Characterisation of the translation efficiency and quasispecies composition of the 5'untranslated region of Hepatitis C Virus in genotype 1 and 3 infected patients

by

Morteza Motazakker, M.D.

A thesis presented for the degree of Doctor of Philosophy

in

The Faculty of Biomedical and Life Sciences
at the
University of Glasgow

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

May 2005

Acknowledgements

Firstly, I would like to thank Professor Richard Elliott for providing research facilities in the Institute of Virology.

I would like to thank my supervisor, Dr Liz McCruden for her valuable comments and encouragement throughout this project and especially for her critical reading of this thesis so promptly.

I am grateful to Dr Frazer Rixon, my second supervisor for his helpful advice and support.

I am grateful to Prof Richard Elliot for providing me pRL vector and some other reagents used in this project. I also wish to thank Dr John McLauchlan for providing the core product and anticore monoclonal antibody, and for his helpful discussions. I am grateful to Dr. Karin Oien for providing me with the liver biopsy scoring data. Many thanks to Gwen Gunning at the Regional Virus Laboratory for her help in quantitation of viral load using the COBAS system. Thanks Dr Rory Bowden, for his helpful comments on work presented in Chapter 6.

I wish to thank all past and present members of lab 311/301 for both technical assistance and for providing a friendly environment in which to work. Many thanks to Dr Petra Preiskchat for providing sequencing data and PCR products. Special thanks to Carol-Anne Smith for looking after the sera and for teaching me how to use the real time PCR system. Thanks Selma Erwa, for carrying out part of work described in chapter 7. Thanks Victoria McKechnie for providing me the HVR clones used in chapter 7. I also wish to thank the washroom and media staff who supplied the communal glassware and solutions used in this project

Many thanks to all my friends in the Institute and elsewhere in the university, especially Babak Sioofy and Rahim Rezaee who have made the bad days seem not so bad and the good days seem even better.

I have to thank my parents, for their continuous strength and support throughout the course of my study. A big thank you to all of the members of our family in Iran. My wife is the last in this list but deserves the biggest thanks for all her love, encouragement and making the most of our life together in Glasgow. My daughter, Denise for giving me the most enjoyable moments in my life.

The author was the recipient of Islamic republic of Iran student scholarship. I thank the ministry of Health and Medical Education for generous financial support. Unless otherwise, all work presented in this thesis was carried out by the author.

Morteza Motazakker May 2005

Contents

List of Figures and Tables

Abbreviations

5'UTR	5' untranslated region
3'UTR	3' untranslated region
%	percentage
°C	degree Centigrade
µg	microgram(s)
µl	microlitre(s)
A	adenosine
aa	amino acid(s)
ALT	alanine aminotransferase
AST	aspartate aminotransferase
APS	Ammonium persulphate
ARFP	alternate reading frame protein
as	antisense
ASGP-R	asialoglycoprotein receptor
ATP	(2'-deoxy)adenosine 5'-triphosphate
BHK	Baby hamster kidney cells
BHKsinT7	BHK cells expressing T7 RNA polymerase
bp	base pair
BSA	Bovine serum albumin
BVDV	bovine viral diarrhoea virus
C	cytosine
CAT	Chloramphenicol Acetyl Transferase
C-terminal	Carboxy-terminal
$CaCl_2$	Calcium chloride
cDNA	Complementary DNA
CIP	Calf intestinal phosphatase
cm	centimetre
CMV	cytomegalovirus
CO_2	Carbon dioxide
COS-1	African Green Monkey kidney cells
C_T	threshold cycle
CTL	cytotoxic T lymphocyte
(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
EBV	Epstein Barr virus
E.coli	*Escherichia coli*
ECL	Enhanced chemiluminescence
EDTA	ethlenediaminetetra-acetic acid (disodium salt)
eIF	Eukariotic initiation factor
EMCV	encephalomyocarditis virus

ER	Endoplasmic reticulum
EtBr	ethidium bromide
FCS	Foetal calf serum
FLuc	firefly luciferase
FMDV	Foot-and-mouth disease virus
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GBV-B	GB virus strain B
GMEM	Glasgow minimal Eagles Medium
GST	glutathione S-transferase
GT	genotype
(d)GTP	(2'-deoxy)guanosine 5'-triphosphate
G	guanosine
HAV	hepatitis A virus
HBV	hepatitis B virus
HCC	hepatocellular carcinoma
H_2O	Water
HCl	Hydrochloric acid
hGH	human growth hormone
hnRNPk	heteronuclear ribonucleoprotein kinase
HRP	Horseradish peroxidase
HoCV	hog cholera virus
HVR	hypervariable region
HuH7	Human hepatocellular carcinoma cell line
IFN	Interferon
IgG	Immunoglobulin G
IL	interleukin
IRES	internal ribosome entry site
IRF-3	IFN regulatory factor-3
ISDR	IFN sensitivity determining region
IU	International units
IVDU	intravenous drug user
kDa	kilodalton
kb	kilobase pair
KCl	Potassium chloride
KH_2PO_4	Potassium dihydrogen orthophosphate
KOAc	Potassium acetate
KOH	Potassium hydroxide
La	Lupus antigen protein
L-broth	Luria Bertani medium
LDL	Low density lipoprotein
LT-βR	lymphotoxin-β receptor
M	molar
mAb	monoclonal antibody
mg	miligram
$MgCl_2$	Magnesium chloride
$Mg(OAc)_2$	Magnesium acetate
$MgSO_4$	Magnesium sulphate
MHC	major histocompatibility complex
min	minute(s)
ml	millilitre(s)

mM	millimolar
mm	millimetre(s)
Molt4	human T lymphoblastic leukaemia cell
mRNA	messenger RNA
N-terminal	Amino-terminal
NaCl	Sodium chloride
Na_2HPO_4	di-sodium hydrogen orthophosphate anhydrous
NANBH	non-A, non-B hepatitis
NaOH	Sodium hydroxide
NBCS	New born calf serum
$(NH_4)_2SO_4$	Ammonium sulphate
ng	nanogram(s)
nm	nanometre(s)
NMR	Nuclear magnetic resonance
NS	non-structural
nt(s)	nucleotide(s)
(d)NTP	(deoxy)nucleotide triphosphate
OD	Optical density
ORF	open reading frame
PABP	polyadenylate binding protein
PAGE	Polyacrylamide gel electrophoresis
PBMC	peripheral blood mononuclear cells
PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
PCR 1	First round PCR
PCR 2	Second round PCR
PEG	polyethylene glycol
p.f.u.	plaque forming units
PKR	RNA-dependent protein kinase
pmoles	picomoles
P_n	the ratios of number of polymorphic sites to number of nucleotides sequenced
pRL	empty bicistronic vector (Collier et al., 1998)
pRLN	pRL vector containing 5'UTR region
pRLNC	pRL vector containing 5'UTR plus core region
PTB	Polyptrimidine tract binding protein
Rluc	renilla luciferase
RdRp	RNA dependent RNA polymerase
RFLP	Restriction fragment length polymorphism
RLU	Relative light unit
RNA	Ribonucleic acid
RNase	Ribonuclease
RNasin	Ribonuclease inhibitor
r.p.m	revolutions per minute
rRNA	ribosomal RNA
RT	Reverse transcription
RT-PCR	Reverse transcription PCR
s	sense
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis

SGB	Stacking gel buffer
SPP	signal peptic peptidase
ss	single stranded DNA
SSCP	single stranded conformation polymorphism
SL	stem loop
SV40	simian 40 virus
TAH	transfusion associated hepatitis
TBE	Tris-boric acid-EDTA buffer
TEMED	N,N,N'N-tetramethylethylenediamine
T	thymidine
TMB	Tetramethyl benzidine
TMD	transmembrane domain
TNF-R1	tumour necrosis factor-receptor 1
TP broth	Tryptose phosphate broth
tRNA	transfer RNA
(d)TTP	(2'-deoxy)thymidine 5'-triphosphate
U	unit
UTR	untranslated region
UV	ultra violet
U	uracil
V	volts
VLP	virus like particle
v/v	volume per volume
WHO	World Health Organisation
w/v	weight per volume

Reference HCV isolates:

Genotype 1a	H77	Accession no: AF011751
Genotype 3a	NZL1	Accession no: D17763

All patients studied in this thesis were infected with either genotype 1a or 3a but for simplicity are usually referred to as genotype 1 or 3 (GT1 or GT3).

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

Summary

Hepatitis C virus (HCV) infects over 170 million people worldwide. Chronic infection occurs in 50–80% of cases and eventually leads to cirrhosis and hepatocellular carcinoma. HCV can be classified into six genotypes. 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 patients infected with HCV genotype 1 and genotype 3. 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.

The HCV lifecycle is only partly understood owing to the lack of a productive cell culture system. There is no vaccine to prevent infection by HCV. Given the predicted future impact of the disease, there is a great need to understand the molecular basis of the HCV life cycle. Protein translation is one of the important processes in HCV replication. It involves an internal ribosome entry site (IRES) in the 5'untranslated region (5'UTR). Comparison of the sequence and function of the 5'UTR from different genotypes might differentiate features essential to the virus life cycle in all genotypes from those relevant only to individual genotypes. In this study, the 5'UTR region of genotype 3 was compared with that of genotype 1 with respect to translation initiation and quasispecies composition. The association between translation efficiencies, serum viral loads and the histology of the liver was also investigated.

The work presented in chapters 3 and 4 was undertaken to compare the translation efficiencies of genotypes 1 and 3 IRESs obtained from a number of infected patients. There are conflicting reports with respect to the translation efficiency of the 5'UTR in genotypes 1 and 3 (Buratti et al., 1997; Collier et al., 1998). These studies used only a single sequence as representative of its genotype. In this study, the system developed by Collier et al. (1998) in this Institute was used to measure the translation efficiency of a number of 5'UTR sequences matched with the relevant majority sequence from

genotype 1and 3 infected patients. In this system, the upstream reporter (renilla luciferase) is driven by the T7 promoter sequence and the IRES sequences are inserted before the downstream firefly luciferase reporter gene. Sixteen constructs (8 genotype 1 and 8 genotype 3) containing different 5'UTR sequences from our patient cohort were constructed and their activities were measured in two different cell lines, BHKsinT7 and HuH7. Differences in the 5'UTR nucleotide sequence resulted in variations in IRES activity. Translation efficiencies differed in the two cell lines tested. Mean translational activities of genotype 1 isolates were statistically significantly higher than genotype 3 in BHKsinT7 cells in 3 out of 5 experiments but further consideration of the variation in the raw data led us to the conclusion that the differences may not be biologically relevant. We proceeded to repeat the experiments in HuH7 cells in which the differences in mean translational levels between genotypes were shown not to be significant.

To compare the translation efficiencies of the 5'UTR combined with the putative core encoding region between genotype 1 and 3 patients, an attempt was made to construct clones matched with the majority sequence obtained from patients. Most previous reports suggested that the presence of core resulted in downregulation of HCV IRES activity. Unfortunately, no completely matched clone was identified. Therefore, the translation efficiency of three random clones from each of 6 patients was measured. The results suggested that the presence of core region did not downregulate the translation efficiency of 5'UTR as suggested previously. It was also shown that IRES activities of genotypes 1 and 3 constructs did not differ significantly.

The work described in chapter 5 was carried out to investigate the association between IRES activity, serum viral load and the histopathological appearance of liver biopsies from the patients. Serum viral loads were measured using real-time PCR (Taqman) and relevant liver biopsies scored using the Ishak scoring system. The viral loads ranged from 10^4 to 10^7 IU/ml. Statistical analyses of these data found no association between translation efficiency, viral load and liver histology. However, a strong correlation was observed between the presence of steatosis in liver biopsies taken from genotype 3 infected patients compared to genotype 1.

The work presented in chapters 6 and 7 was undertaken in order to determine the quasispecies composition of 5'UTR sequences in serum and matched liver samples from 6 HCV infected patients using 2 different methods including cloning and single stranded conformation polymorphism (SSCP). Previous work carried out by P. Preikschat in our laboratory showed the presence of an identical majority sequence in paired serum and liver samples from 26 patients (11 genotype 1 and 15 genotype 3). It was hypothesized that a minor population of virus with low translational activity could replicate in liver cells in order to escape from immune surveillance and result in virus persistence. The majority sequence data did not either support or refute this hypothesis. Therefore, amplified 5'UTR products from matched serum and liver samples were cloned and sequenced in both directions (chapter 6). It was shown that the majority of clones in serum and liver were identical to the majority sequence obtained from each patient. In addition, single, distinct variants were observed suggesting the presence of a quasispecies in 5'UTR. However, there is a limit to the number of clones which can be sequenced and nucleotide misincorporation due to *Taq* polymerase and reverse transcription errors will occur. In order to address whether any of the clones detected by the above method were present as more than 2.5% of the quasispecies, SSCP analysis was performed (chapter 7). An identical SSCP profile was observed in serum and liver samples from each patient. Taken together, the results obtained from cloning and SSCP analysis of 5'UTR region from matched serum and liver samples did not support the hypothesis that a minor subpopulation of virus in liver cells exists which translates the polyprotein with a lower efficiency.

Chapter 1

Introduction

1.1. Historical background

1.1.1. Non-A, non-B hepatitis

The discovery of the hepatitis B surface antigen (HBsAg) in the mid 1960's led to the introduction of serological tests to screen blood donors for hepatitis B infection. Hepatitis B virus (HBV) was the major cause of post-transfusion hepatitis prior to introduction of screening for HBsAg. The introduction of screening tests for HBV in 1970 reduced transfusion associated hepatitis (TAH) by approximately 80%. Despite the removal of all HBsAg-positive blood donations from the supply, transfusion associated hepatitis persisted. Later, Feinstone *et al.* (1975) discovered the hepatitis A virus (HAV) using immunoelectron microscopy on faecal samples from volunteers infected with this virus. Tests to detect IgM and IgG antibodies were then introduced and it became possible to screen non-HBV cases for HAV. It became apparent that another blood-borne agent was responsible and these cases were defined as having non-A, non-B hepatitis (NANBH)(Alter et al., 1975).

In 1978, two studies by Alter *et al.* (1978) and Tabor *et al.* (1978) reported the transmission of the NANBH agent to chimpanzees. Alter *et al.* administered plasma obtained from patients with NANBH to chimpanzees intravenously. 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) extracted samples of human plasma which contained infectious doses of NANBH virus with chloroform which is a potent lipid solvent, and showed that the extract was non-infectious when tested in chimpanzees. This suggested that the agent might be a virus with a lipid envelope.

Subsequently, He *et al.* (1987) determined the approximate size of one strain (H) of the NANBH viral agent by filtering it through polycarbonate membranes. The study indicated that strain H was 30 to 60 nm in diameter.

1.1.2. Cloning of the NANB hepatitis infectious agent

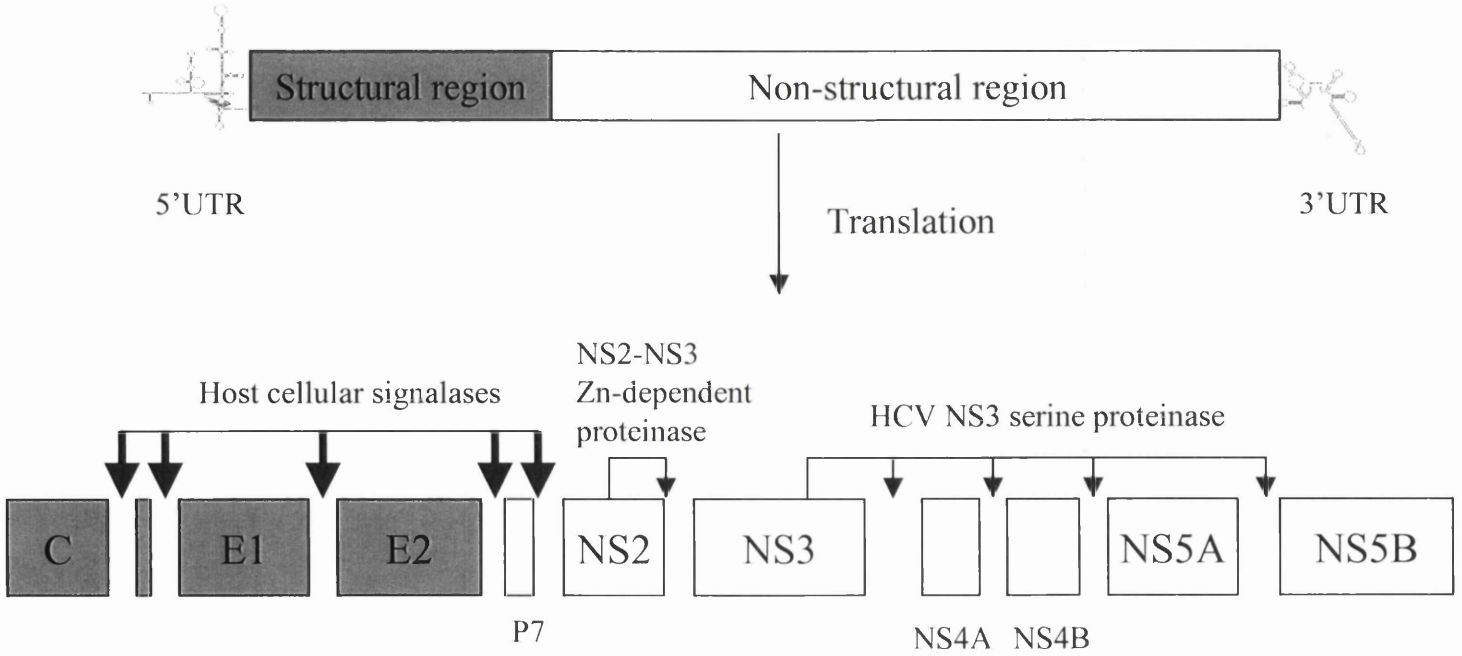
After more than a decade, attempts at identifying the NANBH agent by conventional immunological methods failed, possibly because of insufficient concentrations of viral antigen. Houghton *et al.* using their previous experience in cloning of the genome of hepatitis D virus, used molecular biological approaches to clone and express the genome (Choo et al., 1989) and then to develop immunoassays to detect antibody to the protein products of these clones (Kuo et al., 1989). In order to obtain infectious material with a high concentration of agent, they pooled large volumes of plasmapheresis samples from a chronically infected chimpanzee which had been shown to have an unusually high titre of the presumed virus. The plasma was centrifuged until a pellet was obtained. Because it was not known whether the virus had a DNA or a RNA genome, random primers were used to synthesise cDNA, which was cloned into the cloning vector λgt11 and expressed in *Escherichia coli* (*E.coli*). This allowed the amplification of the cDNA and the proteins to be expressed. After the lysis of the bacteria, expressed proteins were screened with serum from a patient who had NANBH as a presumed source of antibodies and then with radiolabeled antiglobulin. After screening approximately 10^6 recombinants, a single positive clone, 5-1-1, was identified. To investigate the origin of clone 5-1-1, a larger

clone, clone 81, was isolated from the same cDNA library using clone 5-1-1 as a hybridisation probe. Clone 81 did not hybridise to control human DNA or RNA derived from non-NANBH infected chimpanzees, however it did hybridise to RNA extracted from the liver tissue of an NANBH infected chimpanzee. This activity was abolished after treatment with ribonuclease but not deoxyribonuclease. The investigators concluded that the clones 5-1-1 and 81 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) presumed to be the genome of the virus. The virus responsible for NANB agent was renamed hepatitis C virus (HCV). The cloning of the HCV led to the development of immunoassays for detection of HCV antibodies (Kuo *et al.*, 1989). Using the original clone 5-1-1 as a probe, three overlapping clones were isolated and ligated together to build clone c100 consisting of part of NS4 region of the HCV genome. This clone was expressed in yeast and the resultant polypeptide was used to capture HCV antibodies from serum samples.

1.1.3. Virus morphology

Before the identification of HCV in 1989, Feinstone *et al.* (1983) reported inactivation of the non-A, non-B hepatitis agent when infectious material was treated with chloroform suggesting that the particles were surrounded by a lipid envelope. Microfiltration experiments demonstrated that this agent had a diameter of less than 80 nm (Bradley et al., 1990). Further filtration experiments by He *et al.* (1987) suggested a diameter of 30-60 nm. After identification of HCV, immunogold electron microscopy studies have shown that HCV is a spherical particle with a lipid membrane envelope containing projections of 6 nm. These particles were found in 1.14 to 1.16 g/ml fractions after sucrose density gradient centrifugation (Kaito et al., 1994). Similarly, Shimizu *et al.* (1996) detected virus-like particles with a diameter of approximately 50 nm 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

Figure 1.1. A schematic representation of the HCV genome and the individual protein products of the polyprotein.(modified from Pawlotsky, 2004).



low density lipoproteins (LDL) (Hijikata et al., 1993c). The higher-density fractions (~1.1 g/ml) possibly represent free virus or particles complexed with immunoglobulin (Choo et al., 1995) and correlated with lower infectivity.

1.1.4. Classification of HCV

The HCV belongs to the *Flaviviridae* family, and is a member of the genus hepaciviruses (Robertson et al., 1998). The other genera in the family are the pestiviruses and the flaviviruses. All viruses in the *Flaviviridea* family have enveloped virions with a positive sense RNA genome which is translated into a single long polyprotein with genes encoding structural protein at the N-terminal and those encoding non-structural proteins at the C-terminal ends of the genome. Individual proteins are produced by cleavage of the polyprotein by host and viral proteases. Comparative sequence analysis of the HCV genome revealed that HCV has a genetic organisation and polyprotein structure similar to the pestiviruses and, to a lesser extent, to the flaviviruses (Choo et al., 1991). There is limited amino acid similarity among these viruses including serine protease, and nucleotide triphosphatase (NTPase) domains of NS3 and the NS5B (Miller and Purcell, 1990). However, despite these similarities, pestiviruses and HCV differ significantly in amino acid sequence. Therefore, a third genus of the *Flaviviridae* family, the hepaciviruses, was proposed for classification of HCV (Robertson et al., 1998).

1.2. The HCV genome and its gene products

1.2.1. 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 has been shown 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., 1991). 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.

1.2.2. Core

Core is the first protein encoded by the HCV open reading frame. Three species of core have been identified. p23 is a 191 amino acid product which contains the signal sequence which directs E1 to the endoplasmic reticulum (ER). Cleavage of the polyprotein occurs between residues 191 and 192 to generate the N-terminal end of E1 (Hijikata et al., 1991). The presumed mature form of core is produced by an additional cleavage between residues 174 and 191 to give a 21 kDa species (p21) (Hussy et al., 1996). Following expression in mammalian cells, p21 is the major form of core detected (Moradpour et al., 1996). A third core-related species, termed p16, has also been detected in studies using HCV-1, the prototype strain of the virus (Lo et al., 1994). Production of this species by *in vitro* translation does not require the presence of membranes and the protein has the same *N*-terminus as full-length core polypeptide. In the absence of E1 coding sequences, p16 is the predominant form of core produced *in vitro* and in transfected cells (Lo et al., 1994). Core is highly conserved between the 6 genotypes and the hydrophobicity profile of core identifies 3 domains within the protein (Hope and McLauchlan, 2000). Domain 1 (aa 1-122) contains clusters of lysine and arginine 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 is the segment removed during processing of core that acts as the signal sequence for E1. It has been shown that, after cleavage by signal peptidase, the signal peptide is further processed by the intramembrane-cleaving protease signal peptide

peptidase (SPP) which promotes the release of core protein from the ER membrane (McLauchlan et al., 2002). This processing event allows the core to be free for subsequent trafficking to lipid droplets. Immunoelectron microscopy of expressing cell lines showed that the core protein was located along membranes outside the ER cisternae and on the surface of cytoplasmic lipid droplets (Barba et al., 1997; Moradpour et al., 1996). Domain II of core was reported to be essential for lipid droplet association (Hope et al., 2002). This region in HCV core is present in the corresponding protein of GBV-B. It is possible that interaction between core protein and lipid droplets is responsible for a common condition seen in HCV infected patients known as steatosis (McLauchlan, 2000).

Core is presumed to form the capsid shell of the virus 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. The core protein present in virus like particles (VLPs) produced in insect cells transfected with a baculovirus expressing HCV structural proteins is similar in size to that expressed from a vaccinia recombinant in mammalian cells (Baumert et al., 1998). Because VLPs are not released from infected cells, complete virus assembly cannot be analysed.

It has been shown that core is an RNA-binding protein (Santolini et al., 1994). The RNA binding region maps to the N-terminal 75 residues within domain I. The virus like particles produced by the baculovirus system have also been shown to encapsidate positive-sense HCV RNA (Baumert et al., 1998). It has been suggested that core protein could 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 indicated the presence of an interaction between core and E1 but not E2 (Lo et al., 1996). A more recent study reported the involvement of the first domain of E1 in an interaction with core protein (Ma et al., 2002).

There are several cellular proteins with which core can interact. The region of core that associates with these proteins is located in domain I of the p21 species (McLauchlan,

2000). 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 (Owsianka and Patel, 1999). The interaction of core with TNF-R1 and LT-βR, both of which are known to be involved in apoptosis, may be indicative of an effect of core protein on apoptosis (Ruggieri et al., 1997). In contrast, it has been reported that expression of core in a human B cell line did not modify the main apoptosis pathways (Giannini et al., 2002). Core has also been implicated in cellular transformation and the development of hepatocellular carcinoma (Moriya et al., 1998), and transcriptional regulation (McLauchlan, 2000). It has been reported that core protein affects T cell responses by modulation of IL2 which may contribute to the persistence of HCV infection (Bergqvist et al., 2003). It seems that core has multiple effects on cells. However, these effects need to be confirmed in a model system which reproduces viral replication and the pathological conditions seen in humans. Many of these studies typically rely on systems which tend to produce high levels of protein. There are many conflicting data. In addition, some of differences may be attributed to core products derived from different HCV strains and different length of core protein have been expressed.

An additional HCV protein, the F protein or ARFP (alternate reading frame protein) of 17-20 kDa has been identified to be generated as a result of a +1 frameshift in the N-terminal core coding region of genotype 1a (Varaklioti et al., 2002; Xu et al., 2001). The frame shift site is located in the A rich sequence at amino acids 10-12 of the polyprotein. The F protein is very unstable and short lived (Roussel et al., 2003) and primarily associated with the ER (Xu et al., 2003).

1.2.3. E1 and E2

HCV encodes two major glycoproteins, E1 and E2, which are believed to be exposed on the surface of the virus. The genes encoding these 2 putative envelope proteins are located downstream of the core protein and the proteins have molecular weights of approximately 31 and 70 kDa respectively (Grakoui et al., 1993b).

