1 of hepatitis C virus is correlated with specific humoral immune responses. *Journal of Virology* **69**, 773-778.

Vyas, G. N. & Yen, T. S. B. (1999). Hepatitis B Virus: biology, pathogenesis, epidemiology, clinical description and diagnosis. In *Viral Hepatitis: Diagnosis, Therapy and Prevention*, pp. 35-63. Edited by S. Specter. Totowa, NJ: Humana Press Inc.

Wang, C., Sarnow, P. & Siddiqui, A. (1993). Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *Journal of Virology* **67**, 3338-3344.

Wang, T. H., Rijnbrand, R. C. & Lemon, S. M. (2000). Core protein-coding sequence, but not core protein, modulates the efficiency of cap-independent translation directed by the internal ribosome entry site of hepatitis C virus. *Journal of Virology* **74**, 11347-11358.

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., Geysen, H. M., Christopherson, C., Hall, J. E., Mason, T. J., Saracco, G., Bonino, F., Crawford, K., Marion, C. D., Crawford, K. A., Brunetto, M., Barr, P. J., Miyamura, T., McHutchinson, J. & Houghton, M. (1992). Evidence for immune selection of hepatitis C virus (HCV) putative envelope glycoprotein variants: Potential role in chronic HCV infections. *Proceedings of the National Academy of Sciences* **89**, 3468-3472.

WHO (1997). Hepatitis C: global prevalence. *Weekly Epidemiological Records*, 341-348.

Witherell, G. W. & Beineke, P. (2001). Statistical analysis of combined substitutions in nonstructural 5A region of hepatitis C virus and interferon response. *Journal of Medical Virology* **63**, 8-16.

Wolk, B., Sansonno, D., Krausslich, H.-G., Dammacco, F., Rice, C.M., Blum, H.E. & Moradpour, D. (2000). Subcellular localization, stability, and trans-cleavage competence of the hepatitis C virus NS3-NS4A complex expressed in tetracycline-regulated cell lines. *Journal of Virology* **74**, 2293-2304.

Wood, J., Frederickson, R. M., Fields, S. & Patel, A. H. (2001). Hepatitis C Virus 3'X Region Interacts with Human Ribosomal Proteins. *Journal of Virology* **75**, 1348-2129.

Wunschmann, S., Medh, J. D., Klinzmann, D., Schmidt, W. N. & Stapleton, J. T. (2000). Characterization of hepatitis C virus (HCV) and HCV E2 interactions with CD81 and the low-density lipoprotein receptor. *Journal of Virology* **74**, 10055-10062.

Xiang, Z. Q., Spitalnik, S., Tran, M., Wunner, W. H., Cheng, J. & Ertl, H. C. (1994). Vaccination with a plasmid vector carrying the rabies glycoprotein gene induces protective immunity against rabies virus. *Virology* **199**, 132-140.

Xu, L. Z., Larzul, D., Delaporte, E., Brechot, C. & Kremsdorf, D. (1994). Hepatitis C virus genotype 4 is highly prevalent in Central Africa (Gabon). *Journal of General Virology* **75**, 2393-2398.

Yagnik, A. T., Lahm, A., Meola, A., Roccasecca, R. M., Ercole, B. B., Nicosia, A. & Tramontano, A. (2000). A model for the hepatitis C virus envelope glycoprotein E2. *Proteins* **40**, 355-66.

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 USA* **94**, 8738-8743.

Yanagi, M., StClaire, M., Shapiro, M., Emerson, S. U., Purcell, R. H. & Bukh, J. (1998). Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious *in vivo*. *Virology* **244**, 161-172.

Yanagi, M., Purcell, R. H., Emerson, S. U. & Bukh, J. (1999a). Hepatitis C virus: an infectious molecular clone of a second major genotype (2a) and lack of viability of intertypic 1a and 2a chimeras. *Virology* **262**, 250-263.

Yanagi, M., St. Claire, M., Emerson, S., Purcell, R. H. & Bukh, J. (1999b). In vivo analysis of the 3' untranslated region of the hepatitis C virus after in vitro mutagenesis of an infectious cDNA clone. *Proceedings of the National Academy of Sciences USA* **96**, 2291-2295.

Yap, P. L., McOmish, F., Webster, A. D. B., Hammarstrom, L., Smith, C. I. E., Bjorkander, J., Ochs, H. D., Fischer, S. H., Quinti, I. & Simmonds, P. (1994). Hepatitis C virus transmission by intravenous immunoglobulin. *Journal of Hepatology* **21**, 455-460.

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., 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.

You, L.-R., Chen, C.-M., Yeh, T.-S., Tsai, T.-Y., Mai, R.-T., Lin, C.-H. & Lee, Y.-H. W. (1999). Hepatitis C virus core protein interacts with a cellular putative RNA helicase. *Journal of Virology* **73**, 2841-2853.

Yuki, N., Hayashi, N., Moribe, T., Matsushita, Y., Tabata, T., Inoue, T., Kanazawa, Y., Ohkawa, K., Kasahara, A., Fusamoto, H. & Kamada, T. (1997). Relation of disease activity during chronic hepatitis C infection to complexity of hypervariable region 1 quasispecies. *Hepatology* **25**, 439-444.

Zein, N. N., Rakela, J., Krawitt, E. L., Reddy, R. K., Tominaga, T. & Persing, D. H. (1996). Hepatitis C virus genotypes in the United States: epidemiology, pathogenicity, and response to interferon therapy. *Annals of Internal Medicine* **125**, 634-639.

Zein, N. N. (2000). Clinical significance of hepatitis C virus genotypes. *Clinical Microbiology Reviews* **13**, 223-235.

Zeuzem, S., Lee, J. H. & Roth, W. K. (1997). Mutations in the nonstructural 5A gene of European hepatitis C virus isolates and response to interferon alfa. *Hepatology* **25**, 740-744.

Zeuzem, S., Feinman, S. V., Rasenack, J., Heathcote, E. J., Lai, M. Y., Gane, E., O'Grady, J., Reichen, J., Diago, M., Lin, A., Hoffman, J. & Brunda, M. J. (2000). Peginterferon Alfa-2a in Patients with Chronic Hepatitis C. *New England Journal of Medicine* **343**, 1666-1672.

Zhao, H., De, B. P., Das, T. & Banjeree, A. K. (1996). Inhibition of human parainfluenza virus-3 replication by interferon and human MxA. *Virology* **220**, 330-338.

Zhou, A., Paranjape, J. M., Der, S. D., Williams, B. R. & Silverman, R. H. (1999). Interferon action in triply deficient mice reveals the existence of alternative antiviral pathways. *Virology* **258**, 435-440.

Zhu, N., 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.

Zhu, S., Romano, P. R. & Wek, R. C. (1997). Ribosome Targeting of PKR Is Mediated by Two Double-stranded RNA-binding Domains and Facilitates in Vivo Phosphorylation of Eukaryotic Initiation Factor-2. *Journal of Biological Chemistry* **272**, 14434-14441.

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**, 653-61.