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., 1991). Cleavage at the C-terminus of E2 is not 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).

Deglycosylation studies of HCV envelope proteins have shown that these proteins are highly modified by N-linked glycans (Dubuisson et al., 1994). E1 and E2 possess up to 6 and 11 potential glycosylation sites, respectively (Goffard and Dubuisson, 2003). It has been shown that the glycosylation of E1 occurs post-translationally and is improved by the coexpression of E2 in *cis* (Goffard and Dubuisson, 2003).

Hydrophobic domains have been identified in the C-termini of the HCV glycoproteins which act as membrane anchors. Deletion of the hydrophobic domain results in secretion of the protein and this allows the transmembrane domain (TMD) to be mapped to specific residues. 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). 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.

Keck *et al.* (2004) identified an human monoclonal antibody (H-111) against a highly conserved region near the N terminus of E1 which was able to interact with E1 genotypes 1a, 1b, 2b, and 3a. It blocked binding of HCV-like particles to infected target cells, suggesting the involvement of this epitope in virus binding and entry.

Hypervariable regions have been identified in the E2 envelope glycoprotein sequence (Kato, 2001; Weiner et al., 1991b). The first 27 amino acids of the E2 form the HVR1. The biological role of HVR1 is unknown. It has been reported that the anti-HVR1 antiserum induced protection against homologous HCV infection in chimpanzees but not against the emergence of neutralization escape mutants (Farci et al., 1996). Penin *et al.* (2001) showed that, despite the sequence variability of HVR1, the physicochemical properties of the residues at each position and the overall conformation of the HVR1 are highly conserved among the various genotypes. Another hypervariable region, HVR2, has been described in the E2 glycoprotein of HCV genotype 1b strains. HVR2 is a stretch of 7 amino acids (positions 91-97) showing up to 100% sequence diversity (Kato, 2001).

1.2.4. p7

p7 is a small protein of 63 aa located between E2 and NS2 in the polyprotein (Lin et al., 1994). 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 incomplete leading to the production of both fully processed E2 and uncleaved E2-p7 (Lin et al., 1994). The p7 polypeptide has a double membrane spanning topology with both N- and C-terminals oriented towards the ER lumen (Lin et al., 1994). In addition, the C-terminal transmembrane domain of p7 has a signal sequence function. The export of a fraction of p7 at the plasma membrane suggests that this polypeptide might have a functional role in several compartments of the secretory pathway (Carrere-Kremer et al., 2002). It has been demonstrated that core, E1 and E2 without p7 are sufficient for recombinant VLP formation in the insect cell expression system (Baumert et al., 1998). Very recent data indicate that the expressed

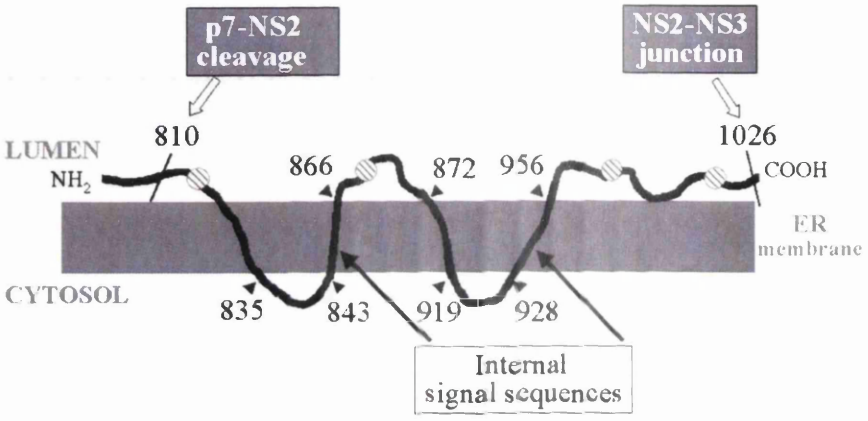


Figure 1.2. A predicted model of the NS2 membrane topology.

The numbers represent the amino acids identified by the deletion mapping experiments as possible locations for TMDs. These amino acids do not represent the boundaries of the transmembrane domains, but rather, they predict the possible locations of transmembrane domains. Thus, four transmembrane domains are predicted to reside within amino acids 810-835, 843-866, 872-919, and 928-956. The striped circles represent the introduced glycosylation sites that were used. Arrows indicate the locations of the two internal signal sequences (taken from Yamaga & Ou, 2002).

form of p7 formed hexamers and functioned as a calcium channel in lipid membranes. This activity could abrogated by the antiviral drug Amantadine (Griffin et al., 2003) .

1.2.5. NS2

The mature NS2 is a hydrophobic transmembrane protein with molecular mass of 23 kDa (reviewed by Kato, 2001). The function of NS2 in the virus life cycle is unclear. Its deletion did not abolish the replication of HCV replicons in cell culture, indicating that it is not required for viral replication (Blight et al., 2000). It has been shown that membrane association of NS2 is P7-independent and occurs co-translationally suggesting the presence of internal signal sequences in NS2. Four putative transmembrane domains (TMD) with both the N- and C-terminals in the ER lumen have been identified (Yamaga and Ou, 2002) (Figure 1.2). The C-terminal end of NS2 protein together with the N-terminal third of NS3 has proteolytic activity that is responsible for cleavage at the NS2/NS3 junction. Kolykhalov *et at* (2000) showed that point mutation in the NS2/3 region can abolish HCV infectivity in chimpanzee. Cleavage at the NS2/3 site is catalysed by zinc. This activity is inhibited by metal chelators such as EDTA (Hijikata et al., 1993a). NS2 has also been shown to be involved in regulation of NS5A phosphorylation (Liu et al., 1999).

1.2.6. NS3

The NS3 protein (about 70 kDa) is a multifunctional protein (Grakoui et al., 1993a). The protease activity is localised in the N-terminal 180 amino acids (Bartenschlager et al., 1994) and the C-terminal 450 amino acids encode the helicase and nucleotide triphosphatase activity (Kim et al., 1995). 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 sites. NS3 has no membrane anchor, but it forms a noncovalent complex with NS4A which is membrane-anchored (Tanji et al., 1995a; Wolk et al., 2000). It has been shown that expression of NS3 in the absence of the NS4A cofactor was diffusely distributed in the cytoplasm and nucleus. The co-expression of NS4A, however, directed NS3 to the ER

or an ER-like modified compartment (Wolk et al., 2000). The serine protease activity is required for infectivity of the HCV genome in chimpanzee (Kolykhalov et al., 2000). 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). The crystal structure of the NS3 serine protease (Yan et al., 1998) and RNA helicase (Cho et al., 1998) domains have recently been published which suggests the presence of three distinct domains which are separated by deep clefts forming a Y shaped structure (Penin et al., 2004).

It has been shown that the NS3 helicase was able to unwind RNA and DNA homo and heteroduplexes in a 3' to 5' direction (Tai et al., 1996) and this required Mg^{2+} or Mn^{2+} and ATP (Jin and Peterson, 1995). The NTPase activity is stimulated by double or single stranded RNA and DNA particularly by poly (U). The minimal requirement for both NTPase and helicase activities has been mapped to the C-terminal 466 amino acids of NS3 (Kim et al., 1997) and mutations in either the NTPase or the helicase motifs affect both functions (Kim et al., 1997). It has been shown that the helicase activity is required for viral replication in chimpanzees. 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 (Kolykhalov et al., 2000).

1.2.7. NS4A and NS4B

The NS4 region of the polyprotein contains two proteins, NS4A and NS4B. NS4A consists of 54 aa and is 8 kDa in size. It acts as cofactor for the NS3 serine protease for efficient processing at NS3/4A, NS4A/B and NS4B/5A and stimulates cleavage at the NS4A/NS4B site. It has been shown that the N-terminal hydrophobic domain of NS4A mediated localisation of NS3 to the ER membrane (Wolk et al., 2000). It has also been reported that NS4A may bind directly to NS5B (with or without NS3) and enhance its polymerase activity (Ishido et al., 1998). 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 and Bartenschlager, 1999). A recent study reported an interaction between NS4A and NS2, although the functional role of this complex is unknown (Flajolet et al., 2000).

The NS4B protein has a molecular weight of approximately 27 kDa. There is little information available about the function of this protein. This hydrophobic protein is ER membrane associated and displays properties of a cytoplasmically orientated integral membrane protein (Lundin et al., 2003). Computer predictions of the membrane topology of NS4B suggested that it has four transmembrane segments (Lundin et al., 2003). It was recently reported that expression of NS4B induces the formation of a ER-derived membranous web that harbors all HCV structural and non-structural proteins (Egger et al., 2002) as well as replicating viral RNA (Gosert et al., 2003). It has been shown that expression of NS4A/B slows the rate of ER-to-Golgi traffic. A reduction in protein secretion rate was observed for several proteins during NS4A/B expression in particular for major histocompatibility complex class I (MHC-I) molecules in the presence of a full-length HCV replicon (Konan et al., 2003; Lundin et al., 2003)

1.2.8. NS5A

Expression of the NS5A gene produces two proteins of 56 and 58 kDa in size. It is found in a basally phosphorylated form of 56 kDa and in a hyperphosphorylated form of 58 kDa (Kaneko et al., 1994; Tanji 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, the production of p58 needs NS3, NS4A and NS4b expression in *cis* (Neddermann et al., 1999). The work carried out by Koch *et al* (1999) suggested that a continuous NS3-5A sequence is required for NS5A hyperphosphorylation. Mutations at various positions in the NS3-4B region, not affecting polyprotein processing, could reduce or enhance this NS5A modification. It was concluded that structural integrity of each of these proteins, forming a protein complex, is

essential for differential phosphorylation of NS5A. The exact function of NS5A is not clear. Adaptive mutations have been found to cluster in the central region of NS5A in the replicon system (Blight et al., 2000) suggesting that NS5A is involved in the viral replication process either directly or by interaction with cellular proteins.

NS5A has been shown to form a multisubunit complex with NS3, NS4A and NS4B (Macdonald and Harris, 2004) and NS5A has been co-precipitated with the remaining two non-structural proteins, NS2 and NS5B. This suggests that the HCV non-structural 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 intracellular membranes and NS5A is localised in the cytoplasmic membrane surrounding the nucleus (Tanji et al., 1995b).

There are controversial reports suggesting that NS5A isolated from certain genotypes can bind to IFN-induced double-stranded RNA-dependent protein kinase (PKR) (Gale et al., 1998; Gale et al., 1997), via a 40 amino acid stretch called the interferon sensitivity-determining region (ISDR). It has been reported that certain amino acid alterations or "mutations" in this region appears to have a correlation with response to IFN therapy in patients infected with HCV genotype 1b (Enomoto et al., 1995). However, other studies (Aizaki et al., 2000) failed to find a correlation between ISDR sequence and ability to inhibit IFN activity. Subsequently it has been suggested that sequences in the C terminus of NS5A are also required to inhibit IFN activity (Nousbaum et al., 2000). It has been shown that deletion of the ISDR did not affect the IFN sensitivity of HCV replicons (Blight et al., 2000). Polyak *et al.* (2001) demonstrated that NS5A expression in human cells induced the IL8 RNA and protein and this effect correlated with inhibition of the antiviral effects of IFN in an *in vitro* assays suggesting a further mechanism for inhibition of this antiviral pathway. Co-localisation of NS5A with core protein at the surface of lipid droplets has been shown (Shi et al., 2002), although the significance of this interaction has not been clarified.

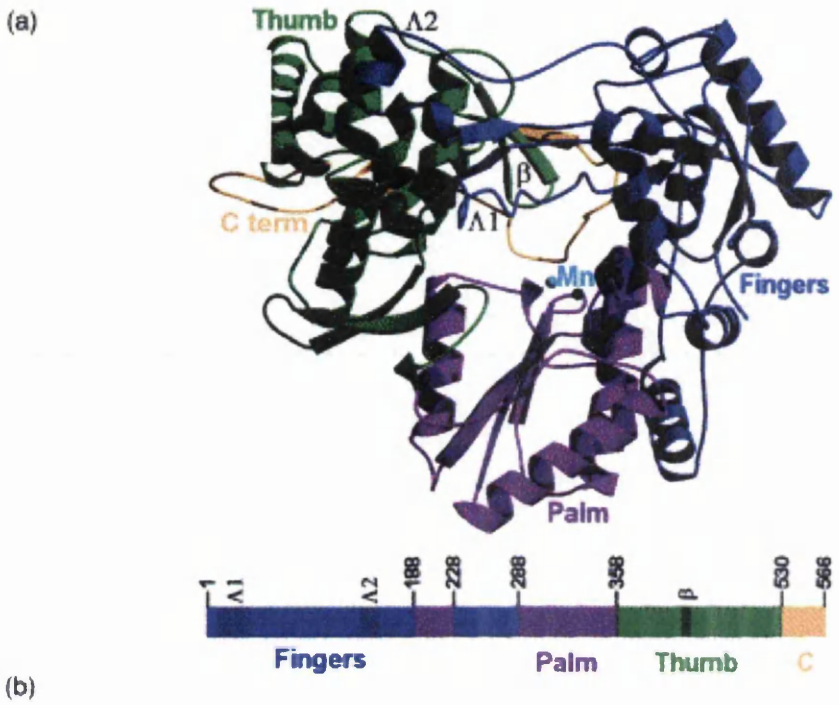


Figure 1.3. Crystal structure of the HCV NS5B polymerase.

Ribbon diagram of the NS5B protein of HCV genotype 1b. Fingers (blue), Palm (purple) and Thumb (green) subdomains are marked, as is the C-terminal Arm (yellow) with appropriate aa domain boundaries noted on the bar below. Subdomains and structural motifs (Λ1, Λ2 and ß) are also labelled. (taken from O'Farrell *et al.*, 2003).

1.2.9. NS5B

NS5B is 65 kDa in size and forms the C-terminus of the HCV polyprotein. It has been identified as an RNA dependent RNA polymerase (RdRp) (Behrens et al., 1996). It contains a GDD motif (Gly-Asp-Asp), residues 2737 to 2739, which is conserved in all RNA polymerases, and is essential for polymerase activity. Several crystal forms of NS5B have been produced from both HCV genotype 1b, BK strain (Bressanelli et al., 1999; Lesburg et al., 1999) and the HCV genotype 1b, J4 strain (O'Farrell et al., 2003). The crystal structure analysis of NS5B revealed the presence of the typical right-handed "fingers-palm-thumb" structure, with the RNA-binding tunnel lying between the "fingers" and "thumb" (Bressanelli et al., 1999) (Figure 1.3). It has been reported that NS5B interacted with NS3 and NS4A to form a replication complex and this complex was localised to the endoplasmic reticulum (Ishido et al., 1998). NS5B expressed alone was also associated with intracellular membranes (Hwang et al., 1997) so it is likely that HCV replicates on intracellular membranes. It has been demonstrated that a recombinant full-length NS5B is capable of copying the full-length HCV RNA genome *in vitro*, without the need for addition of other factors, 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 HCV RNA and 3' end of minus strand as templates for RNA synthesis (Oh et al., 1999). This specificity for the 3'X tail is suggested to be due to the recognition of specific stem-loop structures (SL2 and SL4) in the 3'UTR (Cheng et al., 1999).

1.3. HCV untranslated regions

1.3.1. 5'UTR

The 5' UTR is the most conserved region of the HCV genome (Bukh et al., 1992) reflecting its importance in both viral replication and translation. Overall, the 5'UTR from different HCV strains share over 85% nucleotide sequence identity (Smith et al., 1995b). This 341-342 nt region is much longer than the 5' UTRs of flaviviruses, which have an average length of 100 nts but is similar to that of pestiviruses such as bovine viral

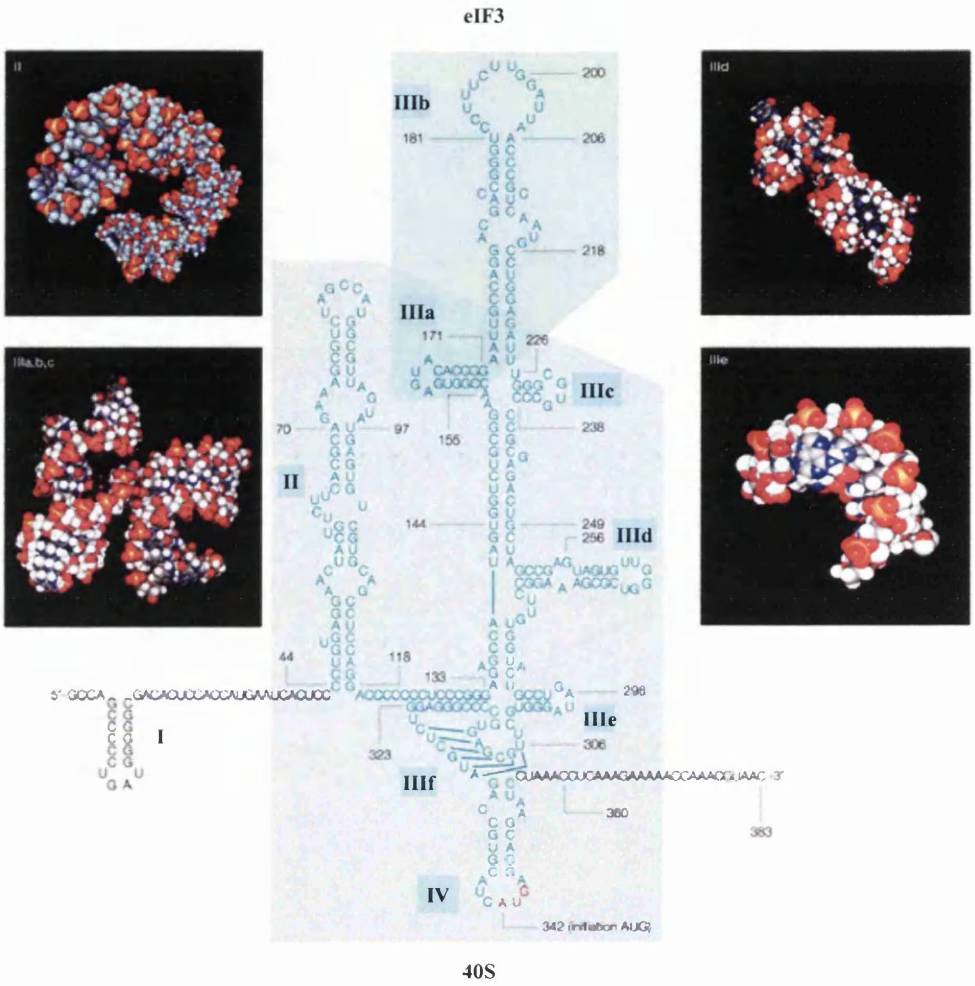


Figure 1.4. Sequence and proposed secondary structure of the HCV IRES.

The sequence and secondary structure of the 5 UTR is shown which contains the internal ribosome-entry site (IRES), including the four domains (I, II, III and IV) and subdomains of domain III (a, b, c, d, e and f). The IRES sequence is indicated in blue, with the location of the start codon (AUG) indicated in red. Numbers refer to nucleotide positions. The binding sites of eukaryotic translation-initiation factor 3 (eIF3) and the 40S ribosomal subunit are indicated in boxes. Solution structures of domain II and subdomains IIIa,b,c, IIId and IIIe are also depicted (taken from Tan et al., 2002).

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 structures. The first model of the secondary structure of the 5'UTR of HCV was proposed by Brown *et al.* (1992). This model was based on a comparative analysis of the sequences of multiple strains of HCV and members of the genus pestivirus such as bovine viral diarrhoea virus (BVDV) and hog cholera virus (HoCV). This model was modified by Wang *et al.* (1995) following the identification of a pseudoknot within the 5'NTR which is required for translation, and it was further refined by Honda *et al.* (1996a). Later, the same group (Honda et al., 1999) presented a different prediction of the secondary structure of domain II of the HCV 5'NTR based on a comparison of the HCV sequence with that of a newly discovered, hepatotropic tamarin virus, GB virus B. Kieft and co workers (1999) provided the first evidence that the IRES folds into a distinct three-dimensional structure at physiological salt concentrations using enzymatic and probing techniques. Point mutations that prevent folding of one or more domains disrupt IRES translation initiation activity, providing strong evidence that HCV IRES function requires formation of this tertiary structure.

The proposed HCV 5' UTR contains four major structural domains designated I to IV with a pseudoknot structure upstream of the initiation codon (Figure 1.4). The most 5' structure appears to be a small stem-loop spanning nucleotides 1-22. This segment of the 5'UTR has the lowest level of sequence homology between HCV, the pestiviruses and GBV-B. The role of 5' end in IRES activity will be discussed later.

Domain II extends from nt 44 to nt 118 and consists of multiple stems and loops. Many of the nucleotides in the loop regions of this domain are identical in each of these viruses (HCV, pestiviruses, GBV-B) in contrast to the majority of nucleotides present in base-paired regions. Mutagenesis studies have shown that the structural integrity of this domain is crucial for efficient IRES-directed translation (Kalliampakou et al., 2002; Odreman-Macchioli et al., 2001) but deletion of this domain does not lead to complete loss of activity (Tsukiyama-Kohara et al., 1992). Sequence analysis studies have shown that the unpaired regions of domain II contain conserved nucleotide motifs which are also present

in the corresponding IRES sequences from GBV-B and pestiviruses (Honda et al., 1999). These include nucleotides 81- 85 (apical loop) and 71-73 and 92 -96 (adjacent bulge). A recent report using cryo-electron microscopy suggested that domain II is responsible for the induction of a conformational change in the 40S subunit which could play an important role in translation initiation by holding the HCV encoding RNA in the decoding site of the ribosome in position until the translational machinery is correctly assembled (Spahn et al., 2001).

Domain III is the largest RNA structure within the 5'UTR and consists of several essential elements, including a large four-way junction (IIIabc)(Kieft et al., 2001) and the smaller stem-loop structures IIId, IIIe and IIIf. A pseudoknot structure in the HCV IRES was shown to be required for IRES activity, as mutations that destabilized tertiary interactions between residues of loop IIIf with complementary residues in domain IV were accompanied by a marked reduction in translation initiation (Wang et al., 1995). A structural element containing stem–loops IIIa, b and c facilitated binding of eIF3 (Kieft et al., 2001). It has been reported that subdomains IIId, e and f, together with the pseudoknot structural element, constituted the binding site for the 40S ribosomal subunit (Spahn et al., 2001). Nuclear magnetic resonance (NMR) analysis has suggested that stem-loop IIId folded into a loop which contains rich hydrogen-bonding potential for the formation of RNA-protein or RNA-RNA interactions (Lukavsky et al., 2000). The three guanosines in the loop are required for full IRES activity. Mutation of the three loop guanosines to cytosines had been previously shown to be deleterious to IRES activity *in vitro* (Kieft et al., 1999). In addition, the sequence of the loop of domain IIId is absolutely conserved among all HCV isolates. The domain IIId hairpin loop clearly plays an important role in the IRES–40S subunit interaction. Stem-loop IIIe folds into a tetraloop structure suspected to have a role in the initial assembly of IRES-40S complexes.

Domain IV contains the AUG initiation codon. This stem loop structure has been predicted to exist only in HCV and GBV-B (Smith et al., 1995b). The stem-loop may play a key role in regulating the initiation of translation. It has been reported that the stability of the stem-loop is inversely correlated with the efficiency of internal initiation of translation

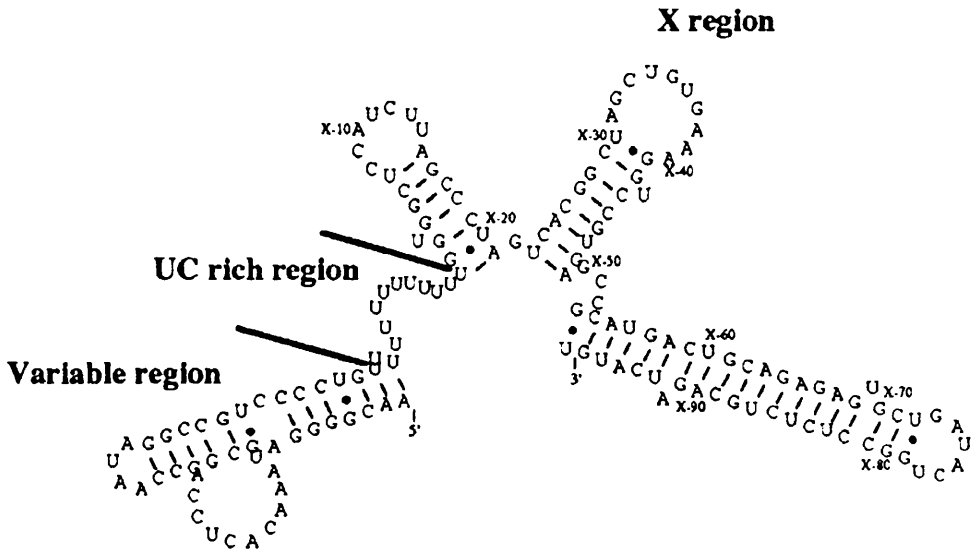


Figure 1.5. Predicted model of the secondary structure of the HCV 3'UTR RNA. (Taken from Ito & Lai, 1997).

(Honda et al., 1996a). There has been some controversy as to whether HCV sequences located downstream from the initiation codon influence the efficiency of IRES-mediated initiation. This will be discussed later.

1.3.2. 3'UTR

The predicted structure of the 3' UTR contains three distinct regions: a variable region of approximately 40 nucleotides in length, a poly (U/UC) tract of variable length and a highly conserved sequence of 98 nt called the 3'X region (Tanaka et al., 1996). Computer predictions of the secondary structure show that the 3'UTR can form stable stem-loops (Blight and Rice, 1997) as shown in Figure 1.5. The importance of the 3'UTR for *in vivo* replication of HCV has been confirmed in a chimpanzee model (Yanagi et al., 1997). By injecting RNA of infectious clones with various deletion in the 3'UTR, it has been shown that the 3'X tail and the poly (U) region were both required for infectivity of HCV in chimpanzees, whereas the 5' end of the variable region is not (Kolykhalov et al., 2000). Viral proteins may bind to 3'UTR. The NS3 protein was shown to bind to the poly (U) rich region and possibly to unwind the RNA secondary structure through its helicase activity (Kanai et al., 1995). It has been shown that NS5B bound to both the U-rich and 3'X regions in the HCV 3'UTR (Oh et al., 2000). There are several cellular proteins that have been reported to interact with the 3'UTR. So far, the following proteins have been identified: Polypyrimidine tract-binding protein (PTB) which bound to both the poly (U) tract and the 3'X region (Tsuchihara et al., 1997) , heterogeneous nuclear ribonucleoprotein C (hnRNP C) (Spangberg et al., 2000) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Petrik et al., 1999). La protein (Spangberg et al., 1999) also bound the 3'UTR and inhibited premature degradation of viral mRNA (Spangberg et al., 2001). The 5'UTR has been reported to bind PTB, suggesting a possible interaction between the 5' and 3'UTRs mediated by PTB. The functions of GAPDH, hnRNP C and ribosomal proteins in the HCV replication remain unknown. The role of the 3'UTR in translation of the HCV polyprotein will be discussed later.

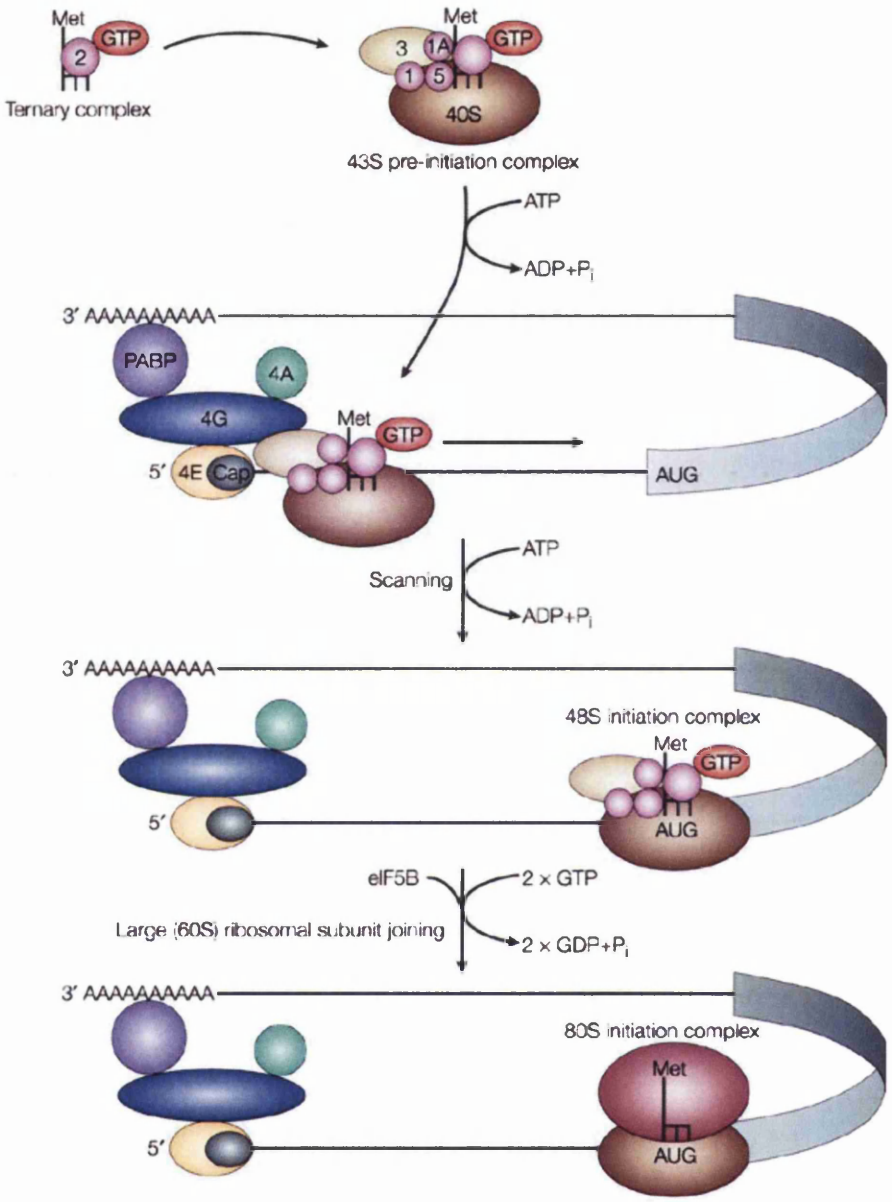


Figure1.6. Schematic of proposed mechanism of translation initiation by cap-dependent mechanism.

Eukaryotic initiation factors (eIFs) are depicted as coloured, numbered shapes in the figure.(taken from Gebauer, 2004)

1.4. Translation

Protein synthesis takes place on ribosomes, large ribonucleoprotein assemblies of approximately 4 MDa acting in association with a number of accessory factors to 'translate' the genetic information contained in messenger RNA (mRNA) molecules. The translation process can be divided into three phases; initiation, elongation and termination. Translation initiation in eukaryotes is a complex event that is assisted by more than 25 polypeptides (Gebauer and Hentze, 2004).

1.4.1. Cap-dependent translation

The majority of the capped eukaryotic mRNAs are translated by the scanning mechanism. In eukaryotes, mRNA molecules usually carry a "cap" structure which consists of m^7GpppN (where m^7G represents 7-methylguanylate, p represents a phosphate group and N represents any base). This is located at the 5' end of eukaryotic mRNAs. This structure enhances translation by facilitating binding of translation initiation factors and the 40S ribosome subunit to the mRNA. A model of the scanning mechanism for eukaryotic translation initiation was proposed in 1978 by Kozak (1989). In this model (Figure 1.6), the methionine-loaded initiator tRNA (L-shaped symbol) binds to GTP-coupled eIF2, to form the ternary complex. This complex then binds to the small (40S) ribosomal subunit, eIF3 and other initiation factors to form the 43S pre-initiation complex. The pre-initiation complex recognizes the mRNA by the binding of eIF3 to the eIF4G subunit of the cap-binding complex. In addition to eIF4G, the cap-binding complex contains eIF4E, which directly binds to the cap, and eIF4A, an RNA helicase that unwinds secondary structure during the subsequent step of scanning. eIF4G also contacts the poly (A)-binding protein (PABP) and this interaction is thought to circularise the mRNA. The 43S pre-initiation complex scans the mRNA in a 5'→3' direction until it identifies the initiator codon AUG. Scanning is assisted by the factors eIF1 and eIF1A. Stable binding of the 43S pre-initiation complex to the AUG codon forms the 48S initiation complex. Subsequent joining of the large (60S) ribosomal subunit results in the formation of the 80S initiation complex. Both AUG recognition and joining of the large ribosomal subunit induces GTP

hydrolysis of eIF2 and eIF5B, respectively. Subsequently, the 80S complex catalyses the formation of the first peptide bond. (Gebauer and Hentze, 2004).

1.4.2. IRES-mediated translation

The initiation of translation in this system does not require a 5' cap structure but requires a RNA segment upstream of the initiation codon which directs the 40S ribosome subunit to the site of translation initiation. This segment of highly structured RNA has been referred to as the "internal ribosome entry site" (IRES).

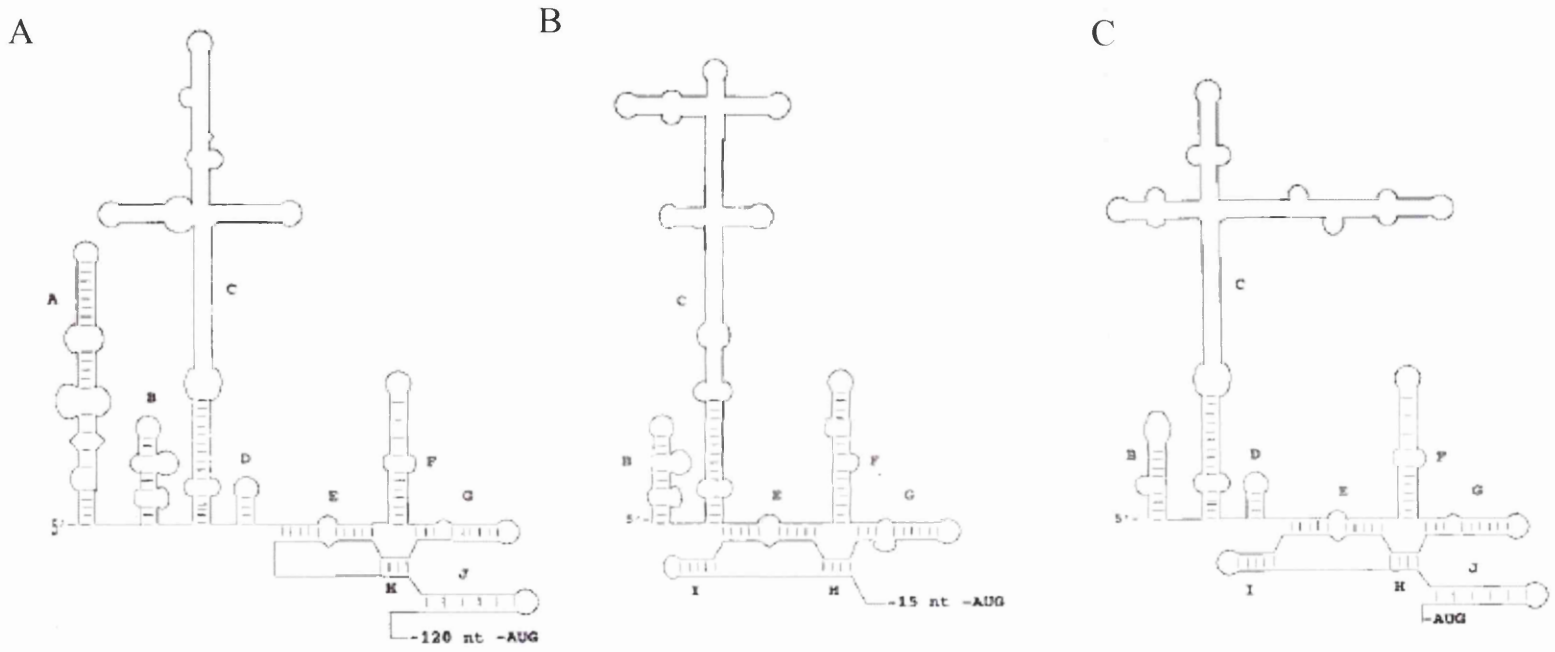
In 1988, it was discovered that uncapped picornaviral mRNAs were translated by an mechanism distinct from scanning (Jang et al., 1988). The development of bicistronic expression vectors containing the poliovirus 5'UTR (Pelletier and Sonenberg, 1988) or the encephalemyocarditis virus (EMCV) 5'UTR (Jang et al., 1988) located in the intercistronic region showed that this region (5'UTR) could confer internal initiation of translation on RNA. Translation of the downstream cistron occurred even when translation of the upstream cistron was abolished.

So far, all picornaviral RNAs have been found to contain IRES elements. The RNA genome of many other RNA viruses including HCV, classical swine fever virus, murine leukemia virus, simian immunodeficiency and cricket paralysis viruses use IRES-mediated translation. In addition, some cellular mRNAs, including translation initiation factors, transcription factors, oncogenes and growth factors, contain IRES elements in their 5'UTRs (Hellen and Sarnow, 2001).

1.4.3. IRES-mediated translation in picornaviruses

Picornaviruses have several characteristics that preclude their use of the cap-dependent scanning mechanism for initiation of translation. The RNA genomes of picornaviruses do not contain a cap structure but have a covalently linked protein called VPg at the 5' end. A

Figure 1.7. Schematic diagrams of the picornavirus IRESs. (A) Type 1 IRES, (B) Type 2 IRES and (C) Type 3 IRES. (taken from Le., 1998)



second feature of picornavirus genomes is the long highly structured 5′ untranslated noncoding region (~600–1300 nucleotides) which makes cap-dependent translation initiation unlikely. The third feature is the presence of several upstream non-authentic start codons which could preclude a scanning ribosome mechanism (Bedard and Semler, 2004).

The picornavirus IRESs were initially divided into two groups based on sequence similarity and structure homology (Pilipenko et al., 1989). The type 1 IRESs are present in enteroviruses and human rhinoviruses, and the type 2 IRESs are found in cardioviruses and aphthoviruses (Figure 1.7). The type 1 IRES element has RNA structural domains A-H and J, and type 2 has structural domains B-C and E-I. A third type of IRES element has been identified in hepatoviruses which preserves the structural features of both the type 1 and type 2 IRES elements. It includes the structural domains D and J which are absent in the type 2 IRES and the structural domain I which is absent in the type 1 IRES. However, the RNA structure of the type 3 IRES is closer to that of type 2 IRES than that of the type 1 IRES.

At picornaviral IRES elements, almost all initiation factors (eIFs) were found to be required for internal initiation, except the cap-binding protein eIF4E (Pestova et al., 1996b). The presence of eIF2, eIF3 and ATP was absolutely essential for the binding of ribosomal 40S subunits to the EMCV IRES, while the additional presence of eIF4A, eIF4B and eIF4F served to improve this binding (Pestova et al., 1996a).

1.5. Initiation of translation by the HCV IRES

The first line of evidence supporting the presence of an IRES element within the 5'UTR of HCV came from *in vitro* translation studies by Tsukiyama-Kohara *et al* (1992). They constructed a bicistronic mRNA system in which the HCV 5'UTR was flanked by a chloramphenicol acetyle transferase (CAT) reporter gene as the first cistron and partial coding sequences (nts 342-1772) of the HCV as a second cistron. In this system, translation of CAT was driven by T7 RNA polymerase (cap-dependent) and the partial

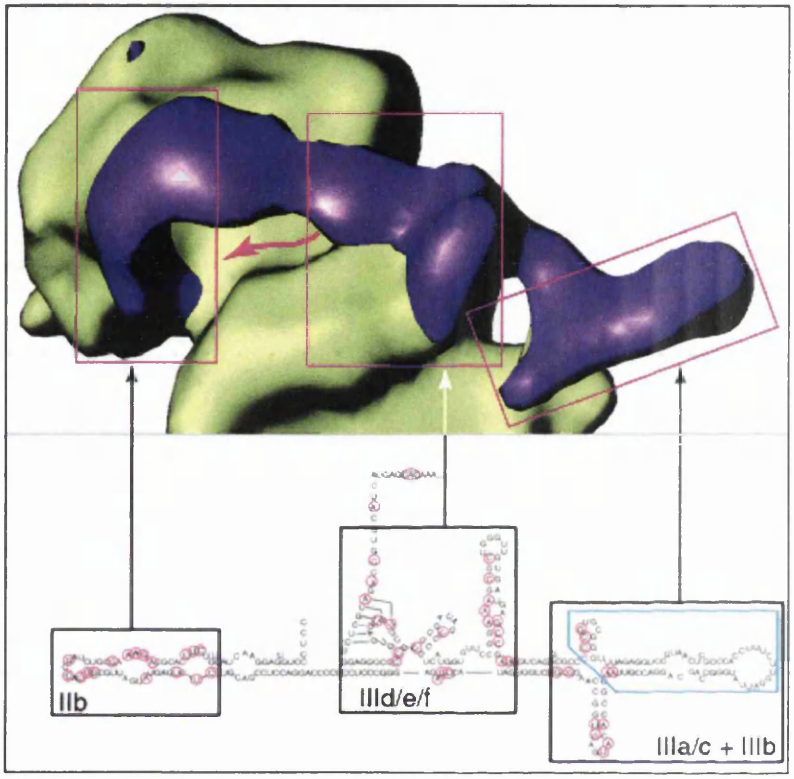


Figure 1.8. Proposed structure of the HCV IRES RNA-40*S* subunit complex.

Domains II, III and IV of IRES are boxed within the secondary structure. Arrows link the tertiary structure domains to the corresponding elements within the IRES RNA cryo-electron microscopy structure. Nucleotides that were previously reported to be inaccessible to solvent upon 40*S* subunit binding are circled in red. The location of the eIF3 binding site on domain IIIb and IIIa is boxed in blue. A magenta arrow pointing to the 3' end indicates the probable location of the coding RNA in the mRNA binding groove (taken from Spahn et al., 2001).

coding region of HCV by the 5'UTR. Efficient translation of both genes was demonstrated in rabbit reticulocyte lysates (*in vitro*) and translational products from the HCV RNA were confirmed using anti-HCV antibodies. However, in coxsackievirus infected cell extracts, in which cap-dependent translation was suppressed, only the core protein was detected suggesting that translation of second cistron occurred even when ribosome entry did not occur at the 5' end of the bicistronic mRNA. Their data indicated that protein synthesis started at the fourth AUG preceding the large open reading frame.

The data presented by Wang *et al.* (1993) further supported the presence of an IRES using an *in vivo* system by introducing synthesized RNA into HepG2 cells. The 5'UTR of HCV was inserted between two reporter genes, CAT and firefly luciferase. Since the *in vitro* synthesised RNAs were uncapped, the CAT activity was not detectable. However, the expression of luciferase occurred both *in vitro* and *in vivo* providing evidence for the presence of an IRES element within the HCV 5'UTR.

An important discovery in the eukaryotic translation field was the finding that the HCV IRES could bind 40S subunits in the absence of any initiation factors, including the eIF2/GTP/initiator tRNA ternary complex (Pestova et al., 1998). The complex structure of 5'UTR determined the correct positioning of the initiation codon in the ribosomal "P" (peptidyl) site in binary complexes. Subsequent addition of the ternary eIF2/GTP/initiator tRNA complex to IRES/40S subunit complexes was necessary and sufficient for formation of 48S complexes.

The 40S subunit appears to interact with the HCV IRES at multiple sites as shown by the residues inaccessible to solvent upon 40S subunit binding. Domain IIIb is not involved in 40*S* subunit binding, but is known to make specific contacts with eIF3. The contact between domain IIb and 40S has also been identified as shown in Figure 1.8 (Spahn et al., 2001). It has been shown that a truncated HCV IRES lacking domain II bound to the 40S subunit with nearly wild type affinity (Kieft et al., 2001).

The translation initiation factor 3 (eIF3) binds specifically to both the HCV IRES and the 40S subunit, although it is not necessary for 48S complex formation. Human eIF3 has a molecular mass of ~600 kDa and contains at least 11 subunits; four of these interact directly with the IRES (Sizova et al., 1998), but it is not yet known which subunit(s) are primary determinants of this interaction. The binding site for eIF3 has been mapped to the terminal half of domain III particularly in the apical stem IIIb of domain III (Kieft et al., 2001; Pestova and Hellen, 1999).The role eIF3 in HCV IRES translation is not yet known. eIF3 may destabilize incorrectly assembled 48S complexes (Kolupaeva et al., 2000).

1.5.1. Cellular factors involved in HCV Cap-independent translation

Several cellular proteins other than the eIFs appear to interact with the HCV 5'UTR. In contrast to the type II IRESs observed in picornaviruses, HCV and pestivirus IRES elements can bind to 40S subunits even in the absence of all eukaryotic initiation factors, while subsequent binding of 60S subunits requires eIF2, eIF3 and other yet uncharacterized factors.

Both polypyrimidine tract binding protein (PTB) and La have been shown to play important roles in picornaviral translation (Belsham et al., 1995). This suggests that perhaps PTB may also support initiation from the HCV and pestivirus IRESs. PTB is a 57 kDa nuclear protein that binds polypyrimidine tracts (Rijnbrand and Lemon, 2000). PTB was shown to bind to the HCV IRES (Ali and Siddiqui, 1995). However, these authors could not demonstrate that the binding of PTB has functional consequences for the efficiency of HCV translation. Immune-depletion of PTB from reticulocyte lysates resulted in reduced HCV IRES activity. However, attempts to restore the original activity by adding PTB failed, suggesting that other factors required for translation had been affected by the depletion procedure. Later, the data presented by Gosert *et al* (2000) showed that the transient expression of PTB from the upstream cistron of dicistronic transcripts stimulated cap-independent translation of a downstream cistron (12 fold) directed by HCV IRES elements placed within the intercistronic space *in vivo*. In contrast, another report by Kaminsky (1995) demonstrated that PTB was not essential for translation mediated by the HCV IRES. In addition to the weak binding of PTB to the

HCV IRES, two other regions in the HCV RNA have been found to bind PTB. One is located in the 3' part of the core coding sequence and includes an oligopyrimidine tract (Ito et al., 1998), and the other is located in the 3'UTR region of the HCV RNA. The PTB-binding region in the 3' part of the core strongly inhibited translation, whereas the 3'UTR appeared to relieve this inhibiting effect and enhances HCV translation *in cis* (Ito et al., 1998). Possibly these two PTB binding regions interact with each other via the PTB protein.

The second protein that appears to be involved in HCV translation is La. La antigen is a multifunctional 52kDa phosphoprotein which was originally identified in patients with autoimmune disorders such as systemic lupus erythematosus. Although a fraction of La antigen is found in the cytoplasm, the majority of the protein is localized in the nucleus. However, cellular stress such as that resulting from a viral infection causes redistribution of the nuclear La to the cytoplasm (Meerovitch et al., 1993). This protein was one of the first cellular proteins identified to interact with IRES elements and induce HCV IRES-mediated translation (Ali et al., 2000; Ali and Siddiqui, 1997). It has been reported that the La protein binds to a region in the 3'-part of the 5'UTR, between nt 291 and 347 which overlaps the polyprotein initiation site and stem-loop IV of HCV RNA. It was reported that the addition of very small amounts of recombinant La were able to stimulate translation directed by the HCV IRES by at least 60-fold in reticulocyte lysates (Ali and Siddiqui, 1997). The data presented by Ali *et al.* (2000) also support the functional requirement for La protein for HCV IRES activity in a liver-derived cell line (HuH7 cells).

Two other cellular proteins have been described to interact with the HCV IRES. Hahm *et al.* (1998) reported a cellular protein (68 kDa) called heterogeneous nuclear ribonucleoprotein L (hnRNP L) that specifically interacted with the 3' border of the HCV IRES spanning part of the core-coding sequence. It bound to a region of about 60 nucleotides between the authentic HCV AUG at nt 342 and nt 402 in the HCV core protein coding sequence, i.e., directly downstream of the binding site for La mentioned above. This sequence represents the 3′ end of the IRES and contributes to IRES function. Binding of hnRNP L to this sequence is correlated with IRES activity, suggesting that

hnRNP L is also involved in the regulation of HCV IRES activity. The close vicinity of the binding sites for La and hnRNP L pointed to a possible interaction between these two proteins in HCV translation.

It has been reported that a 25 kDa protein bound to a sequence in domain II of the HCV IRES. Mutations affecting the binding of this 25 kDa protein also affected translation (Fukushi et al., 1997). Pestova *et al.* (1998) identified a 25 kDa protein cross-linked to HCV RNA. They identified the protein as ribosome protein S9. Another group suggested that ribosomal protein S5 interacted with the HCV IRES elements (Fukushi et al., 2001).

The cellular proteins described above shown to interact with viral IRES elements and stimulate IRES-mediated translation may act as "RNA chaperones" stabilising IRES secondary and tertiary structures to allow efficient translation to take place (Belsham and Sonenberg, 2000).

1.5.2. Assembly of the translation complex

Using data from the literature, a model of the HCV translation initiation pathway has been proposed (Lytle et al., 2002; Rijnbrand and Lemon, 2000). According to this model, the 40S ribosome subunit containing eIF2, Met-tRNA and S9, interacts with the folded viral 5'UTR such that the AUG codon is placed near the anticodon of the Met-tRNA. Two possible scenarios resulting in formation of the 48S preinitiation complex have been suggested (Figure 1.9). In the first proposal, the 40S ribosome subunit binds both eIF3 and the ternary complex including eIF2-GTP-Met-tRNA to form a 43S preinitiation complex which is similar to that occurring in cap-dependent translation. The 43S complex subsequently binds to the 5'UTR. Then, the 48S complex is formed by a conformational change in the IRES which places the AUG codon against the Met-tRNA anticodon (Figure 1.9.A). The alternative scenario suggests that eIF3 and the 40S ribosome subunit are bound by the IRES. In a subsequent step, the ternary complex is incorporated into the 40S-RNA complex resulting in a conformational change in the viral RNA that places codon and anticodon opposite each other to form a 48S complex (Figure 1.9.B). Then, in both

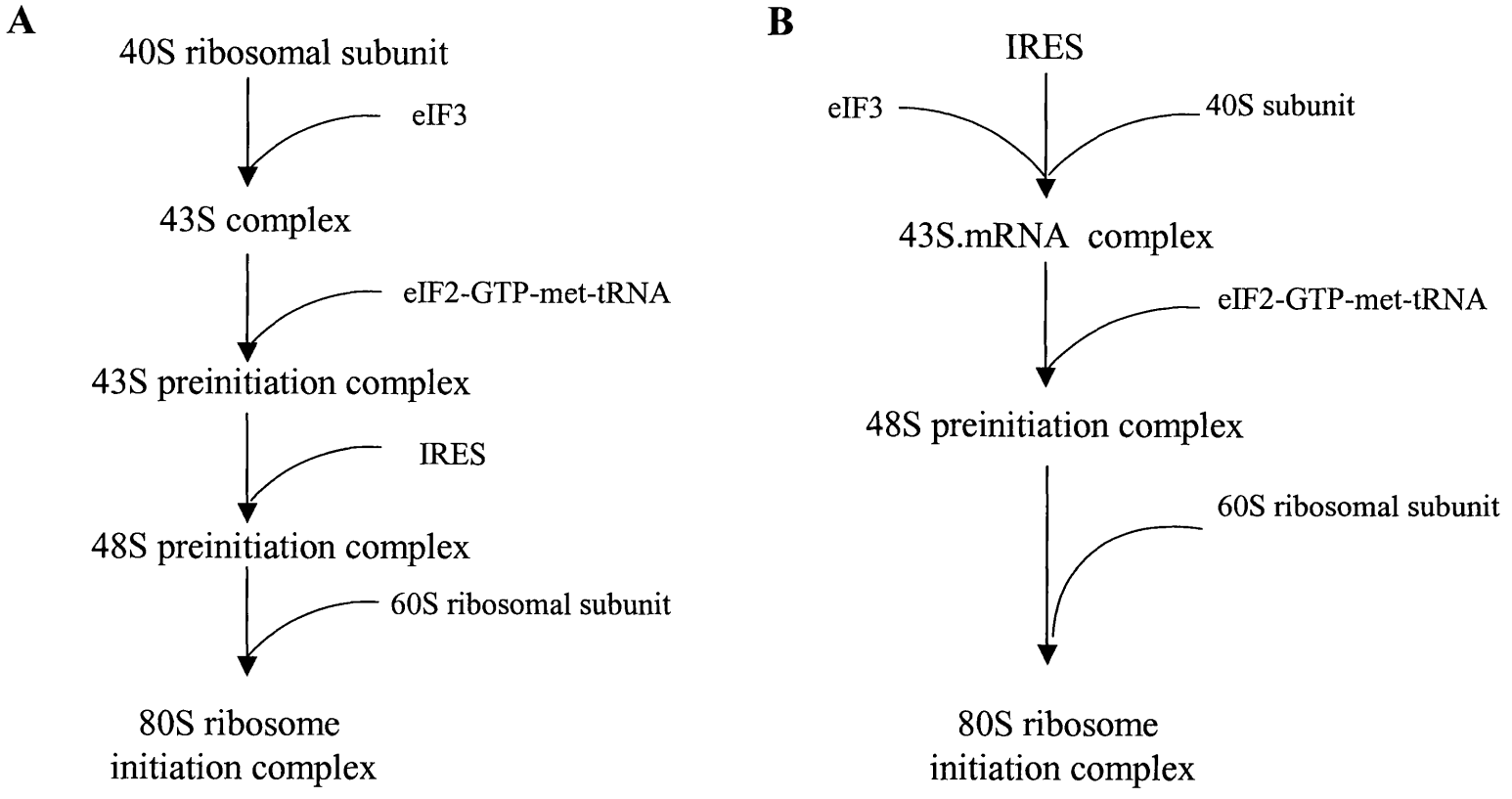


Figure 1.9. Two possible mechanisms in assembly of the translation complex on the HCV 5'UTR
(taken from Rijnbrand & Lemon., 2000)

scenarios, subsequent binding of the 60S subunit results in the assembly of a functional 80S ribosome, which starts to translate the HCV polyprotein. It should be noted that the precise order of these events remains uncertain.

1.5.3. Mapping the HCV internal ribosome entry site

Several approaches have been used to identify the minimal sequence required for HCV IRES dependent translation and to map the 5' and 3' borders of the HCV IRES elements. The commonly used technique has been the analysis of bicistronic RNAs in which two reporter protein coding sequences are separated by an IRES sequence. Translation of the upstream gene occurs in a 5'cap-dependent mechanism, while translation of the downstream reading frame is driven by the IRES element. Quantifiable reporter proteins like luciferase (Tsukiyama-Kohara et al., 1992; Wang et al., 1993) or chloramphenicol acetyl transferase (CAT) (Rijnbrand et al., 1995) have been used by several groups. However, the HCV core protein coding region (Fukushi et al., 1994; Tsukiyama-Kohara et al., 1992) or the complete HCV open reading frame have been used by other groups as the reporter sequence in order to study the translation efficiency in a more natural condition (Honda et al., 1996b).

1.5.3.1. The 5' border of HCV IRES

An early study by Tsukiyama-Kohara *et al.* (1992) reported the 5' border to be located between nts 100 and 156. Later, Fukushi *et al.* (1994) reported that the entire 5'UTR was needed for IRES activity. Despite this report, most studies have shown the 5' border of the IRES to be at or near the 5' end of domain II (nt 44) (Honda et al., 1996b; Kamoshita et al., 1997; Reynolds et al., 1996). Honda *et al.* (1999c) showed that deletion of nts 32-37 did not have any effect on IRES activity, while substitutions at nts 45-46 had a strong, negative effect on translation (Honda et al., 1999). They suggested that the 5' border of the HCV IRES was between nts 38 and 46, i.e., just upstream of stem-loop II. As can be seen, most studies agreed that stem-loop I is not required for IRES activity. In fact, some studies have shown that removal of stem loop I enhanced the translation efficiency of the

IRES (Honda et al., 1996b; Rijnbrand et al., 1995). In addition, the data presented by Kamoshita *et al.* (1997) suggested that the inhibitory effect of stem-loop I on IRES activity may be cell-type specific. They showed that RNA transcripts lacking nts 1-22 were translated more efficiently in a HeLa cell lysate, but less efficiently in African green monkey kidney cells, suggesting that host cell-specific factors may interact with stem-loop I. More recently, Friebe *et al.* (2001) observed that deletions of stem loop I or the spacer sequence (nts 20-44) decrease RNA translation by two- to threefold in a bicistronic system in HuH7 cells. Luo *et al.* (2003) also reported that deletion and nucleotide substitutions in the 5' proximal stem-loop structure resulted in reduction of translation of a reporter gene (CAT) by two- to five fold, suggesting that the 5' proximal stem-loop RNA element also modulates HCV RNA translation.

1.5.3.2. The 3' border of HCV IRES

The 3' boundary of the IRES is less certain. The IRES extends in a 3' direction as far as the initiator AUG codon. It has been shown that the RNA pseudoknot involving stem-loop IIIf was essential for translation of HCV (Wang et al., 1995). In early studies, efficient IRES activity was observed in dicistronic constructs containing between 3 (initiator AUG codon) and 8 nucleotides of the core-coding sequence (Rijnbrand et al., 1995; Tsukiyama-Kohara et al., 1992). In contrast, Reynolds *et al* (1995) reported a significant increase in IRES activity if 12 to 30 nts of core coding sequence were included in the RNA transcript. The importance of the first 24 nucleotides of the core-coding sequence in IRES-dependent translation was supported by the observation that chimeric polioviruses containing the HCV IRES required at least 24 nts of core-coding sequence for viral replication (Lu and Wimmer, 1996). A chimeric virus containing 370 nts of the core protein coding region showed the most efficient replication.

Two different explanations have been proposed for the discrepancy between these observations. First, Reynolds *et al.* suggested that a sequence homology existing between the firefly luciferase and HCV core nt sequences may be responsible for the IRES activity observed by Wang *et al.* (Tsukiyama-Kohara et al., 1992; Wang et al., 1993) in constructs

containing the luciferase sequence fused directly to the HCV initiator AUG. Another explanation was that the fusion of some reporter gene sequences to the HCV 5'UTR may result in an RNA structure which may have an inhibitory effect on translation initiation (Honda et al., 1996a). It has been shown that the 40S ribosome interacted with the RNA directly at the site of the initiator AUG codon and structures around the AUG reduced IRES activity (Pestova et al., 1998). So it may be that the inclusion of the 5' core coding sequence simply ensures the absence of an unfavourable base pairing structure around the AUG initiation codon. Taken together, the findings suggested that the minimal sequence for IRES function includes stem loops II and III as far as the initiator AUG codon and extends into the core encoding region which makes up part of the domain IV structure.

1.5.4. The impact of the 3' UTR on HCV translation

The RNA of HCV, GBV-B and pestiviruses differs from other RNAs containing IRESs by the lack of a poly (A) tail at the 3' end of the RNA. Instead there is a variable poly (U-C) stretch plus a conserved X region at the 3' end (Kolykhalov et al., 1996). The X region seemed to form a three-stem-loop structure and bound to polypyrimidine tract-binding protein (PTB) (Tsuchihara et al., 1997). PTB has been reported to bind to the IRES region of HCV, and regulate its translation (Ali and Siddiqui, 1995). There are controversial reports regarding the effect of the 3'UTR on translation of HCV. One study published by Ito *et al.* (1998) suggested that the X region at the 3' end of the HCV genome enhanced IRES-dependent translation weakly. They showed that HCV RNA containing the X region was translated three- to five fold more than the corresponding RNAs without this region. Mutations that abolished PTB binding in the X region reduced, but did not completely abolish, the enhancement of translation. However, they did not use the entire HCV 3'UTR. In contrast, Murakami *et al.* (2001) observed that the complete 3'UTR downregulated HCV translation *in vitro*. This inhibition was removed when the poly (U/UC) or stem–loop III (SL3) regions of the 3'UTR were deleted. The absence of regulation of HCV IRES activity by the 3'UTR has been reported by other studies. Fang & Moyer (2000) demonstrated that the presence or absence of the 3'UTR sequence did not affect translation efficiency in an *in vitro* system. Recently, the impact on translation

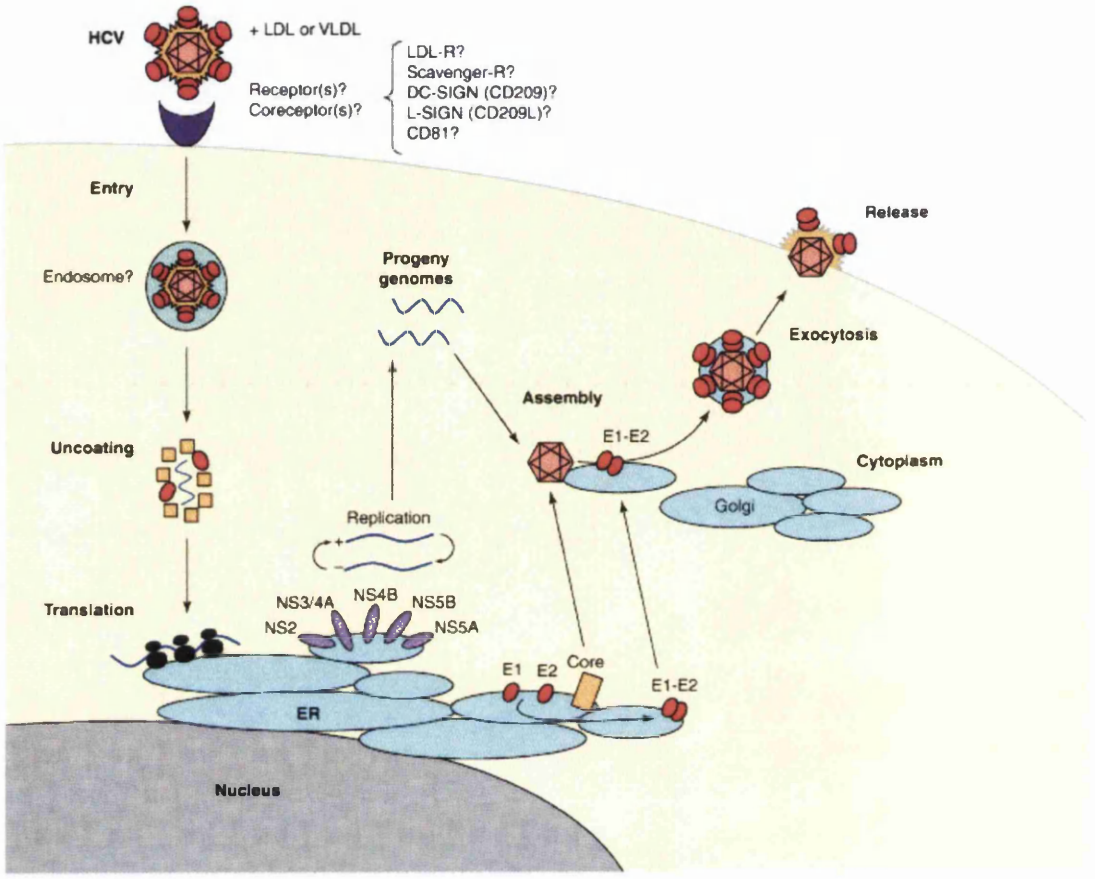


Figure 1.10. Putative hepatitis C virus (HCV) replication cycle.
ER=endoplasmic reticulum. (taken from Racanelli & Rehermann, 2003).

efficiency of a series of mutations on the 3'UTR was analysed. The results showed that complete deletion of the variable region, the poly (U/UC) tract or the 3'X region did not modify HCV IRES activity (Friebe and Bartenschlager, 2002). Moreover, Kong & Sarnow (2002) have shown that the HCV 3'UTR modulates neither the translation nor the stability of a chimeric mRNA. These results were supported by the data presented by Imbert *et al.* (2003) suggesting that the HCV 3'UTR does not affect IRES-dependent translation efficiency, even in the presence of HCV structural or non-structural proteins in hepatic and non-hepatic cell lines tested.

1.6. Replication of HCV

Little is known about the mechanism of replication of HCV due to lack of an efficient culture system for its study. Because of the lack of convenient animal models and cell culture systems, the proposed model of replication has been suggested based on the relationship of HCV with the other members of *Flaviviridea* family and characterisation of recombinant HCV proteins (Figure 1.10). The detection of HCV non-structural proteins and viral RNA in the livers of infected patients (Blight et al., 1993) or experimentally infected chimpanzees showed that the liver is a main site of HCV replication. It has been reported that HCV can also replicate in peripheral blood mononuclear cells (PBMCs) (Cribier et al., 1995) and experimentally infected B and T cell lines (Mizutani et al., 1996).

How HCV enters the cell is unknown, but it may be by interaction with a receptor. Several putative receptors have been identified. The first to be identified was CD81(Pileri et al., 1998). CD81 is a 25 kDa protein and a member of the tetraspanin superfamily of cell surface proteins. Binding of E2 was mapped to the major extracellular loop of CD81. This is the most variable part of the molecule but it is highly conserved in humans and chimpanzees, the only species permissive for HCV infection (Pileri et al., 1998). Analysis of inhibition of the E2-CD81 interaction by using a panel of anti-E2 monoclonal antibodies has shown that amino acids 412-417 of E2 are involved in this interaction

(Owsianka et al., 2001). A more recent report suggested that CD81 is an entry coreceptor and that other cellular factors are involved during binding and entry (Cormier et al., 2004). The LDL receptor has also been proposed as a candidate for the HCV receptor. Agnello *et al.* (1999) showed a direct correlation between cell surface expression of LDL-R and the number of infected cells positive for the presence of HCV RNA using different cell lines. It has also been observed that HCV is not able to bind to COS-7 cells unless they express the LDL receptor (Monazahian et al., 1999). However, endocytosis of HCV has been reported in LDL-deficient fibroblasts (Agnello et al., 1999).

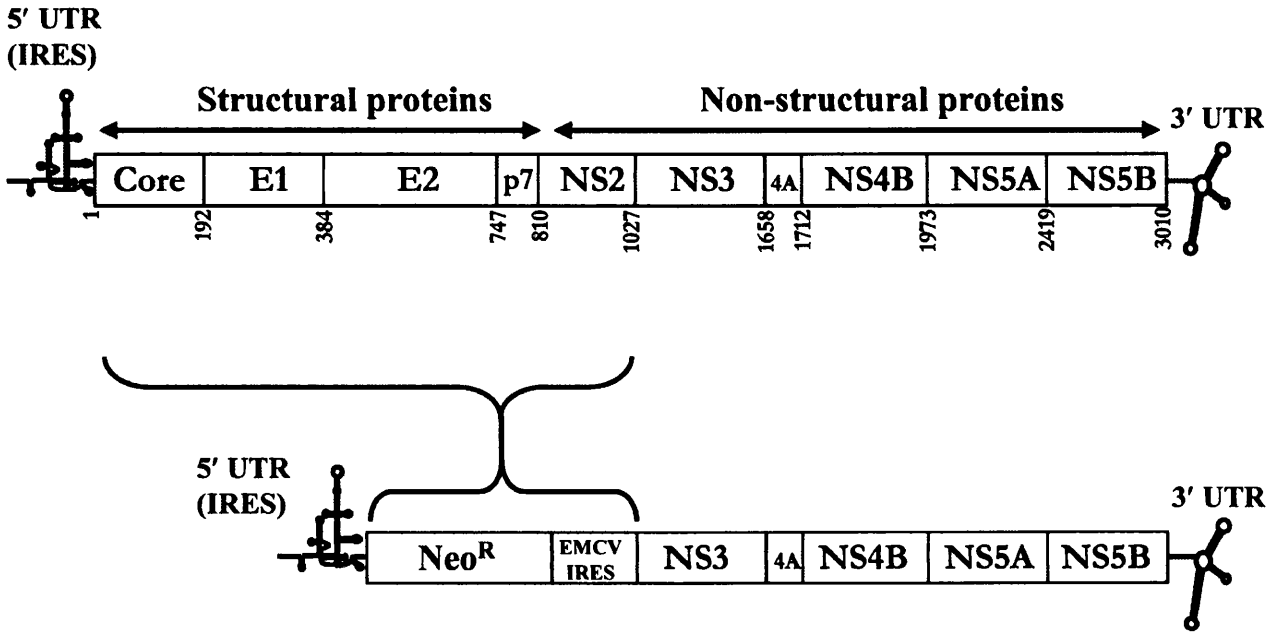
Other putative receptors that have been reported are human scavenger receptor class B type I (SR-BI) (Scarselli et al., 2002), asialoglycoprotein receptor (ASGP-R) (Saunier et al., 2003) and liver/lymph node-specific intercellular adhesion molecule-3-grabbing integrin (L-SIGN) which is a calcium-dependent lectin expressed on endothelial cells of liver and lymph nodes (Gardner et al., 2003).

Coprecipitation experiments have indicated that the non-structural proteins probably form a replicase complex on the cytoplasmic side of the ER membrane (Ishido et al., 1998) which is similar to other flaviviruses and pestiviruses. 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 (Ito et al., 1996; Lanford et al., 1994; Rumin et al., 1999). These systems suffer from both poor efficiency and poor reproducibility and none permitted detailed analysis of the replication cycle. The development of subgenomic HCV RNA replicons has allowed the study of RNA replication in the human hepatoma cell line (HuH7) (Lohmann et al., 1999; Pietschmann et al., 2002). This system will be covered in more detail in section 1.6.1.

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

Fig. 1.11. Structure of the HCV genome and a typical subgenomic replicon.

The subgenomic replicon was constructed by replacing the core through to NS2 region with the gene encoding neomycin phosphotransferase (neo) and a second IRES, derived from the picornavirus EMCV. The numbers underneath the genome represent polyprotein amino-acid sequence numbers of the genotype 1b infectious clone (J4 isolate). (taken from Macdonald & Harris, 2004).



molecular infectious clones of HCV. A mouse model for HCV infection has been reported 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 (3×10^4 to 3×10^6 copies/ml) for a period of 15-17 weeks. Successful infections were established with viral genotypes 1a, 1b, 3a and 6a, with rapid increases in viral RNA titres to levels easily detected by standard commercial assays. Although promising, the transplantation of primary human hepatocytes into mice a few days after birth is technically difficult and requires specialised techniques.

1.6.1. Subgenomic replicon system

Recently, Lohmann *et al.* (1999) developed selectable subgenomic replicons which can replicate to high levels after transfection into the human hepatoma cell line HuH7 (Figure 1.11). A consensus genotype 1b genome (Lohmann et al., 1999) was constructed from material cloned from the 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 the neo gene was directed by the HCV IRES. Encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) directed translation of the HCV non-structural genes containing NS2 or NS3 up to NS5B including the 3'UTR. The RNA replicons were generated by *in vitro* transcription from DNA plasmid constructs. These bicistronic RNA transcripts were transfected into HuH7 cells and those cells harbouring the replicon were selected with G418. Selected colonies carried large amounts of HCV RNA (1000-5000 RNA molecule per cell). Immunoprecipitation analysis confirmed the presence of the NS3, NS4, NS5A and NS5B proteins in the cytoplasm.

Blight *et at.*(2000) reported the identification of multiple adaptive mutations which cluster in a 30 amino acid region in the NS5A during propagation of the cells containing the replicon. These mutations enhanced RNA replication, which increased the G418

transduction efficiency. However, since the complete structural region of the genome was deleted, virus assembly and particle release could not be studied with this system.

To establish a system that allows efficient replication of the complete HCV genome in cell culture, Pietschmann *et al.* (2002) constructed a full length replicon of genotype 1b carrying cell adaptive mutations and showed that these genomes replicated stably and expressed all viral proteins for prolonged cultivation periods. However, no evidence of virus particle assembly was found, suggesting that host cell factors important for virus production are not present in HuH-7 cells. Later Blight *et al.* (2003) reported the establishment of efficient RNA replication systems for genotype 1a strain H77 containing adaptive amino acid substitutions in both NS3 and NS5A in the highly permissive HuH-7.5 hepatoma subline. Although the production of infectious virus particles was not successful, it probably provides the framework to proceed toward this goal. Future research will may identify cellular factors needed to support HCV assembly and release in cell culture. This would allow construction of a model of the entire HCV replication cycle in cell culture.

1.7. Immune response to HCV

One of the first host defence mechanisms against viral infections is the non-specific immune response involving type I IFN secretion and natural killer (NK) cell activation. Recent data suggest that HCV replication induces the secretion of endogenous type I IFN but does not efficiently inhibit it in the human (Thimme et al., 2001). It has been reported that dendritic cells recovered from chronically HCV-infected patients showed an impaired capacity to activate natural killer (NK) cells in response to IFN-α stimulation. This impairment might affect IFN responsiveness in the acute phase of infection (Jinushi et al., 2003). Various HCV proteins have been shown to inhibit the antiviral effects of IFN-α *in vitro.* Tseng and Klimpel (2002) suggested that binding of the HCV envelope glycoprotein E2 to CD81 might inhibit NK cell functions such as proliferation, cytokine production and cytotoxic granule release and also their ability to produce IFN gamma. In addition, the HCV NS3/4A serine protease blocks the phosphorylation and effector action

of interferon regulatory factor-3 (IRF-3) which is a key cellular antiviral signalling molecule (Foy et al., 2003)

Antibodies targeting various HCV epitopes are generally detected 7 to 31 weeks after infection. Antibodies directed against HVR1 have been shown to be neutralizing, protecting chimpanzees against HCV infection after *in vitro* neutralization of the corresponding strain (Farci et al., 1994). However, it is not known whether control of HCV replication by neutralizing antibodies plays a major role in viral clearance. It has been suggested that variability in the HVR1 region could generate new HCV variants which are able to escape from neutralizing antibodies. However, several recent findings argue against a major role of 'escape variants' in the persistence of infection. Penin *et al.* (2001) demonstrated that, despite strong amino acid sequence variability, the chemicophysical properties and conformation of the HVR1 were highly conserved. The data presented by Cerino *et al.* (2001) demonstrated that it is possible to induce a broadly cross-reactive monoclonal antibody response to HVR1 variants which could also recognize viral particles. The recent development of infectious retroviral HCV pseudotypes, comprising HIV capsids bearing HCV envelope glycoproteins, have allowed the study of neutralizing antibodies during HCV infection (Bartosch et al., 2003). Logvinoff *et al.* (2004) found no association between the development of neutralising antibodies against these pseudo particles and viral clearance which occurred in two of seven acutely infected patients. Therefore, the role of the neutralizing responses in the control of HCV infection or persistence is still unclear.

The role of cellular immune responses in HCV has been better studied. Spontaneous HCV clearance has been shown to be associated with a sustained, vigorous and virus-specific CD4+ T-cell response in peripheral blood. This response is maintained for several years after viral clearance (Bertoletti and Ferrari, 2003). By contrast, the CD4+ T-cell response is weak, delayed or transient in patients who develop persistent infection. CD8+ T-cell (cytotoxic T lymphocyte, CTL) responses also play a major role in virus elimination. Spontaneous viral clearance in man and chimpanzees is associated with a strong, sustained and multispecific CD8+ T-cell- mediated response (Cooper et al., 1999; Thimme et al.,

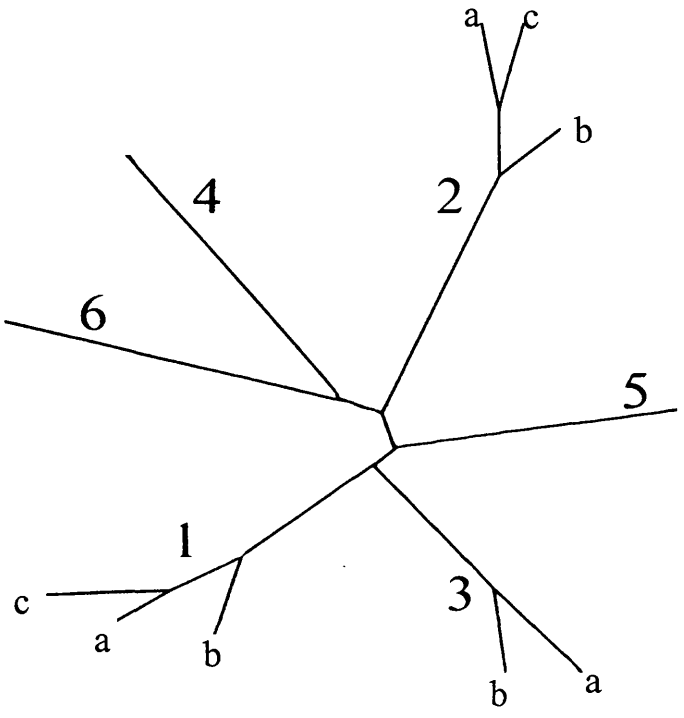


Figure 1.12. Phylogenetic analysis of nucleotide sequences in part of the HCV NS5A region.
They are classified into six genotypes and subtypes a, b, etc. (taken from Simmonds, 1995).

2001). It has been reported that the ability of dendritic cells (DC) to stimulate the proliferation of allogeneic T cells following infection was dramatically impaired in patients with chronic HCV infection in comparison to naive-DCs (Bain et al., 2001). Other possible mechanisms include the defective functions of HCV-specific CD8+ T cells in chronically infected patients (Wedemeyer et al., 2002) and the appearance of escape mutations in epitopes which impaired class I MHC binding and/or CTL recognition (Erickson et al., 2001).

1.8. Genetic variability of the HCV genome

1.8.1. Genotypes

Analysis of full length or partial sequences of HCV strains isolated from various geographical regions of the world led to the classification of six HCV genotypes (Figure 1.12) and these genotypes have been further subdivided into subtypes within each genotype. The types have been numbered 1 to 6 and the subtypes a, b, c (Simmonds et al., 1994). Differences between genotypes over the complete virus genome are relatively high with nucleotide differences of 30%-35%. Subtypes within a genotype have around 80% nucleotide identity. Isolates within a specific subtype have more than 90% nucleotide identity (Simmonds, 2004).

Direct sequence comparison of full-length HCV genomes is the ideal method for typing HCV but it is laborious and time consuming. A more rapid method is analysis of the 5'UTR region by restriction fragment length polymorphism (RFLP) for clinical purposes (Davidson et al., 1995). For subtyping of HCV isolates, sequence analysis of additional regions of the genome such as the core, E1 or NS5 is needed.

HCV genotypes differ throughout the geographical regions. Genotypes 1, 2 and 3 have a broad worldwide distribution, although their relative 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 (Smith et al., 1995a) 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 variety of genotype 3 subtypes can be found in Southern Asian and Indonesia, but genotype 1 remains the predominant type (Smith et al., 1995a; Tokita et al., 1994). Genotype 4 is common in Egypt and Central Africa (Chamberlain et al., 1997a; Simmonds, 1995). Both genotype 5 and 6 have restricted distributions in South Africa (Ohto et al., 1994) and Hong Kong respectively (Chamberlain et al., 1997b; Di Bisceglie, 1998). In Europe, types 1b and 2 are widely distributed particularly in older age groups who acquired HCV through blood transfusion whereas genotypes 1a and 3 are predominant in younger age groups likely to be those infected through intravenous drug use (Simmonds et al., 1996).

So far no correlation has been found between HCV genotype and the rate of persistence after acute infection, the severity of chronic liver diseases or the development of HCV associated extra-hepatic disorders (Pawlotsky, 2003a). A more recent study reported that acute infection in young Caucasian men with HCV genotype 3 leads more often to spontaneous clearance than infection with HCV genotype 1. Ninety three percent of individuals exposed to HCV genotype 1 but only 63% of individuals exposed to genotype 3 experienced a chronic course of the infection (Lehmann et al., 2004). It has been shown that genotype 3 is more associated with the appearance of steatosis in the liver than other genotypes (Adinolfi et al., 2001a; Ramalho, 2003).

1.8.2. Variability in the HCV genome

Variation has been found throughout the whole HCV genome (Bukh et al., 1995). The 3' X tail and 5'UTR are highly conserved between genotypes suggesting their essential role in replication and translation (Smith et al., 1995b). The core protein is the most conserved of the viral proteins (Bukh et al., 1994) followed by conserved regions in the NS3 and NS5B (Okamoto et al., 1992). Extensive variation has been shown in envelope

proteins (E1 and E2) particularly a 27 amino acid stretch located at the N-terminal region of the E2 known as hypervariable region 1 (HVR1) (Hijikata et al., 1991; Weiner et al., 1991a). This is the most variable region of the HCV genome. Other variable regions include a HVR2 located downstream of HVR1 in the E2 protein (Kato et al., 1992) which seems to be present only in genotype 1.

1.8.3. Viral quasispecies

One of the important characteristics of HCV, like many other RNA viruses, is that its genome shows significant genetic heterogeneity. Even in a single infected individual, HCV circulates as a population of closely related but heterogeneous sequences called a quasispecies (Martell et al., 1992). These variants together represent a "quasispecies" or form a "quasispecies distribution". As discussed in a review, (Smith and Simmonds, 1997), the term of quasispecies has been used incorrectly by many authors to refer to each of the distinct variants which together form the quasispecies. In this thesis, the term is used to describe the whole population of related sequences observed within a single infected patient.

"Quasispecies" was defined by Smith *et al* (1997b) as a population of viruses that share a common origin but which have distinct genomic sequences. The viral population in a single individual is composed of a sequence that is dominant called the master or majority sequence and a number of sequences differing from the majority sequence. The majority sequence does not provide information about the structure of the quasispecies.

The quasispecies nature of HCV has been attributed to the RNA polymerase lacking proofreading 3'-5' exonuclease activity and therefore each round of replication is accompanied by the incorporation of mutations. The average rate of mutation in the HCV genome has been estimated to be 1.44 to 1.92 x 10^{-3} substitutions per site per year (Ogata et al., 1991).

The flavivirus, GBV-C 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, GBV-C isolates from around the world differ by only 14% (Smith et al., 1997a). Therefore the heterogeneous nature of HCV cannot be explained solely by the error prone RNA polymerase.

The quasispecies nature of the virus 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 in the HVR1 region contributes to the high level of chronicity seen in HCV-infected patients (Yuki et al., 1997) . 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 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.

1.9. Natural history of HCV infection

1.9.1. Epidemiology and transmission

Hepatitis C infection has a worldwide distribution. According to World Health Organisation (WHO) estimates, approximately 3% of the world population, or about 170 million people, may be infected with hepatitis C virus (WHO, 1997). The incidence of HCV in healthy volunteer blood donors in the USA varies between 0.17% and 1.4% (Murphy et al., 1996) and in the UK is 0.35% (Mutimer et al., 1995). Egypt has the highest reported prevalence worldwide with an estimate of 24% (Arthur et al., 1997)

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 transmitted primarily by exposure to infected blood. However, in up to 50% of cases, no recognisable transmission route could be identified (Van Damme and Vellinga, 1998).

1.9.1.1. Parenteral Transmission of HCV

By introduction of mandatory blood donor screening for anti-HCV in 1989, the incidence of post transfusion HCV was substantially reduced to less than 1% but was not completely eliminated. The increased sensitivity of third generation screening tests further reduced the risk of infection via blood transfusion. 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 and Wands, 1997; Yap et al., 1994). In thalassemic children who received multiple blood transfusions, the incidence of HCV infection varied from 55% to 83% (Memon and Memon, 2002).

Intravenous drug use is currently the main risk factor for HCV transmission. The reported incidence varies between 31% (Nakata et al., 1994) and 98% (Westh et al., 1993) in different parts of the world. At least 60% of new HCV infections in the United States were related to intravenous drug use (IVDU) (Alter, 1999). Risk of infection has been reduced as a result of change in injection practice by the introduction of needle-exchange programmes (Hagan et al., 1995).

Healthcare professionals working with blood and blood products are at greater risk of contracting HCV via exposure to blood and needle-stick injuries. Transmission of HCV in the healthcare setting can occur from patient to patient and from patients to healthcare worker or vice versa. In most developing countries, unsafe injections have been the major cause of HCV transmission. The high prevalence of HCV in Egypt is associated with injections of anti-schistosomiasis treatment with unsterilised needles (Frank et al., 2000).

Outbreaks of HCV have been reported among patients on chronic haemodialysis possibly as a result of nosocomial transmission due to environmental contamination from patient to patient or incorrect sterilisation procedures of dialysis machines between HCV positive and negative patients (McLaughlin et al., 1997). Procedures such as tattooing, piercing and acupuncture are also associated with potential risk of HCV transmission.

1.9.1.2. Non-parenteral Transmission of HCV

Cases of hepatitis C infection have been reported where there is no apparent evidence of parental exposure. Sexual transmission of HCV does occur, although it is infrequent (Memon and Memon, 2002). Co-infection with HIV seems to increase the risk of HCV sexual transmission (Lissen et al., 1993). Perinatal transmission of HCV occurred in between 0% and 15% and was most likely to occur at the time of delivery (Van Damme and Vellinga, 1998). It has been reported that the risk was increased by higher viral load in the mother (Zanetti et al., 1998). There has been no evidence of transmission of HCV via breast feeding (Polywka et al., 1999).

1.9.2. Prevention of HCV infection

No vaccine or effective prophylaxis is available for the prevention of HCV infection. Difficulties in HCV vaccine development include the heterogeneous nature of HCV which result in rapid changes within an infected individual. Another potential problem is the lack of long term immunity in individuals who resolve the infection. Therefore, there is potential for reinfection with either the same or a different HCV isolate. Prevention of HCV infection will be the best strategy for the near future. Screening is mandatory for donors of blood, organs, tissues, and semen. Otherwise screening is recommended only for those with a higher risk of having HCV infection such as individuals who received a blood transfusion or a transplant prior to donor screening, haemophiliacs, healthcare workers after exposure to HCV positive blood, haemodialysed patients, IVDUs and children born to HCV-infected mothers. However, education of the public on the disease, its transmission and prevention could reduce the number of new infections.

In many developed countries, drug use is the major source of HCV infection. Education has a major role to prevent the initiation of drug use in adults. Easy access to sterile syringes (syringe exchange programs) accompanied by counselling are important strategies for limiting HCV transmission (Taylor et al., 2000). Patients with known HCV infection should be counselled on how to limit their risk of transmitting HCV to others. They should be advised not to donate blood, organs or semen and not to share toothbrushes and razors. Education of healthcare workers to use standard precautions for prevention of transmission of blood-borne pathogens in healthcare is important particularly in the haemodialysis setting. Individuals with stable long-term relationships should be informed of the low risk of transmission, and encouraged to discuss the risk and the use of barrier precautions with their sexual partners. The use of condoms is strongly recommended in individuals with multiple partners and intra-venous drug users (IVDUs) (WHO, 1999)

1.9.3. Clinical features of HCV infection

The natural history of HCV infection is poorly documented because the onset of acute infection is often silent. Also, the interval between infection and the development of cirrhosis can exceed 30 years, which means there are few prospective studies.

1.9.3.1. Acute hepatitis C

Acute infection is usually asymptomatic. Clinical manifestations occur in only one third of patients. Symptoms, when they occur, are malaise, nausea and pain in the right upper quadrant of the abdomen. Jaundice occurs in less than 20% of infected patients (Alter et al., 1992). Symptoms and signs develop within 2 to 12 weeks of exposure to the virus (Marcellin, 1999a). HCV viral RNA becomes detectable by RT-PCR in serum 7 to 21 days after exposure to infection (Farci et al., 1991) and antibodies appear within 20 to 150 days (Hoofnagle, 1997). Serum alanine aminotransferase (ALT) levels begin to increase

before the appearance of clinical symptoms. The rate of chronicity varies from 55% to 91% (Wiese et al., 2000).

1.9.3.2. Chronic hepatitis C

HCV infection is characterised by a high incidence of chronicity demonstrated by the persistence of HCV RNA in serum. Serum ALT levels frequently fluctuate and may be normal or elevated in the same patient at different times. It has been shown that serum ALT levels are not a good predictor of liver fibrosis (Marcellin, 1999a). Fibrosis has been reported to develop into cirrhosis in 20-30% of patients within 2 or 3 decades (Takahashi et al., 1993). The role of viral factors in the progression of fibrotic process is not clear. However, several other factors have been reported to have a major role in progression of liver fibrosis toward cirrhosis. These includes high alcohol consumption, co-infection with HIV or HBV, diabetes, obesity and immunosuppression (Pawlotsky, 2004). HCV-related end-stage liver disease is now a leading cause of liver transplantation (Marcellin, 1999b; Pawlotsky, 2003b). Hepatocellular carcinoma (HCC) is the most severe complication of chronic HCV infection and it has been reported to occur in 1-4% of patients (Pawlotsky, 2004).

1.9.3.3. Extrahepatic manifestations of HCV infection

Many extra-hepatic manifestations have been associated with HCV infection. HCV infection is most closely associated with essential mixed cryoglobulinaemia (aggregation of immunoglobulins in blood vessels in response to cold) and membranoproliferative glomerulonephritis (deposition of immune complexes in the capillaries of the glomeruli) (Manns and Rambusch, 1999; Teoh and Farrell, 2004). Review of 50 different studies showed that the HCV prevalence in patients with porphyria cutanea tarda is approximately 50% higher than in the general population (Fargion and Fracanzani, 2003). An association between HCV infection and diabetes mellitus type 2 has also been reported (Mason et al., 1999; Mehta et al., 2003).

1.9.4. Genotypes and natural history of HCV

In contrast to the clear differences between genotypes in their response to antiviral therapy, the issue of whether the 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 were reports that genotype 1b was associated with a greater frequency of cirrhosis and the development of HCC than other genotypes (Bruno et al., 1997; Dusheiko et al., 1994), while other studies have failed to show any relationship between disease severity and genotype (Benvegnu et al., 1997). Infection with genotype 3 is associated with steatosis, a condition in which lipids accumulate in the liver (Rubbia-Brandt et al., 2000). It has been suggested that steatosis resulted from direct cytopathic damage by the virus to hepatocytes from a block in lipoprotein secretion (Serfaty et al., 2001).

1.10. Treatment of HCV

1.10.1. Antiviral therapy

Interferon alpha (IFN-α) was the first recommended therapy for chronic hepatitis C and a 6 months course of treatment led to normalisation of ALT levels, loss of detectable virus in blood and reduction of inflammation in liver biopsies in a few patients. However, IFN-α monotherapy showed a sustained response (defined as undetectable HCV RNA in serum at 24-weeks after the end of treatment) in only a minority of patients. Extension of duration of treatment up to 12 months still resulted in a sustained response being achieved in only 13-25% of patients (Poynard et al., 1996). The combination of ribavirin and IFN-α enhanced the sustained response rate, particularly for patients infected with HCV genotype 1 (McHutchison et al., 1998). Ribavirin is a guanosine analogue that has antiviral efficacy against respiratory syncytial virus. When used as monotherapy against HCV for 24 months, it reduced serum ALT levels but had no impact on HCV RNA levels in serum or liver (Di Bisceglie et al., 1995). Treatment of HCV infected patients with the

combination of a standard interferon-α with ribavirin (in the respective doses of 3 million units of IFN-α three times a week and 800–1200 mg/day of ribavirin in a divided dose, depending on bodyweight) increased the overall sustained response up to approximately 50%. Several factors have an important role in the response rate including genotype, viral load and presence of cirrhosis, age and gender. Patients with genotype 3 without cirrhosis had a sustained response rate of 65–75% after 6 months of combination treatment whereas, for genotype 1, high viral load and cirrhosis, it was 10–20% (Teoh and Farrell, 2004). More recently, the development of pegylated IFN-α has improved sustained responses in comparison to monotherapy or the combination of IFN-α plus ribavirin (Fried et al., 2002). The attachment of polyethylene glycol to IFN increases the molecular weight. This slows the rate of absorption from subcutaneous sites resulting the prolongation of serum half-life and the biological effects. It allows once weekly rather than thrice weekly injections. The combination of pegylated IFN-α plus ribavirin provided sustained responses of approximately 80% for patients infected with non genotype 1 and of about 50% for patients infected with genotype 1 (Fried et al., 2002). It has been suggested that patients with genotype 1 infection should be treated for 48 weeks with a standard dose of ribavirin. Patients with HCV genotypes 2 or 3 could be treated for 24 weeks with a lower dose of ribavirin (Hadziyannis et al., 2004).

1.10.2. Predicting response to treatment

The molecular mechanisms responsible for failure of IFN-α treatment are not well known, but evidence indicates that both viral and host factors are involved. It has been shown that younger females responded better to IFN-α therapy than older males (Booth et al., 2001). Hayashi *et al.* (1998b) reported that the rate of sustained response to IFN-α therapy was 33.3% in men aged 39 years or younger, 25.0% in men aged 40 years or older, 75.0% in women aged 39 years or younger, and 15.6% in women aged 40 years or older. Several clinical studies have reported that African American HCV infected patients had poor responses to treatment with IFN-α monotherapy or even with combination of IFN-α plus ribavirin (Layden-Almer et al., 2003; Reddy et al., 1999). Martinot-Peignoux *et al* (1995)

reported that the average age of complete responders was 35 years, which was 5 years younger than the average age of nonresponders. One possible explanation is that older HCV patients are likely to have more advanced liver disease, such as fibrosis and cirrhosis. The presence of an inadequate immune response in the elderly may be another important factor responsible for decreasing the number of successful responses to IFN-α treatment in older patients. In fact, absence of cirrhosis has been associated with an increased response to IFN-α therapy (Jouet et al., 1994). Higher body mass index (greater than 30 kg/m^2) and hepatic steatosis are associated with a poorer response to treatment (Bressler et al., 2003; McCullough, 2003).

Several viral factors have been reported to be predictive of responses to treatment. Viral genotype has been repeatedly shown to be the strongest predictive factor for treatment response, followed by pre-treatment serum viral titres and the duration and status of HCV infection at the time of treatment (Gao et al., 2004). It has been shown that being infected with genotypes 2 and 3 was predictive of a better response than with genotype 1. A higher pre-treatment viral load was associated with decreased rates of response (Hayashi et al., 1998a). It has been reported that patients with fewer than $2x10^6$ copies/ml of viral RNA showed a 44% response rate, compared to 27% for those with greater than $2x10^6$ copies/ml using combination therapy of IFN-α 2b plus ribavirin (Poynard et al., 1998). High quasispecies complexity at the beginning of treatment has been shown to be predictor of poor responses to treatment (Pawlotsky et al., 1998). 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-α treatment. However, greater diversity may simply be related to higher viral load.

Alcohol consumption is considered one of the important exogenous factors associated with poor responses to treatment. Patients should be advised not to drink alcohol while they receive IFN-α treatment. One possible mechanism is direct inhibition of IFN activated signals in hepatocytes by alcohol (Nguyen et al., 2000). The fibrosis score and its rate of progression have been suggested as independent predictors of response to IFN treatment (Banner et al., 1995). Patients with cirrhosis are more resistant to treatment than

those who have fibrosis, whereas patients with fibrosis are less responsive to treatment than those without fibrosis (Gao et al., 2004)

1.11. The background to the work in this thesis

Previous studies have reported differences in IRES efficiencies when comparing 5'UTR sequences from two or three different HCV genotypes. A few studies compared translation efficiency of genotype 1 and 3 and reported contradictory results. One study reported that the 5'UTR of GT3 is 50% less able to direct translation initiation that the 5'UTR of GT1 (Buratti et al., 1997), while another suggested the presence of similar translation efficiency (Collier et al., 1998). Patients infected with GT3 have a better response to treatment with interferon-α (IFN-α) either alone, or in combination with ribavirin. The causes and mechanisms of HCV resistance to IFN-α treatment are not understood. However, a lower translational efficiency of GT3 could contribute to the greater effectiveness of interferon therapy. Previous studies on translation efficiency measured the efficiency of translation of a single sequence as a representative of a genotype. No attempt had been made to ensure the sequences matched those actually found in patients so that it is not clear whether the observed IRES activity was genotype or actually only sequence specific. The effect of the core gene on translation of polyprotein by the HCV IRES is controversial. Several studies have found that including sequences encoding all or nearly full-length core protein in constructs appears to reduce the efficiency of reporter gene translation. It is not clear whether the core protein itself or the core encoding RNA sequence leads to suppression of IRES translation initiation. It was reported that the HCV core protein itself reduced the efficiency of HCV translation by binding to the IRES (Shimoike et al., 1999). In contrast, another study suggested that the core protein did not appear to have any specific effect on HCV IRES directed translation, and instead, it was reported that suppression of IRES activity resulted from an RNA–RNA interaction (Wang et al., 2000).

The significance of translational efficiency has not been explored in the clinical context. HCV replicates in the liver, yet no study has been reported in which translation efficiency has been related to serum viral loads and histological changes observed in the liver of infected patients. It is conceivable that a more efficient IRES may be correlated with a higher viral load and a greater ability to resist the inhibitory effects of interferon. The presence of a correlation between viral load and severity of histopathological changes in

the liver has been controversial. Several groups reported that higher RNA levels in serum were associated with the presence of severe liver disease (Adinolfi et al., 2001b; Gretch et al., 1994; Kumar et al., 1994). In contrast, no correlation was reported by other studies (Anand and Velez, 2004; Lee et al., 2001; Zeuzem et al., 1996).

HCV infection is characterised by a high incidence of chronicity. Our group had hypothesized that the liver might contain minor variants with lower translational activity which would allow sufficient protein expression for viral replication but insufficient levels to induce immune recognition of infected liver cells resulting in persistent infection. Reviewing the literature, only two reports were found in which the differences between quasispecies composition in the 5'UTR between serum and liver were reported (Cabot et al., 1997; Jang et al., 1999). The presence of 5'UTR variants specific to liver samples which were not found in serum samples from the same individuals has been reported (Jang et al., 1999). Previous work by P. Preikschat in our laboratory showed the presence of identical majority sequences of 5'UTR in matched serum and liver samples from 26 chronically HCV infected patients. However, majority sequence analysis is not able to detect minor variants in the quasispecies. Their detection would require other methods such as single-strand conformation polymorphism (SSCP) or sequencing of many cloned PCR products.

More detailed discussion of the background to each of the experiments performed is provided at the beginning of the relevant chapters.

1.12. Aims and objectives

The aims of this project were to carry out experiments to answer the following questions:

A) Is the translation efficiency of HCV genotype 3 consistently lower than that of genotype 1?

B) Does the translation efficiency of the HCV IRES correlate with viral load and liver histology in HCV infected patients?

C) Is there evidence that variants within the HCV quasispecies with low translation efficiencies exist in the liver?

The following objectives were achieved:

1. To construct 5'UTR and 5'UTR plus core gene clones matched with majority sequences obtained from a number of genotype 1a and 3a infected patients in order to examine the translation efficiencies and determine whether there is a consistent difference between genotypes.

2. To measure serum viral load using real time PCR and histological index of liver biopsies using the Ishak scoring system to investigate possible correlations between translation efficiency, viral load and liver histology.

3. To clone and sequence the 5'UTR region from serum and matched liver samples from 3 genotype 1a and 3 genotype 3a infected patients to analyse the quasispecies composition.

4. To employ the technique of SSCP to analyse the 5'UTR region in the same samples to confirm, or otherwise, the cloning and sequencing data.

Chapter 2

Materials and methods

2.1. Materials

2.1.1. Bacterial Strains

Plasmids were manipulated and propagated in the *Escherichia coli (E. coli)* competent Top 10F cells (Invitrogen) and Epicurian Coli XL1-Blue cells (Stratagene).

2.1.2. Vectors

pRL vector (Collier et al., 1998)	Kind gift from Professor Richard Elliott
pCRII-TOPO	Lifetechnologies

2.1.3. Synthetic oligonucleotides

Ordered from MWG-Biotech (Germany) and later from Sigma-Genosis Ltd

2.1.4. Kits and enzymes for RNA/DNA modification

Advantage -HF2 PCR kit,	BD Biosciences
AdvanTaq Plus PCR kit	BD Biosciences
QIAquick Gel Extraction Kit	Qiagen

QIAprepSpin Miniprep Kit	Qiagen
QIAmp RNA extraction kit, SuperFect transfection reagent,	Qiagen
Restriction Enzymes, CIP, T4 DNA ligase, NTPs	Roche
Reverse transcriptase, RNase H	Life Technologies
RNasin	Promega
Taq DNA Polymerase in Storage Buffer A	Promega

2.1.5. Mammalian cell lines and culture media

Laboratory stocks of human hepatoma cell lines (HuH7), baby hamster kidney (BHK-21) cells, baby hamster kidney cells expressing T7 RNA polymearse (BHKsinT7) and an African green monkey kidney cell line (CV1) were used.

Glasgow minimal Eagles Medium (GMEM), Dulbecco's modified Eagles Medium (DMEM), new born calf serum (NBCS), foetal calf serum (FBS), penicillin/streptomycin, L-glutamine, puromycin, non-essential amino acids and S.O.C medium were supplied by Life Technologies.

Phosphate buffered saline (PBS), PBS (A), versene, trypsin, L-broth, tryptose phosphate (TP) broth, were produced "in-house" by the media department.

2.1.6. Human sera

The sera were taken with written informed consent from the patients and consent for the study was given by West Glasgow Research Ethics Committee. Sera were stored in an MRC funded tissue store at Gartnaval General Hospital. Sera were separated within 4 hours of sampling by C.A. Smith, L. Conroy, A. Orr and K. Stewart. Sera were stored in small aliquots at -20°C or in liquid nitrogen.

2.1.7. Luciferase assay

Dual-Luciferase Reporter Assay	Promega

2.1.8. Antibodies

Donkey anti-Goat IgG, horse radish peroxide (HRP)	Promega
Goat anti-Luciferase antibody	Promega
Mouse anti-core (JM122)	Kind gift from Dr.J.McLauchlan
Rabbit anti-mouse HRP	Sigma
Mouse monoclonal anti-B Actin Clone AC-15	Sigma

2.1.9. Chemicals

All chemicals were obtained from Sigma Chemical Co., BDH Chemicals Ltd or Roche.

ECL, Hybond nitrocellulose membrane	Amersham
30% acrylamide, ammonium persulphate	Bio-Rad
Dried skimmed milk	Marvel

2.1.10. SSCP

Mutation Detection Enhancement Gel (MDE)	Cambrex
SilverXpress	Invitrogen

2.1.11. Commonly used solutions

Miniprep solution I	25mM Tris HCl (pH 8.0), 50mM glucose, 10mM EDTA (pH8.0)
Miniprep solution II	0.2M NaOH, 1% SDS
Miniprep solution III	3M potassium acetate, 5M acetic acid

Ethanol	100% and 70% diluted with distilled water
Agarose gel loading buffer	0.1M EDTA, 50% sucrose, 1μg/ml bromophenol blue.
L-broth	10g NaCl, 10g Bactopeptone and 5g yeast extract per litre.
L-broth agar	L-broth plus 1.5% (w/v) agar.
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.
Resolving gel buffer	1.5M Tris-HCl (pH 8.9), 0.4% SDS.
Running gel buffer	40mM Tris, 185mM glycine, 0.1% SDS.
Stacking gel buffer	0.5M Tris-HCl (pH 6.8), 0.4% SDS.
Stripping buffer	62.5 mM Tris-HClpH6.7, 2% SDS, 100 mM
Sample loading buffer for protein analysis	100 mM Tris-HCl pH6.9, 2% SDS, 10% Glycerol, 5% β-mercaptoethanol, 1μg/ml bromo-phenol-blue
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.
TE buffer	10mM Tris-HCl (pH 7.5), 1 mM EDTA
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. Virus stocks.

vTF7.3 is a recombinant vaccinia virus that expresses T7 RNA polymerase (Fuerst et al., 1986) and was obtained from Professor R.Elliott.

2.1.13. 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.

2.2. Manipulation of DNA and RNA

2.2.1. Small scale preparation of plasmid DNA (Minipreps)

2.2.1.1. Using QIAprep Spin Miniprep Kit

The extraction of plasmid DNA from bacterial culture was performed using the QIAprep Spin Miniprep Kit according to the manufacturer's 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. Two 1.5ml aliquots of culture were centrifuged at 13,000 rpm for 30sec, the supernatant removed and the cell pellet resuspended in 250 μl Pre-Lysis solution (Buffer P1). This was followed by the addition of 250 μl of Cell Lysis solution (Buffer P2), which was mixed gently and thoroughly by inversion until the solution was clear and viscous. 350 μl of Neutralising solution (Buffer N3) was added before centrifugation for 10 min at 13,000 rpm. The supernatant was transferred to a QIAprep column and was centrifuged for 1 min before the addition of 750μl of Wash solution (Buffer PE) and a further 1min centrifugation. DNA was eluted by the addition of 50μl dH_2O or elution buffer (Buffer EB). The plasmid DNA was stored at -20°C.

2.2.1.2. "In-house" Miniprep method

Colonies of transfected bacteria were inoculated into 2-3ml of LB broth containing ampicillin at 100μg/ml and incubated in a shaker at 37°C overnight. One ml of the overnight culture was transferred to a 1.5ml microcentrifuge tube and centrifuged at 13,000 rpm for 30 sec. After decanting the supernatant, the pellet was resuspended in 100 μl of ice cold Miniprep solution I and left at room temperature for 5 min. 200 μl of freshly prepared Miniprep solution II was added, mixed gently and left on ice for 2-3 min. 150μl of Miniprep solution III was added, mixed by few inversions and left on ice for 2-3 min. 150μl of phenol/chloroform was added to the mixture, vortexed and centrifuged at 13,000 rpm for 5 min. The aqueous phase was transferred to a fresh tube, 800μl of 100% ice cold ethanol was added to precipitate the plasmid DNA and centrifuged for 10 min. The pellet was then washed with 70% chilled ethanol, air dried and finally resuspended in 50μl dH_2O containing RNase at 20 μg/ml to remove any contaminating RNA.

2.2.2. Large scale preparation of plasmid DNA (Midipreps)

Large quantities of plasmid DNA were isolated from cells using the QIAfilter Plasmid Midi Kit as specified by the manufacturers instructions. 50ml of L-broth containing 100 μg/ml ampicillin was inoculated with bacteria and incubated with shaking overnight at 37°C. Cells were pelleted by centrifugation at 3,000 rpm for 15 min at 4°C. The supernatant was removed completely and the cell pellet resuspended in 4ml of cold Resuspension solution (Buffer P1) (containing RNase). 4ml of Cell Lysis solution (Buffer P2) was added and mixed by inversion before leaving at room temperature for no longer than 5 min. 4ml of chilled Neutralisation solution (Buffer P3) was added, immediately mixed by inversion and incubated in QIAfilter cartridge at room temperature for 10min. The QIAGEN-tip 100 columns were equilibrated with 4ml 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 (Buffer QC) were applied to the column and the DNA was eluted by the addition of 5ml of Elution solution (Buffer QF). The DNA solution was mixed with 3.5ml of room temperature isopropanol and the DNA precipitated

by centrifugation at 11,000 rpm for 30 min at 4°C. The DNA pellet was washed with 70% ethanol and centrifuged at 11,000 rpm for 10 min at 4°C. Finally, the DNA was air-dried and resuspended in 100µl of dH_2O and stored at -20°C.

2.2.3. Quantification of plasmid DNA and oligonucleotides

DNA was quantified by reading the optical density of a 10^{-2} dilution of miniprep or midiprep DNA at 260 and 280 nm using a Heλiosα version 4.55 spectrophotometer. The formula below was used to determine the concentration of nucleic acids.

$$\text{Concentration (µg DNA/ml)} = O.D._{260} \times N \times 100$$

O.D= value of 1 OD_{260}

N= OD reading

100= dilution factor

A value of 1 OD_{260} is equivalent to: 50 µg/ml for double strand DNA

40 µg/ml for single strand DNA

20 µg/ml for oligonucleotides (less than 30 bases in length)

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.

2.2.4. Restriction enzyme digestion of DNA

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. DNA fragments were analysed by agarose gel electrophoresis as described in 2.2.9.1.

2.2.5. Dephosphorylation of linearised plasmid DNA

Removal of the 5' terminal phosphates of digested vectors was carried out using calf intestinal phosphatase (CIP). Reactions were carried out in a total volume of 50µl containing 5µl of 10x CIP buffer (Roche) and 1 unit of CIP per 5µg of DNA. The reaction was incubated at 37° C for 60 minutes followed by inactivation of enzyme by adding 5 µl of 0.2 mM EDTA and further incubation at 65°C for 10 minutes. The linearised, dephosphorylated vector was purified as described in 2.2.9.2 in preparation for ligation.

2.2.6. Extraction of RNA from human sera

HCV RNA was extracted from human sera using the "QIAamp viral RNA" kit from Qiagen, as specified by the manufacturer's instructions. For lysis of virus particles, a 140µl aliquot of serum was added to 560µl of AVL buffer (viral lysis buffer) containing carrier RNA and mixed, then centrifuged for 1 min at 8,000 rpm. After incubation at room temperature for 10 min, 560µl of 100% ethanol was added and mixed. A 630µl aliquot of the solution was added to a QIAamp spin column and centrifuged for 1 min at 8,000 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 8,000 rpm 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 or 50 µl of pre-heated (80°C) RNase-free water (DEPC H_2O), incubation at room temperature for 1 min and centrifugation at 8,000 rpm for 1 min. The viral RNA was stored at -20°C until required.

2.2.7. Reverse transcription of RNA

2.2.7.1. 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 VtagRT 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

2.2.7.2. Superscript reverse transcriptase

Reverse transcription was performed using "SUPERSCRIPT II" reverse transcriptase from Life Technologies. In a total volume of 12.5µl, 40 pmoles of primer (Vtag RT) 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 then 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 II reverse transcriptase to give a final volume of 20µl. The reaction was performed at 42°C for 50 min and then inactivated at 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. Polymerase chain reaction (PCR)

2.2.8.1. Amplification of 5'UTR

Amplification of 5'UTR was carried out using DNA templates containing 5'UTR and core regions generated by P. Preikschat in our laboratory as described in 2.2.8.2. Second round

of PCR reactions were performed using the Clontech Advantage HF 2 PCR kit which contained AdvanTaq DNA polymerase, a small amount of a proofreading polymerase and TaqStart antibody to provide automatic hot-start PCR. For the second round PCR reaction, 1µl of the first round product was added to a 19µl reaction mixture containing HF2 PCR buffer, 0.2 mM each of dATP, dCTP, dGTP and dTTP, 5 pmoles of each primer and 0.5µl of the polymerase. PCR reactions were carried out under the following reaction conditions: a denaturation step of 94°C for 2 min and then thermal cycling of 94°C for 50 sec (strand separation), 56°C for 30 sec (primer annealing) and 72°C for 50 sec (strand elongation) for 25 cycles. Finally, the reaction was heated to 72°C for 6 min for a final extension step and cooled to 4°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".

2.2.8.2. Amplification of the 5'UTR and core

First round of PCR reactions were carried out by P. Preikschat using the Clontech Advan Taq Plus PCR kit as described here. 4µl of cDNA was added to a 16µl reaction mixture containing PCR buffer, 0.2mM each of dATP, dCTP, dGTP and dTTP, 5 pmoles of each primer and 0.5µl of the polymerase. PCR reactions were carried out under the following reaction conditions: a denaturation step of 94°C for 2 min and then thermal cycling of 94°C for 1min (strand separation), 67°C for 50 sec (primer annealing) and 72°C for 2min (strand elongation) for 38 cycles. Finally, the reaction was heated to 72°C for 6 min for a final extension step and cooled to 4°C.

The entire 5'UTR and core regions were re-amplified using PCR products generated by P. Preikschat as template. PCR was performed using the Advantage-HF2 PCR kit (Clontech). In this round, 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 a different set of primers. Following a 2 min denaturation step at 94°C, the reactions underwent thermal cycling of 94°C for 50 sec, 56°C for 30 sec and 72°C for 70 sec for 25

cycles. Finally, the reaction was heated to 72°C for 6 min for a final extension step and cooled to 4°C. For purification of the PCR product, PCR reaction volume was increased to 50µl.

2.2.8.3. Amplification of cDNA for quantitation of viral load

For quantitation proposes, the cDNA was amplified by PCR using a Applied Biosystems "5700 sequence detection system" and all PCR reagents and primers were obtained from Applied Biosystems. Extraction of RNA and RT were carried out as described previously. 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 (sense) primer, 1.225 pmoles of EMCHCV R (antisense) primer and 5.1 pmoles of EMCMGBP fluorescent probe as optimised by C.A. Smith previously. 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 and quantified RNA to generate a standard curve.

2.2.9. Electrophoretic separation and isolation of DNA

2.2.9.1. Agarose gel electrophoresis

Electrophoresis of DNA fragments produced by PCR or restriction enzyme digestion were carried out in horizontal slab gels containing 1-1.5% (w/v) agarose gel in 1x TBE or buffer containing a final concentration of 0.5µg/ml of ethidium bromide. DNA samples contained 0.2 volumes of loading dye. 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.

2.2.9.2. Purification of DNA from agarose gels

100 ml horizontal slab gels (140mm x120mm x 5mm) containing 0.8% (w/v) agarose and 0.5μg/ml of ethidium bromide were submerged in 1 x TAE buffer. 80 μl of PCR product with 20 μl gel loading dye were loaded and electrophoresed at 80-100 V for 60-90 minutes. DNA was examined, photographed and the band of interest excised using brief exposure to long wave ultraviolet light. The DNA was recovered using the "QIAquick Gel Extraction Kit" (Qiagen), which uses a silica gel based system. Three volume of binding buffer (Buffer QG) containing the silica gel matrix was added to one volume of sliced gel and incubated at 50°C until the gel had dissolved. The solution was applied to the column after adding x1 volume of isopropanol. The mixture was centrifuged at 13,000 rpm for 1 min. DNA was washed with 750μl wash solution (Buffer PE) and the centrifugation step repeated. The pellet underwent further centrifugation for 1 min and the DNA was eluted by the addition of 30 to 50 μl of elution buffer (Buffer EB). The DNA was stored at -20 C

2.2.9.3. 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 13,000 rpm. 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 2.2.9.4.

2.2.9.4. Ethanol precipitation

The DNA was recovered from the aqueous phase by precipitation by adding 2 ½ volumes of ethanol and 1/10 volume of 3M sodium acetate. After mixing, the solution was placed on dry ice for 15 min and the precipitated nucleic acids collected by centrifugation at 13,000rpm for 10 min. The pellet was washed with 70% ethanol and centrifuged at 13,000rpm for 3 min. The pellet was then air-dried and resuspended in 10-50µl of dH_2O.

2.2.10. Ligation of DNA fragments.

2.2.10.1. Ligation of DNA fragment into pRL vector (Collier et al., 1998)

Ligation reactions were carried out in a reaction volume of 10 µl containing 2 µl 5x ligation buffer (Gibco BRL), 1 unit T4 DNA ligase, 100 ng linearised vector, DNA insert and made up to final volume with dH_2O. Vector and "insert" DNA fragments were ligated in a 1:3 molar ratio. Reactions were incubated at 25°C overnight before transformation of competent *E.coli.*

2.2.10.2. Ligation of DNA fragment into pCRII vector (TA cloning)

Ligation reactions were carried out in a reaction volume of 6 µl containing 1 µl of salt buffer, 1 µl of DNA insert, 1 µl of pCRII vector and made up to final volume of 6 µl with dH_2O. Reactions were incubated at room temperature for 15 min and transformed into competent *E.coli* (Top10 cells) as described in 2.3.2

2.2.11. Single Stranded Conformation Polymorphism (SSCP) analysis

2.2.11.1. Generation of double stranded DNA.

Generation of double stranded DNA was carried out as detailed in section 2.2.8.1 using primers PP-AC5 and NCR4 (chapter 7, Table 7.2).

2.2.11.2. Purification of double stranded DNA

80µl of PCR product was run on a 1% agarose gel in TAE buffer as detailed in section 2.2.9.2 and the DNA recovered by QIAquick Gel Extraction Kit. The DNA was eluted in 50 µl of elution buffer. Concentration of DNA was estimated by running of 1µl of product on a 2% agarose TBE gel alongside a DNA mass ladder marker (Invitrogen).

2.2.11.3. SSCP of PCR product

50ng of PCR product was mixed with 2x volume of loading solution (0.05% xylene cyanol, 20mM EDTA, 95% formamide). This mixture was heated and denatured at 95°C for 5 min and snap-cooled on ice and then subjected to modified nondenaturing PAGE (0.5 x MDE) at 200V for 18 hours at room temperature. The 0.5x MDE gels were prepared as specified by the manufacturer's instructions and consisted of 25 ml of MDE Gel Solution (2X), 6 ml of 10 x TBE Buffer, 40 µl of TEMED, 400 µl of 10% APS made up to 100 ml with dH_2O and poured into a vertical gel apparatus (Protein II kit from Biorad). Two lanes of molecular weight markers (Marker VI) and a standard clone (2c) were also run on each gel to test reproducibility. SSCP distinguishes DNA according to conformation so the markers did not indicate the molecular weights of sample fragments. The gel was stained in l µg/ml ethidium bromide for 30 mins followed by 3 washing steps (using 0.6 x Tris-borate-EDTA buffer) and the bands were visualised under ultraviolet light.

2.2.11.4. Silver staining method

Silver staining carried out using Silver Xpress® Silver staining Kit (Invitrogen) as described in manufacturer's instructions. Briefly, the gel was incubated in "Sensitising Solution" (198 ml dH_2O and 2 ml sensitiser) for 20 min followed by 2 washing steps with dH2O for 5 min. The gel was incubated in gel "Staining Solution" (5ml of stainer A, 5ml

Stainer B and 90ml dH_2O) for 30 min followed by 2 washes for 5 min. Then the gel was incubated in "Developing Solution" (5ml Developer and 95ml dH_2O) for 3-15 min and when the desired intensity was reached, 5 ml "Stopping Solution" was added and incubated for 10min. Finally, the gel was washed 3 times with 200ml of H_2O for 10 min.

2.2.11.5. Extraction of DNA from silver stained gel

Extraction of DNA from silver-stained polyacrylamide gel was performed as described in (Laskus et al., 1998). The bands of interest were cut out of the gel using a clean disposable scalpel and then crushed with a pipette tip in a 1.5ml microcentrifuge tube. 300 ml of dH_2O water was then added and the tube was shaken for six hours at 37°C. The solution was filtered through a 0.45μm syringe filter, purified using a "QIAquick PCR Purification Kit" (Qiagen), and eluted in 30μl of elution buffer.

2.2.12. Automated DNA sequencing

Automated DNA sequencing was carried out using an ABI PRISM BigDye™ terminator Cycle Sequencing Ready Reaction Kit. Sequencing was carried out by Dr. G. Riboldi-Tunicliffe (Microarray and DNA Analysis Unit of the Sir Henry Wellcome Functional Genomics Facility, University of Glasgow).

2.3. Transformation of competent *E.coli* cells

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 30 seconds to "heat shock" the bacteria and then chilled on ice for another 2 min prior to the addition of 250μl S.O.C. medium. The cells were incubated at 37°C for 1 hour in a shaking incubator and then plated onto agar plates containing 100μg/ml ampicillin and incubated at 37°C overnight. Occasionally.

Epicurian Coli XL1-Blue (Stratagene) cells were used for transformation using the same methods as described for TOP10F cells. For blue/white selection, 40 µl of X-gal and 40 µl of IPTG were spread on each LB plate and incubated at 37°C until ready for use.

2.4. Maintenance of mammalian cells

2.4.1. BHK cells

Baby Hamster Kidney (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. Cells were passaged by trypsinization (trypsin:versine 1:1 v/v) (10ml) for 2-3 min at room temperature and resuspended in supplemented GMEM to allow seeding of new flasks. Cells were incubated in a humidified CO_2 incubator 37°C.

2.4.2. BHKsinT7 cells

BHKsinT7 cells expressing bacteriophage T7 RNA polymerase with puromycin selection, were treated as for BHK cells but were propagated in Glasgow minimal Eagles Medium (GMEM) supplemented with 1µg /ml of Puromycin.

2.4.3. HuH7 cells

HuH7 cells were cultured in Dulbecco's modified Eagles Medium (DMEM) containing 10% foetal calf serum (FCS), 4mM L-glutamine, 2% tryptose phoshate (TP) broth, 1% 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.

2.4.4. CV-1 cells

CV-1 cells (an African green monkey kidney cell line) were used to grow vTF7.3. The cell line was maintained in Dulbecco's modified Eagles Medium (DMEM) containing 10% foetal calf serum (FCS), 100 units/ml of penicillin/streptomycin at 37°C with 5% CO_2 in 160cm^2 tissue culture flasks.

2.5. Transfection of mammalian cells

2.5.1. Transfection of pRL constructs into BHKsinT7 cells

On the day prior to transfection, 6 x 10^4 cells/well were seeded onto in a 24-well plate. On the day of transfection the cells had grown to approximately 60-80% confluency. Plasmid DNA (1.5 μg/well) was mixed with serum-free medium ("Optimem") (100 μl) and subsequently mixed with Optimem (150 μl) containing lipofectin reagent (10μl) provided from R. Elliott's laboratory prepared as described by Rose *et al.*(1991). The mixture was incubated at room temperature for 15 min, during which time, the cells were washed once with Optimem. Following incubation for 15mins, the mixture was immediately pipetted onto cells. Cells were incubated at 37°C for 3 hours whereupon the transfection mix was removed and 1 ml of appropriate cellular medium was added to each well. Cells were harvested 16 hours after the addition of serum. For each construct, two or three replicate wells were transfected.

2.5.2. Transfection of pRL constructs using vTF7.3 vaccinia virus

Subconfluent monolayers of HuH7 cells in 24-well plates were infected with vTF7-3, a vaccinia virus expressing T7 RNA polymerase at 5 PFU/cell in 300 μl of serum-free medium (Optimem; Gibco-BRL) for 30 min at 37°C. The inoculum was removed, and the cells were washed once with Optimem. The cells were then transfected with plasmid DNA (1.5μg/well) in 300 μl of Optimem containing 15 μl of liposomes. Cells were incubated 3

hours at 37°C whereupon transfection mix was replaced with 1ml DMEM growth medium. The cells were harvested 16 h after addition of medium and assayed as described below. For each construct, two or three replicate wells were transfected. All constructs were tested in three separate experiments.

2.5.3. Growth and purification of vTF7.3

A confluent 175 cm^2 flask of CV-1 cells was inoculated with 10 µl of stock (1x10 pfu/µl) in 5ml DMEM-5% FCS and incubated at 37°C for 1 hour with agitation every 15 minutes. The medium was replaced with 20ml DMEM-5% and returned to 37°C for approximately 2 days until cytopathic effects were visualised. The infected cells were centrifuged at 3000 rpm and resuspended in 4ml of 10mM Tris-HCl pH 9.0 at 4°C. The cells were disrupted by 3 cycles of freeze thawing (5 minutes in dry ice, 5 minutes at 37°C) before nuclei were pelleted at 10000 rpm for 5 minutes at 4°C and the supernatant removed to a sterile tube. The pellet was resuspended in 4ml 10mM Tris-HCl pH 9.0, recentrifuged and the two supernatants combined. A one tenth volume of trypsin solution was added and the sample incubated at 37°C for 30 minutes with frequent mixing. Four millilitre aliquots were then layered onto 12ml 36% (w/v) sucrose in 10 mM Tris-HCl pH 9.0 and centrifuged at 13000 rpm in an SW29 rotor for 80 mins at 4°C. The supernatant was discarded and each pellet was resupended in 1ml Tris-HCl pH9.0, aliquoted and stored at -70°C

2.5.4. Titration of vTF7.3 stock virus

Thirty-five mm petri dishes containing 1 x 10^6 CV-1 cells were infected with dilutions of vTF7.3. Dilutions were carried out in a total volume of 1ml PBS and duplicate 100 µl volumes (for each dilution from 10^{-5} to 10^{-9}) were used for titration. The cells were incubated at 37°C for 60 min to allow the virus to adsorb with gentle agitation every 15 min. The inoculum was removed, 2ml DMEM containing 5%FCS added and the cells incubated at 37°C for 2 days. Then the medium was removed and the cells gently covered with CIDEX (a commercial glutaraldehyde preparation diluted 1:1 with PBS) for 3-4 hours to fix the cells. The CIDEX was removed and the cells stained with Giemsa for 10

mins before washing with tap water. Virus plaques were visible as regions within the purple-stained cell monolayer.

2.6. Dual-Luciferase reporter assay

Firefly luciferase is a 61 kDa monomeric protein and is widely used as a bioluminescent reporter because its enzyme activity is closely coupled to protein synthesis. Renilla luciferase has also become widely used as a genetic reporter. In the dual-luciferase reporter assay, the activities of firefly (*Photinus pyralis*) and renilla (*Renilla reniformis*) luciferase are measured sequentially from a single sample. This dual measurement from a single sample is made possible because firefly and renilla luciferases have dissimilar enzyme structures and substrate requirements. These differences make it possible to selectively discriminate between their respective bioluminescent reactions. Using the Dual-Luciferase assay kit, it was possible to quench the luminescence from the firefly luciferase reaction (the experimental reporter) while activating the luminescent reaction of *Renilla* luciferase (the control reporter). Photon emission is achieved through oxidation of beetle luciferin in a reaction that requires ATP, $Mg^{+2}$and O_2. Coenzyme A is incorporated into this reaction to provide more favourable reaction kinetics, resulting in an extended glow-type luminescent signal with greater intensity. The firefly luciferase reporter is measured first by adding luciferase Assay Reagent II (LAR II) to generate this glow-type signal. After quantifying the firefly luminescence, this reaction is quenched, and the renilla luciferase reaction is started by simultaneously adding Stop&Glo® Reagent to the same tube. The Stop & Glo® Reagent also produces a glow-type signal from renilla luciferase. This decays slowly over the course of the measurement.

2.6.1. Cell extract preparation

Dual-luciferase reporter assays were carried out according to the assay kit instructions (Promega)(Sherf, 1996). Cells were transfected and cultured for 16 hours. The cell culture medium was then removed and the cells were washed with 1ml PBSA. Passive Lysis Buffer (Promega) was diluted 1:5 in distilled H_2O to make a 1x solution and 100µl was

added to each well (24 well plate), after removal of the PBSA. The culture plate was agitated for 20 mins until the cells came off the surface of the wells. Cells were pipetted up and down until an even suspension was obtained, which was then transferred to a 0.5 ml microcentrifuge tube and centrifuged at 13,000 rpm for 1 min.

2.6.2. Luciferase quantification assay

Cell lysates, LAR II ® and Stop & Glo buffer were thawed at room temperature. 100 µl of LAR II was predispensed into disposable test tubes. Stop & Glo reagent was prepared by adding 20 µl of Stop & Glo 50 x substrate to 980µl of Stop & Glo buffer, sufficient for 10 reactions. The luminometer (Turner designs TD- 20/20, Promega) was programmed to read dual luciferase levels. 20µl of cell lysates was added to each test tubes together with the LAR II and mixed. The samples were then placed in the luminometer and the experimental firefly luciferase reading in relative light units (RLU) was taken by the luminometer. When prompted, 100µl of Stop &Glo Reagent was added to the tube and briefly mixed. The renilla luciferase reading was taken. The ratio of the firefly luciferase divided by the renilla luciferase was obtained and used as indicative of translation efficiency of 5'UTR.

2.6.3. Validation and normalisation of luciferase assay results

According to manufacturer's instructions, the relationship between the light output catalysed by the interaction of both renilla and firefly luciferases with their substrates is linear over at least a 5 logarithmic range. As shown in Figure 2.1, the renilla activity is within the linear range from somewhere below 10^{1} to 10^{5} RLU. Therefore, to ensure all renilla levels were safely within the linear range, values below 10 RLU were regarded as indicating an invalid transfection level. Furthermore within each experiment, mean renilla levels were calculated and any value >2 standard deviations (SD) below the mean were considered to represent an inadequate transfection for that clone in that experiment.

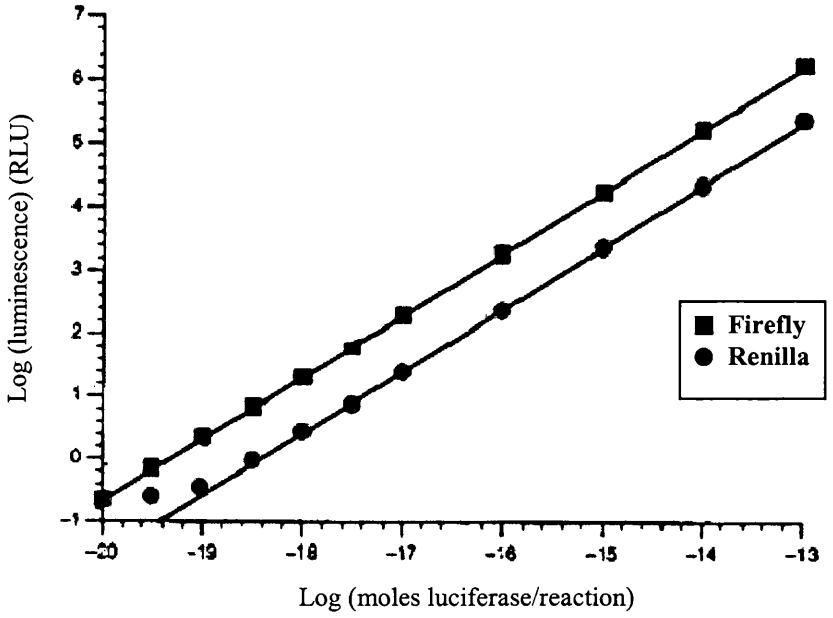


Figure 2.1. Comparison of the linear ranges of firefly and renilla luciferase in the dual luciferase assay.

As shown in this graph, the linear range of the firefly luciferase assay is seven orders of magnitude, providing detection sensitivity of <1 femtogram (approximately 10^{-20} moles) of experimental reporter enzyme. The renilla luciferase assay has a linear range of greater than five orders of magnitude and allows for the detection of approximately 30 femtograms (3×10^{-19}) of control reporter enzyme. Taken from manufacturer's catalogue (Dual-Luciferase ® Reporter assay system, Promega).

The firefly/renilla light output ratios obtained from matched 5'UTR and 5'UTR plus core clones of the H77c were taken arbitrarily as 100% where indicated. The mean firefly/renilla ratio obtained from each replicate was expressed as a percentage of the ratio from H77c.

2.7. Protein analysis by SDS-PAGE

Cells were washed with PBSA and lysed using lysis buffer. Lysates were centrifuged and boiled for 3 min after addition of SDS-PAGE denaturation buffer (sample loading buffer) (0.5ml) and allowed to cool on ice. Samples were subjected to SDS-PAGE using 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. Gels were poured and levelled using dH_2O 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 gels were immersed in 1x running gel buffer and samples loaded alongside "rainbow protein marker" (Amersham Bioscience) (7 µl/lane). Gels were run at 100 V for 3 hour or until the dye front ran off the gel.

2.7.1. Electroblotting to nitrocellulose membrane

Proteins were resolved by SDS-PAGE and were then transferred to a 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 1.5 hours at 4°C.

2.7.2. Immunodetection

Following transfer, membranes were incubated in PBS (A) containing 5% skimmed milk (Marvel) overnight to block non-specific binding of antibody. Membranes were washed in PBSA containing 0.05% (v/v) Tween-20, 6 times for 10 min and probed with the primary antibody diluted in 1% BSA in PBST for 3 hours at room temperature. Membranes were washed as previously and incubated with secondary antibody conjugated to HRP for 2h at room temperature. The membranes were washed for a final six times and the proteins detected using the Amersham enhanced chemiluminescence (ECL) system. The two supplied reagents, I and II in equal ratio, were mixed 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.

2.7.3. Stripping membrane for reprobing

Bound primary and secondary antibodies were removed by incubating the membrane in stripping buffer (0.28 mM β-Mercaptoethanol, 2% v/v SDS, 62.5mM Tris-Cl pH 6.7) at 50°C for 30 min with shaking followed by 3 x 10 min washes in PBS-T. Membranes could then be blocked and reprobed as previously described.

2.8. Computer software

Sequences were initially analysed and edited using the "Sequence Navigator" program (Applied Biosystems). Alignments and final editing of DNA and protein sequences were carried out using the programs PILEUP, PRETTY, GELASSEMBLE AND SEQED 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. EndNote 7.0 was used for creating a bibliography.

2.9. Statistical analysis

Statistical analysis was carried out in chapters 3 and 4 using both the Student t and Mann-Whitney tests. The Student t test is appropriate to test for the significance of observed differences between the means of 2 groups of measurements which fall into a normal distribution. One of the requirements for a valid t-test is that the standard deviations of the two groups being analysed should not be significantly different. If standard deviations do differ significantly, the Mann-Whitney test, a non-parametric test, can be used instead. It does not require the data to have similar standard deviations, nor to be normally distributed. In both tests, p value <0.05 was considered as significant but it should be noted that means have to differ by more to reach significance in the Mann-Whitney test.

The Spearman correlation coefficient test was used to analyse the correlations in chapter 5. This test is a non-parametric test based on ranking the data to measuring the strength of the relationship between pairs of variables. The correlation coefficient is restricted to values within the range -1.0 to $+1.0$ with $+1.0$ showing a perfect positive correlation.

Statistical caluculations were performed using SPSS version 11.0 (SPSS, Inc., Chicago IL).

Chapter 3

Comparison of translation efficiency of 5'UTRs derived from genotype 1 and 3 infected patients

3.1. Introduction

The 5' untranslated region (5'UTR) of the HCV genome is highly conserved and forms a stable secondary and tertiary structure (Brown et al., 1992; Honda et al., 1996a). It has been shown that the 5'UTR of HCV is able to direct translation of the open reading frame (ORF) by a cap independent internal ribosome entry mechanism mediated by an internal ribosome entry site (IRES) situated within the 5'UTR (Tsukiyama-Kohara et al., 1992; Wang et al., 1993).

Four major structural domains can be distinguished in the HCV 5'UTR (Figure 3.1). The short stem loop I is formed by nucleotides 5 to 20. Some reports suggested that it is not required for IRES activity and may have an inhibitory effect on translation (Honda et al., 1996b; Rijnbrand et al., 1995). In contrast, others have shown that deletion of the 5' end resulted in a reduction (two to threefold) in IRES activity (Friebe et al., 2001; Luo et al., 2003). It has been reported that stem loop II (nt 44-118) enhanced translation (Fukushi et al., 1994; Honda et al., 1996b; Reynolds et al., 1996). Lafuente *et al.* (2002), using transcripts that contained the sequences of the HCV IRES domains, showed specific RNA-RNA interactions between domains II and IV, and between domains II and IIIabcd demonstrating the importance of domain II for IRES activity. Stem loop III forms the core of the IRES and showed weak translational initiation activity even in the absence of domain II (Tsukiyama-Kohara et al., 1992). Stem loop IV, which includes the initiator AUG codon, may play an important role in regulation of the initiation of translation on the viral RNA. It has been shown that the stability of this domain correlated inversely with translation efficiency (Honda et al., 1996a).

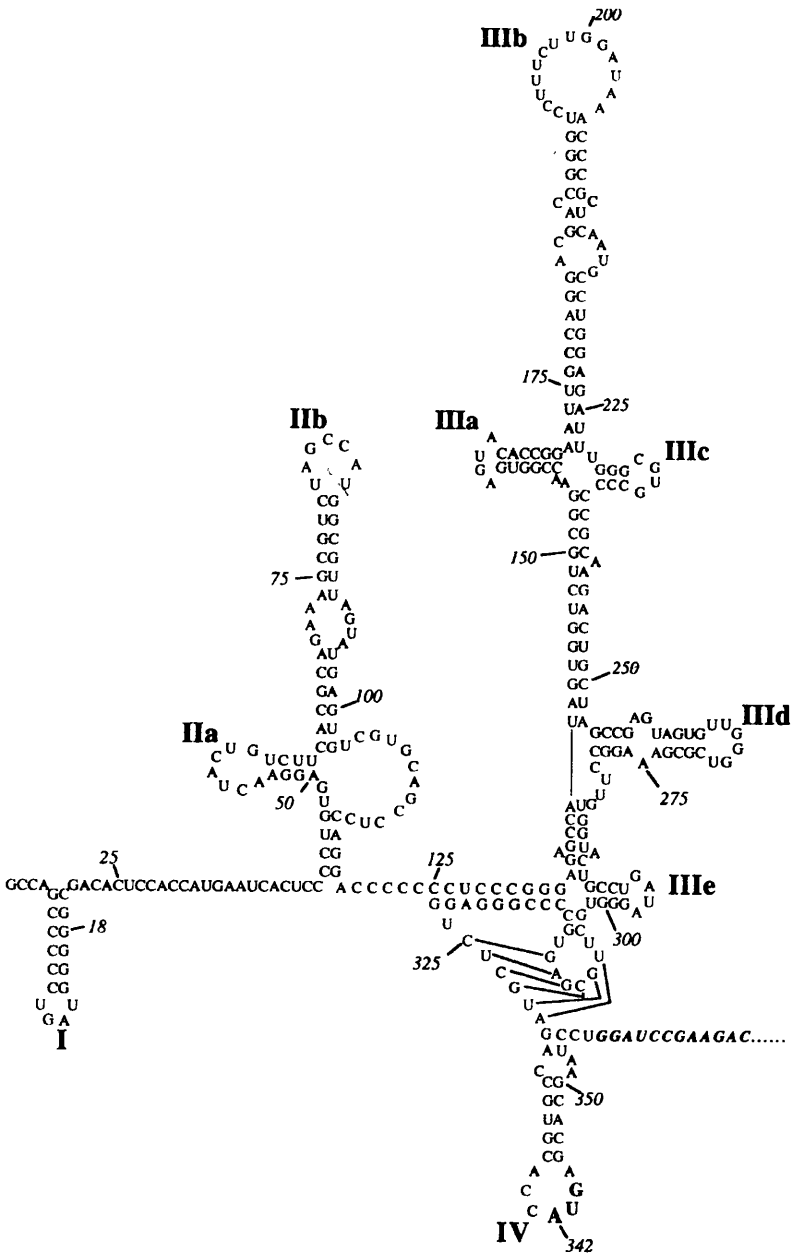


Figure 3.1. Predicted secondary structure of the HCV 5'UTR.

The sequence and numbering shown is that of GT1a (H77c)
and the structure is based on that proposed by Honda et al (1996).

Differences in the efficiency with which the IRES of different genotypes of HCV direct translation from reporter genes have been reported. The first study (Tsukiyama-Kohara et al., 1992) investigated the relative efficiencies of 5'UTRs (lacking 49 nts at the 5' end) from genotypes 1b and 2b in directing translation using an *in vitro* system. They reported a higher level of IRES activity by genotype (GT) 2b than GT 1b. Kamoshita *et al* (1997) used different cell lines to transfect bicistronic constructs containing the 5'UTR of genotypes 1b and 2b, and found similar results to those described by Tsukiyama-Kohara. They also noticed that ratio of 2b IRES activity to that of 1b was different in different cell lines. They suggested that this might be due to different distributions of host cellular factors between cell lines which interact with IRES elements. Honda *et al.* (1996b) reported that the full length 5'UTR sequence of a GT 1a strain was two fold more efficient in directing translation of the downstream core protein than GT 1b (N strain*) in vitro* and *in vivo*. Later, the same group (Honda et al., 1999c) showed that differences were due to specific differences at nts 34-35 of the HCV genome. Substitution of the AG dinucleotide sequence at nts 34 and 35 of HCV-N with GA (present in HCV-H) restored its IRES activity to that of the 1a. Buratti *et al.* (1997) studied the IRES activity of representative sequences from genotypes 1b, 2a and 3. They inserted full-length (nts 1-354) and truncated 5'UTR (nts 45 –342) sequences into a bicistronic expression vector containing human growth hormone (hGH) driven by the SV40 promoter and the CAT gene controlled by the HCV 5'UTR. The hGH levels were used to normalise the amount of cellular lysate used for measurement of CAT activity. They reported that the IRES activity of GT 3 was 50% of that observed for genotypes 1 and 2 in COS-1 cells. Comparable IRES activity was found between genotypes 1b and 2. Later, Collier *et al.* (1998) cloned 7 different 5'UTR sequences (nt 18 of the 5'UTR to nt 15 of the core coding sequence) from six different HCV genotypes (1a, 1b, 2b, 3a, 4a, 5a and 6a) into a bicistronic, dual luciferase reporter system. In contrast to Buratti's observation, similar IRES activity was found between genotypes 1a and 3a. They reported that the genotype 2b IRES had the highest activity in four cell lines (BHK-21, HeLa, HuH7 and HepG2) while GT 6a had the lowest activity.

Genotype is 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 GT 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). Correlation between translation efficiency and response to treatment is an interesting possibility for study. Saiz *et al.* (1999) studied the 5'UTR from 5 patients (a treatment responder and nonresponder from each of genotypes 1b and 3a and a non-responder infected with GT 2). They concluded that the response to IFN therapy and the activity of the IRES are independent. It was shown that the IRES activity of GT 2 and 3a was higher than that of GT 1b.

Bicistronic constructs have commonly been used to identify and assess the ability of candidate IRES elements. In these constructs, translation of the 3' reporter is driven by the putative IRES whereas the translation of 5' reporter is supported by a cap dependent mechanism. One potential problem is that the 3' reporter might actually be translated not from an IRES but from a monocistronic mRNA produced by splicing as suggested by Kozak (2003). A study published by Dumas *et al.* (2003) reported that the HCV 5'UTR DNA sequence contained a cryptic promoter starting at nucleotide 67 which was able to drive the expression of genes inserted downstream. This activity was not detected when the 5'UTR was replaced by the HCV 3'UTR or the poliovirus 5'UTR. They concluded that study of the translational activity of the HCV 5'UTR using bicistronic DNA constructs should be analysed at both translational and transcriptional levels. The presence of a strong promoter in the HCV 5'UTR cDNA could undermine the results reported by most studies using the bicistronic system to evaluate HCV IRES activity.

All previous studies measured the efficiency of translation of a single sequence as a representative of a genotype. Also, little or no attempt had been made to ensure the sequences matched those actually found in patients. It is not clear whether the observed IRES activity was genotype or actually only isolate specific. In the present study, I compared the ability of a number of GT 1 and 3 5'UTR sequences derived from patient samples to initiate translation in cell culture using a dual luciferase reporter system

(Collier et al., 1998) to find out whether there is a consistent difference between the IRES activity of GT 1 and GT 3. Also, it was possible to study the effect of individual substitutions in 5'UTR sequences obtained from different patient on IRES activity. I also conducted an experiment to ensure the results I obtained could not be explained by the present of a cryptic promoter in the DNA encoding the HCV 5'UTR.

The work presented in this and the following chapter was based on initial work carried out by P. Preikschat in our laboratory. She amplified the 5'UTR and core regions from serum and liver samples obtained from 26 patients. Then, all PCR products were sequenced and majority sequences obtained as will be discussed in following sections. This allowed me to select specific variants for the translation study. I also used patients' PCR products made by her as templates for re-amplification of 5'UTR and 5'UTR plus core regions for translation studies (discussed in chapter 4).

3.2. Study patients and samples

Liver biopsies and serum taken at the same time from 26 patients with chronic HCV infection, including 15 infected with GT 1a and 11 with GT 3a, were used in this study. All of these patients had been referred to the liver clinic at Gartnaval General Hospital, Glasgow for assessment and possible treatment of chronic hepatitis C. Informed written consent for taking the samples was obtained from the patients and the study was approved by West Glasgow Ethics Committee. None of the patients was on interferon based therapy at the time when samples were collected.

3.3. Patients' majority sequences

The following steps were carried out by P. Preikschat. PCR products corresponding to the entire 5'UTR and the core coding region were amplified by RT-PCR from the serum and liver of 11 HCV GT 1a and 15 GT 3a infected patients using primers shown in Table 3.1 using conditions described in section 2.2.8.2. The PCR products were sequenced directly. I aligned the majority sequences of the 5'UTR and the 5' end of core (nts 1-360) obtained

Primer	*Sequence 5' to 3'*	*Position*	*Use*	*GT*
PP-AC5 (s)	TTGCTGGATCCGCGACACTCCACCAT	18- 33	PCR 1	1,3
PP1-REV (as)	AGCAAGGATCCCCTCATACACAATACT	988- 972	PCR 1	3
PP3-REV (as)	AGCAAGGATCCGCCTCGTACACAATACT	988- 971	PCR 1	1
MMCG1 (s)	TGAGGATCCGCCAGCCCCCTGATGGGGGCGACACTCCACCAT	1 – 33	PCR 2	1
PP5UTR (s)	TGAGGATCCACCTGCCTCTTACGAGGCGACACTCCACCAT	1 – 33	PCR 2	3
PP5UTR-R (as)	TGAGGATCCAGGAAGTGTGCTCATGTTGC	356-337	PCR 2	3
MM6 (as)	TGAGGATCCAGGATTTGTGCTCATGATGC	356-337	PCR 2	1
MM8 (as)	TGAGGATCCAGGAAGTGTGCTCATGGTGC	356-337	PCR 2	3
MMCG2 (as)	TGAGGATCCGGCTGAAGCGGGCACAGT	914-897	PCR 2	1
MMCG4 (as)	TGAGGATCCAGGATTCGTGCTCATGGTGC	356-337	PCR 2	1
PPCORE-R (as)	TGAGGATCCACTAGCTGCTGGATGAAT	914-897	PCR 2	3
PP2-REV-RT	GACCAGTTCATCATCATATCC	1324-1304	RT	1
MSQ10	GCCATTCGGTGTCCTGAGAG	1300-1280	RT	3

Table 3.1. Oligonucleotide primers for RT-PCR of GT 1a and 3a 5'UTR region.

Nucleotide numbering according to H77c (AF 011751) for
GT 1 and NZL1(D17763) for GT 3 samples.
"s"= sense, "as"= antisense
BamHI restriction site is underlined.

from GT 1 and 3 infected patients. These are shown in the Appendix (Figures A-1 and A-2). No differences were observed between majority sequences derived from serum and liver taken from any individual patient.

3.3.1. Majority sequence of GT 1 samples

Figure 3.2.A summarises the differences between majority sequences obtained from GT 1a infected patients and the H77c reference sequence (EMBL accession no: AF011751). In the majority sequences of the 11 GT1 samples, no deletions or insertions in the 5'UTR were observed and two of the patient sequences matched that of the reference sequence H77c (patients MH and MA). The most frequent differences were seen at position 204 where cytosine was present in 5 and adenosine in 6 patients. Guanosine was present at position 107 in 8 and adenosine in 3. Most of the sequences presented in Figure 3.2.A differed from H77c by 1 nt (patients OS, BK, PE, SA, CD, TW), although differences of 2 nt (LA, BH), and 3 nt (BA) were noted. The positions of the differences observed are shown related to the proposed secondary structure of the 5'UTR (Brown et al., 1992; Honda et al., 1996a) in Figure 3.3.A. Seven different majority sequences were selected for study of translation efficiency as indicated in Figure 3.2.A.

3.3.2. Majority sequence of GT 3 samples

For GT3 samples, NZL1 (EMBL accession no: D17763) was chosen as the reference sequence and no deletions or insertions were present in the GT3 majority sequences (Figure 3.2.B). Similar to results obtained with GT1, the most frequent difference from the reference sequence occurred in stem loop IIIb (Figure 3.3.B). Position 201 (equivalent to 203 in H77c) in loop IIIb was the most variable in GT 3 with guanosine in 9 samples and adenosine in 6. A substitution of cytosine to uracil at position 202 (equivalent to 204 in H77c) was observed in 3 of 15 samples. One sample (JN) was identical to NZL1, 11 differed by a single nt and 2 differed by 2 nt. Nine different sequences were used for further amplification and cloning (including 1 matching NZL1) as highlighted in Figure 3.2.B

A

GT1 Patients	Tissue	nt-Position 43	107	204	233	243	248	340
OS	S			A→C				
	L			A→C				
BK	S			A→C				
	L			A→C				
TW	S			A→C				
	L			A→C				
BH	S		G→A	A→C				
	L		G→A	A→C				
BA	S	C→U		A→C			U→C	
	L	C→U		A→C			U→C	
SA	S		G→A					
	L		G→A					
LA	S		G→A		G→R			
	L		G→A		G→R			
PE	S					A→G		
	L					A→G		
CD	S							C→U
	L							C→U
MH	S							
	L							
MA	S							
	L							

B

GT3 Patients	Tissue	nt-Position 34	89	119	121	203	204	243	340
MO	S					G→A			
	L					G→A			
GR	S					G→A			
	L					G→A			
MF	S					G→A			
	L					G→A			
MW	S					G→A			
	L					G→A			
BJ	S					G→A			
	L					G→A			
JA	S					G→A		G→A	
	L					G→A		G→A	
BC	S						C→U		
	L						C→U		
CP	S						C→U		
	L						C→U		
RI	S	G→A	C→Y				C→U		
	L	G→A	C→Y				C→U		
FV	S			A→C					
	L			A→C					
LJ	S			A→C					
	L			A→C					
RJ	S				C→U				
	L				C→U				
ME	S								A→C
	L								A→C
SJ	S							G→A	A→C
	L							G→A	A→C
JN	S								
	L								

Figure 3.2. Differences between 5'UTR majority sequences obtained from patients.

Sequences derived from serum and liver of 11 patients infected with GT 1(A) and 15 with GT 3 (B) with correspondent reference sequences. H77c was chosen as reference sequence for GT1 and NZL1 for GT 3. Numbering in both tables is according to the H77c sequence. Sequences derived from serum and liver samples from each patient are indicated by S and L respectively. "Y" indicates a mixture of cytosine and uridine. "R" indicates a mixture of adenosine and guanosine. Highlighted samples from each genotype were used for further cloning.

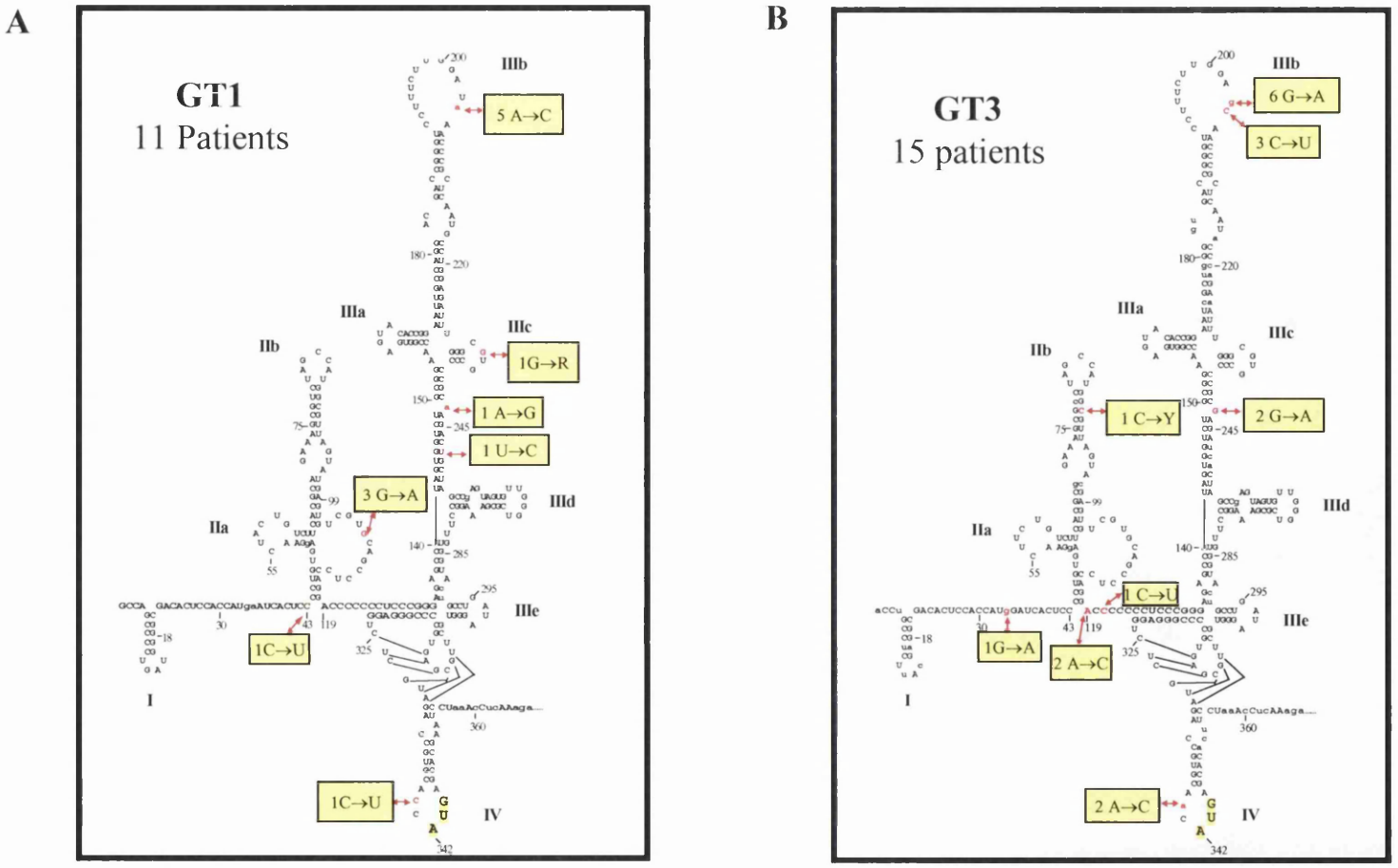


Figure 3.3. Position of nucleotide differences in predicted secondary structure of the HCV 5'UTR.
Differences in majority sequence in GT1(A) and GT3 (B) infected patients. H77c and NZL1 sequences were chosen as reference sequences for GT1a and GT3a samples respectively. The numbers inside boxes show the frequency of substitutions at the indicated position. Secondary structure and numbering are based on those of Honda et al (1996a). "Y" represents a mixed position of cytosine and uridine. "R" represents a mixed position of adenosine and guanosine.

3.4. Amplification of the 5'UTR region for study of relative translational activities of different sequences

I designed oligonucleotides to amplify the entire 5'UTR of HCV and 15 nucleotides of the core gene. The sequences of these primers and the nucleotide positions to which they bind can be seen in Table 3.1. *BamHI* restriction endonuclease sites were incorporated into the 5' ends of the primers for second round amplification to facilitate future cloning. Samples from 7 GT 1 infected patients representing each majority sequence seen in the study were chosen for cloning. I used PCR products from the RT-PCR of 5'UTR and core carried out by P. Preikschat as template to amplify the 5'UTR (nts 1-356) using sense primer MMCG1 and antisense primer MMCG4. For amplification of sample CD, which had 2 substitutions of C340U (cytosine to uracil at nucleotide position 340) and G350A in the majority sequence, primer MM6 was used as antisense primer in the second round of PCR. In GT 3 samples, 9 unique sequences were identified. Amplification of GT 3 samples was carried out using primers PP5'UTR and PP5'UTR-R, except samples ME and SJ which had a substitution of adenosine to cytosine at position 340 so a different antisense primer (MM8) was used. High fidelity DNA polymerase (HF2, Clontech) was used to reduce errors resulting from misincorporation of bases by *Taq* DNA polymerase.

3.5. Construction of plasmids containing the 5'UTR (pRLN)

The empty pRL vector (Collier et al., 1998), kindly provided by R.Elliott, contained the renilla luciferase and firefly luciferase genes flanking a unique *BamHI* restriction site. This site was used for insertion of PCR products containing the 5'UTR (Figure 3.4). *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 pRL vector were cleaved with *BamHI*, purified and then ligated together and transformed into *E.coli*. Individual clones were cultured overnight and plasmid DNAs were extracted using an in-house miniprep method as described in chapter 2.2.1.2. Clones were selected for the presence of the inserted PCR product by digestion with *BamHI*. This released a 356 base fragment. Using the one restriction site (*BamHI*) presented problems in that both orientations of insert are possible.

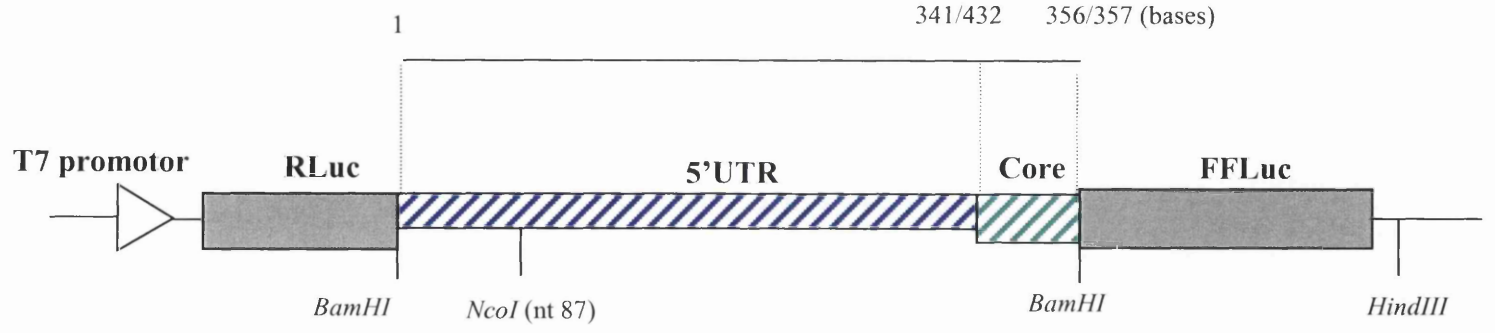


Figure 3.4. Structure of the bicistronic, dual luciferase reporter plasmid containing the 5'UTR and 15 nts of core (pRLN).

The open triangle represents the T7 promoter. The two luciferase reporter genes are indicated by grey boxes; RLuc, renilla luciferase and FFLuc, firefly luciferase. The HCV cDNA corresponding to 5'UTR and core sequences are shown by blue and green hatched boxes respectively. Restriction enzyme cleavage sites used for DNA manipulation are indicated under the cDNA.

However, the orientation of the insert within the vector was determined by cleavage with *NcoI/HindIII*. A single *NcoI* site was identified in the 5'UTR sequence. A unique *HindIII* site was present at the 3' end of the firefly luciferase reporter gene. In total, 275 individual clones were screened for the presence and orientation of the insert. DNAs from clones with the correct orientation of insert were extracted using a Miniprep Kit (Qiagen) as described in chapter 2.2.1.1. and sequenced in both direction using primers 5'luc (5' GCAAGAA GATGCACC TGATG 3') as sense and 3'luc (5' GCGTATCTCTTCATAG CCTT 3') as antisense primer which bind to renilla and firefly luciferase reporter genes respectively. Nucleotide sequences were analysed using the "Sequence Navigator" program (Applied Biosystems) and aligned using the "PILEUP" and "PRETTY" programs [Genetic Computer Group (GCG) Wisconsin Programme Package]. Individual clones from each sample were sequenced until one was found which matched exactly that obtained by direct sequencing of the PCR product from the same patient sample. Construction of a matched clone from patient ME (GT3) was not successful. In total, 65 individual clones were sequenced to identify 16 pRLN matched clones which could be used for translation studies.

3.5.1. Reconstruction of pRLN clones

One problem encountered during the construction of the pRLN matched clones was the occurrence of substitutions particularly at the "left hand" end of the 5'UTR, at the primer binding site, without any other mismatch with the majority sequence from the relevant patient. This would appear to have been due to random misincorporation of bases during primer synthesis. To try and resolve the problem without expensive sequencing of additional clones, a rebuilding strategy was carried out. As shown in Figure 3.5, clone 1 which matched the majority at the 5' *NcoI* restriction site, and clone 2, which matched sequence 3' to the *NcoI* site, were digested with *NcoI /HindIII*. This released a short fragment containing most of the 5'UTR and the firefly luciferase gene and a long fragment containing the first 90 nts of the 5'UTR and the remainder of the pRL vector except for the firefly luciferase gene. Following purification by agarose gel electrophoresis, the long fragment from the first clone was ligated to the short fragment

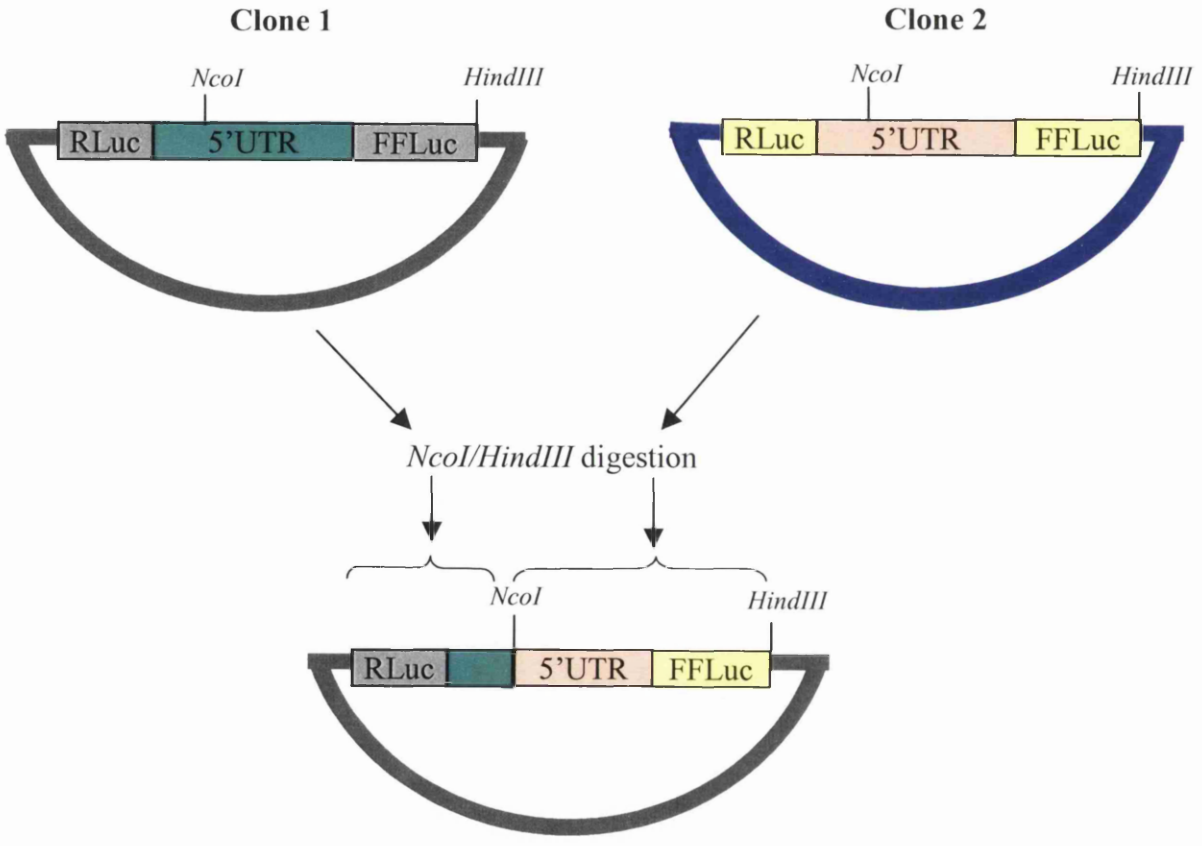


Figure 3.5. Reconstruction of pRL clones containing the 5'UTR of HCV isolates by using fragments from two clones.

The *NcoI-HindIII* restriction fragment including the Rluc vector sequence was taken from clone 1 and the downstream *NcoI-HindIII* fragment including the firefly gene was obtained from clone 2.

from the second clone. The resulting clone was then sequenced and found to match the majority sequence.

3.6. The effect of the first 17 nt of the 5'UTR on IRES activity

At the start of this study, the bicistronic dual reporter vector (pRL) containing the 5'UTR from H77c was obtained from P. Preikschat. Preliminary results of luciferase assays indicated that the IRES activity of most of the samples was higher than the activity of the clone thought to contain the entire IRES of H77c (Figure 3.6.A.B). Later, sequencing of this clone revealed the absence of 17 nucleotides comprising most of domain I of the 5'UTR. In order to include these 17 nucleotides, pRLN H77c was reconstructed as described in section 3.5.1 using the existing pRLNH77c clone and a clone from patient PE (pRLNPE) and the resultant construct was sequenced and shown to match H77c. Both constructs were transfected into BHKsinT7 cells and translation activity was analysed. The relative translation efficiency of the construct lacking the first 17 nts was 40% of that of the full length 5'UTR (Figure 3.6.C.D). The construct pRLNH77c containing the full 5'UTR was used as the arbitrary 100% standard in all subsequent experiments.

3.7. Relative translation activities in BHKsinT7 cells

5'UTR isolates including H77c were tested for translation activity in BHKsinT7 cells. In this system, transcription of plasmid DNA is controlled by the T7 promoter. The mRNA for the renilla luciferase gene is translated by a cap-dependent mechanism and expression of firefly luciferase by the HCV IRES mechanism. Each sample was tested on three different occasions. The first experiment was carried out in triplicate and second, third, forth and fifth experiments in duplicate. During analysis of the data, it was noticed that the renilla and firefly luciferase activities obtained in the third experiment were noticeably lower than those achieved during other experiments (Table 3.2).

Applying the criteria described in methods (section 2.6.3) for evaluation of results of luciferase assays, renilla luciferase light output obtained from both replicates of pRLN

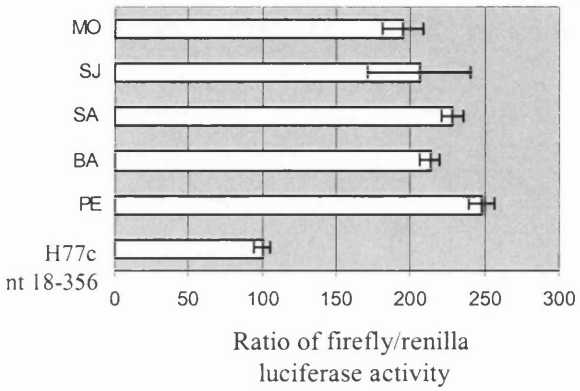
Figure 3.6. The effect of the first 17 nucleotide of H77c on IRES activity in BHKsinT7 cells.

A) Firefly and renilla luciferase light outputs (RLU) obtained from pRLN H77c (nt 18-356) and 5 isolates. Comparison of relative IRES activities of the 5 isolates with H77c lacking the first 17 nucleotides (B). C) Firefly and renilla luciferase light outputs obtained from constructs containing full length (nt 1-365) and truncated (nt 18-356) 5'UTR. D) Comparison of IRES activity obtained from clone pRLNH77c (nt18-356) which was arbitrarily assigned as 100% with full length H77c. Data are presented as means +/- standard deviation (SD).

A

Sample	GT	Firefly output (RLU)	Renilla output (RLU)	Firefly/ Renilla ratio
H77c (nt 18-356)	1	1446	358.3	4.0
H77c (nt 18-356)	1	1276	344.6	3.7
H77c (nt 18-356)	1	1315	359.3	3.7
PE	1	5765	609.8	9.5
PE	1	5145	528.3	9.7
PE	1	3455	375.5	9.1
BA	1	5025	597.0	8.4
BA	1	3828	479.4	7.9
BA	1	2966	372.9	8.0
SA	1	6595	783.1	8.4
SA	1	5161	568.6	9.0
SA	1	4405	512.1	8.6
SJ	3	3894	450.6	8.6
SJ	3	3523	413.0	8.5
SJ	3	2267	358.9	6.3
MO	3	3814	553.0	6.9
MO	3	4058	546.9	7.4
MO	3	3743	471.2	7.9

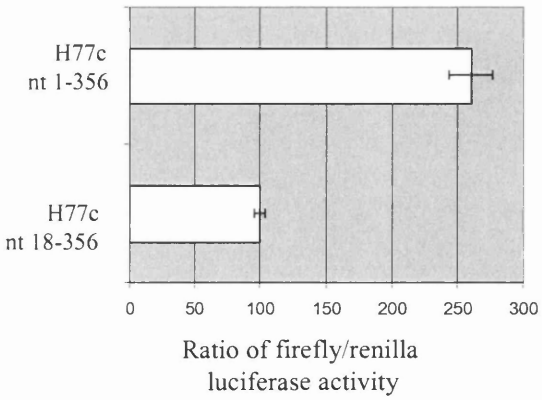
B



C

Sample	GT	Firefly output (RLU)	Renilla output (RLU)	Firefly/ Renilla ratio
H77c (nt 18-356)	1	6022	379.6	15.9
H77c (nt 18-356)	1	5037	341.6	14.8
H77c (nt 18-356)	1	4973	327.4	15.19
H77c (nt 1-356)	1	7494	187.9	39.87
H77c (nt 1-356)	1	6192	174.5	35.49
H77c (nt 1-356)	1	5715	163.2	35

D



H77c in the second experiment was more than 2 standard deviations from the mean renilla activity of all replicates in the experiment. These results were therefore considered invalid. The first two experiments did not include clones obtained from patients BH, RJ and JN. Therefore these clones were tested later along with the H77c clone. Primary data including firefly and renilla luciferase activities are shown in Table 3.2. Data omitted from further calculations because they did not meet the criteria set down in section 2.6.3. are shown in red.

It was impossible to compare IRES activities between experiments because of the lack of a valid H77c reference result in experiment 2. It was decided to analyse the mean firefly/renilla luciferase ratios obtained from each sample without normalisation to H77c taking each experiment separately. Figure 3.7 summarises the results for all 5 experiments. It is notable that, among GT1 isolates, the lowest translational activity was observed in the construct derived from patient OS, which had one substitution of A204C in loop IIIb. Similarly, clones obtained from patients RJ (C121U) and RI (G34A and C204U) had consistently lower activity than other GT3 samples.

3.8. Relative translation activities in HuH7 cells

The IRES activities of 8 GT 1 and 8 GT3 isolates were compared in HuH7 cells. The plasmids were transfected into cells previously infected with vTF7-3, a recombinant vaccinia virus which expresses T7 RNA polymerase to allow cytoplasmic transcription of the bicistronic mRNA. All samples were tested in 3 different experiments. The criteria for validation of results were met for all replicates in all experiments except those indicated by red in Table 3.3. The mean firefly/renilla luciferase ratios obtained from all isolates were normalised against H77c in each experiment.

The relative IRES efficiency of the 8 GT3 isolates varied from 60% to 120% of that of H77c as illustrated in Figure 3.8. Relative IRES activities obtained from RJ (C121U) and RI (G34A and C204U) were consistently lower in comparison to other GT3 isolates in both cell lines tested. However, The 5'UTR derived from patient OS with a substitution

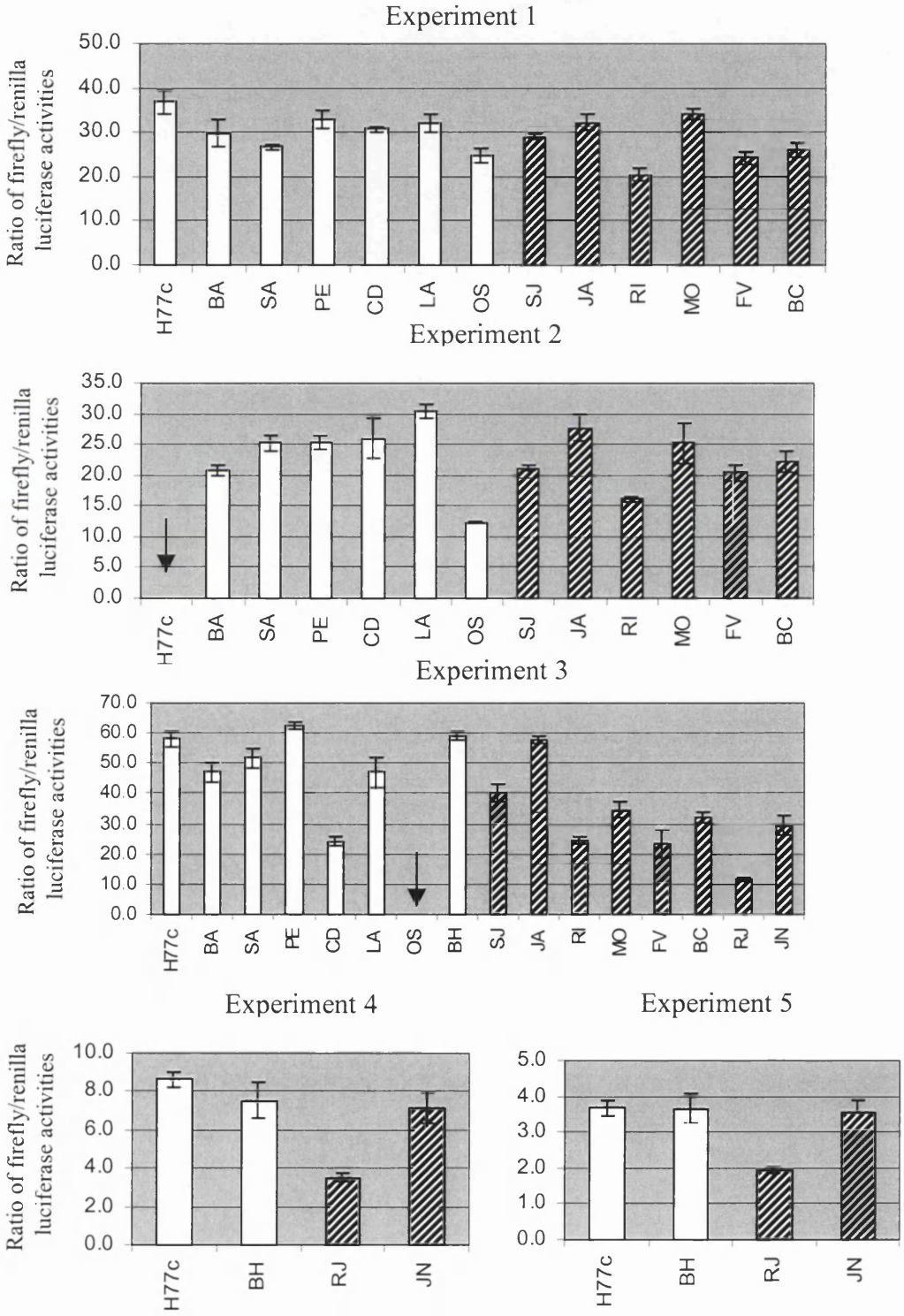
Table 3.2. Firefly and renilla luciferase light outputs in BHKsinT7 cells.

Firefly and renilla luciferase light outputs (RLU) obtained from 5 experiments are shown. Results are shown in the order in which experiments were performed. pRLN constructs from patients BH, JN and RJ were tested separately along with H77c. Renilla values not meeting criteria for acceptance of results are shown in red. ND= not done

		Experiment 1				Experiment 2				Experiment 3			
Sample	GT	Firefly output (RLU)	Renilla output (RLU)	Firefly/ Renilla ratio	Mean of ratios	Firefly output (RLU)	Renilla output (RLU)	Firefly/ Renilla ratio	Mean of ratios	Firefly output (RLU)	Renilla output (RLU)	Firefly/ Renilla ratio	Mean of ratios
H77c	1	7494.0	187.9	39.9		1251	45.8	ND		887.7	14.8	60.1	
H77c	1	6192.0	174.5	35.5	36.8	1174	45.5	ND	ND	910.0	16.2	56.2	58.1
H77c	1	5715.0	163.2	35.0									
BA	1	6320.0	190.2	33.2		2403	118.8	20.2		907.6	18.4	49.3	
BA	1	4661.0	169.2	27.6	29.8	2103	97.9	21.5	20.9	788.6	17.6	44.8	47.1
BA	1	5221.0	181.6	28.8									
SA	1	9481.0	358.7	26.4		3084	126.8	24.3		1100.0	20.3	54.1	
SA	1	8756.0	318.2	27.5	26.7	2865	110.0	26.0	25.2	1043.0	21.2	49.3	51.7
SA	1	7977.0	306.3	26.0									
PE	1	8218.0	233.9	35.1		3630	139.2	26.1		1418.0	22.4	63.3	
PE	1	6956.0	224.6	31.0	33.0	3192	130.4	24.5	25.3	1549.0	25.2	61.6	62.4
PE	1	6982.0	212.8	32.8									
CD	1	9981.0	319.1	31.3		2249	95.1	23.7		332.1	13.3	25.1	
CD	1	8870.0	285.8	31.0	30.7	3457	122.4	28.3	26.0	365.2	15.7	23.2	24.2
CD	1	8765.0	293.4	29.9									
LA	1	6758.0	220.4	30.4		2238	71.4	31.3		969.0	19.2	50.4	
LA	1	7094.0	205.8	34.5	32.2	3531	119.0	29.7	30.5	749.9	17.2	43.6	47.0
LA	1	7200.0	227.5	31.6									
OS	1	2863.0	115.3	24.8		1615	131.0	12.3		110.3	5.8	ND	
OS	1	3448.0	124.0	27.8	25.8	1590	128.3	12.4	12.4	125.6	7.9	ND	ND
OS	1	3246.0	131.3	24.7									
BH		ND	ND			ND	ND			1014.0	16.9	60.2	
BH		ND	ND			ND	ND			1045.0	18.0	58.0	59.1
SJ	3	4699.0	157.7	29.8		2298	110.6	20.8		688.5	16.2	42.4	
SJ	3	4430.0	157.0	28.2	29.1	2520	117.5	21.5	21.1	508.6	13.3	38.2	40.3
SJ	3	4174.0	143.2	29.2									
JA	3	4652.0	151.4	30.7		2445	93.1	26.3		919.3	15.6	58.8	
JA	3	5126.0	161.0	31.9	32.3	2663	91.3	29.2	27.7	894.6	15.8	56.8	57.8
JA	3	5182.0	151.6	34.2									
RI	3	4592.0	215.7	21.3		2007	126.4	15.9		551.0	21.6	25.5	
RI	3	4840.0	229.0	21.1	20.4	2195	133.1	16.5	16.2	389.0	16.5	23.5	24.5
RI	3	4291.0	227.3	18.9									
MO	3	6103.0	183.8	33.2		2865	124.8	23.0		716.9	19.5	36.7	
MO	3	5438.0	161.7	33.6	34.1	3431	124.3	27.6	25.3	478.7	14.7	32.7	34.7
MO	3	5565.0	157.1	35.4									
FV	3	5465.0	230.5	23.7		1126	57.7	19.5		355.8	13.3	26.8	
FV	3	5192.0	224.0	23.2	24.2	1251	58.7	21.3	20.4	220.5	10.9	20.3	23.5
FV	3	4782.0	185.3	25.8									
BC	3	2703.0	109.1	24.8		1821	77.6	23.5		605.9	18.1	33.5	
BC	3	2734.0	98.4	27.8	26.0	1753	83.2	21.1	22.3	446.3	14.6	30.7	32.1
BC	3	2601.0	102.2	25.5									
RJ		ND	ND			ND	ND			196.5	16.0	12.3	
RJ		ND	ND			ND	ND			174.2	15.3	11.4	11.9
JN		ND	ND			ND	ND			334.4	12.3	27.2	
JN		ND	ND			ND	ND			351.3	11.2	31.5	29.3
		Experiment 4				Experiment 5							
H77c	1	2517.0	301.0	8.4		514.9	133.7	3.9					
H77c	1	2365.0	264.2	9.0	8.7	443.0	127.1	3.5	3.7				
BH	1	1495.0	183.5	8.1		735.9	186.4	3.9					
BH	1	1161.0	167.4	6.9	7.5	559.9	166.6	3.4	3.7				
RJ	3	903.8	247.0	3.7		311.0	165.5	1.9					
RJ	3	809.0	239.0	3.4	3.5	315.2	154.8	2.0	2.0				
JN	3	1166.0	151.9	7.7		344.5	143.0	3.8					
JN	3	1129.0	171.2	6.6	7.1	521.0	157.3	3.3	3.6				

Figure 3.7. Relative IRES activities of GT1 and GT3 constructs in BHKsinT7 cells.

Firefly/renilla luciferase ratios obtained from each sample are shown separately in each experiment. GT1 and GT3 samples in each experiment are shown by white and hatched bars respectively. Results are shown in the order in which the experiments were performed. pRLN constructs from patients BH, JN and RJ were tested separately along with H77c. Data are presented as means +/- standard deviation (SD). Arrows indicate values missing because transfection levels were invalid as described (Table 3.2).



A204C which had the least activity in BHKsinT7 cells, showed similar activity to H77c in HuH7 cells.

3.9. Is the IRES of GT 1 more efficient than that of GT 3?

The main aim of this study was to test IRES activity from a number of isolates from genotypes 1 and 3 in order to see whether there is a consistent difference between IRES activity of the two genotypes. It was shown that the IRES activity of each isolate could vary and there was overlap in translation efficiency between GT3 and GT1 isolates. Variation between IRES activities is greater between individual isolates within a genotype than between the mean activities of the 2 genotypes.

In studying the results obtained from BHKsinT7 cells, overall statistical analysis of IRES activities from the 2 genotypes was difficult in that values obtained varied by up to 20-30 % from one experiment to another. Therefore, the IRES activities of GT1 and GT3 isolates in BHKsinT7 cells were compared separately within each experiment. Firefly/renilla luciferase ratios obtained were compared within each experiment using the Mann-Whitney test which does not require the data to have similar standard deviations, nor for the data to be distributed normally as is necessary for the student t test. In all experiments, the mean IRES activity of GT 1 isolates was higher than GT3 isolates and in 3 out of 5 experiments including experiments 1, 3 and 4 was significant at $p<0.05$ (Figure 3.9.A).

In HuH7 cells, the mean translational efficiencies of GT1 isolates were not significantly different from those of GT3 within experiments using either the student t test or the Mann-Whitney test (Figure 3.9.B). All renilla activities for H77c were valid according to criteria described in section 2.6.3. Therefore, it was possible to normalise results by setting the result from the H77c clone arbitrarily at 100%. Repeat analysis of combined data showed no significant difference ($p>0.05$) by applying both student t and the Mann-Whitney tests (Figure 3.9.C).

Table 3.3. Firefly and renilla luciferase light outputs in HuH7 cells.

Firefly and renilla luciferase light outputs (RLU) obtained from 3 experiments are shown. Results are shown in the order in which the experiments were performed. Values are shown in red were excluded. ND: not done

		Experiment 1 A				Experiment 2					Experiment 3			
Patient's initial	GT	Firefly activity	Renilla activity	Firefly / Renilla ratio	Mean of ratios	Patient's initial	Firefly activity	Renilla activity	Firefly / Renilla ratio	Mean of ratios	Firefly activity	Renilla activity	Firefly / Renilla ratio	Mean of ratios
H77c	1	270.3	1615	0.167		H77c	415.5	6986	0.059					
H77c	1	371.0	1961	0.189		H77c	276.4	5086	0.054		209.2	7092	0.029	
H77c	1	288	1565	0.184	0.180	H77c	438.5	6936	0.062	0.060	219.2	6831	0.032	0.031
BA	1	331.7	1954	0.170		BA	437.1	7727	0.057					
BA	1	437.8	2235	0.196		BA	621.7	9977	0.062		156.0	5668	0.028	
BA	1	398.0	2300	0.173	0.180	BA	573.5	10056	0.057	0.059	235.0	7637	0.031	0.03
SA	1	321.7	1655	0.189		SA	382.1	6333	0.06					
SA	1	424.4	2055	0.206		SA	499.3	7549	0.066		149.7	5820	0.026	
SA	1	534.0	2508	0.213	0.203	SA	520.3	8812	0.059	0.062	277.4	8774	0.032	0.029
PE	1	403.5	2083	0.194		PE	190.4	4731	0.04					
PE	1	400.4	1963	0.204		PE	293.9	6402	0.046		189.5	5734	0.033	
PE	1	463.8	2408	0.193	0.197	PE	287.4	6081	0.047	0.044	237.7	7393	0.032	0.033
CD	1	444.6	2142	0.208		CD	301.8	5997	0.05					
CD	1	468.2	2133	0.220		CD	102.5	2416	0.042		120.3	5108	0.024	
CD	1	439.1	2024	0.217	0.215	CD	285.2	5564	0.051	0.048	234.7	7207	0.033	0.029
LA	1	374.2	1771	0.211		LA	307.9	6699	0.046					
LA	1	373.9	1811	0.207		LA	543.1	9875	0.055		143.6	6155	0.023	
LA	1	353.0	1817	0.194	0.204	LA	390.9	7482	0.052	0.051	200.6	7141	0.028	0.026
OS	1	348.1	1833	0.190		OS	144.5	2310	0.063					
OS	1	353.4	2019	0.175		OS	167.9	2559	0.066		288.6	8928	0.032	
OS	1	465.0	2141	0.217	0.194	OS	178.8	2860	0.063	0.064	333.6	9486	0.035	0.034
		Experiment 1B				BH	346.5	5850	0.059					
H77	1	235.7	1809	0.130		BH	678.9	9703	0.07		278.0	7975	0.035	
H77	1	210.6	1683	0.120		BH	697.2	10309	0.068	0.066	283.8	8260	0.034	0.035
H77	1	233.3	1778	0.131	0.127	SJ	284.1	6078	0.047					
BH	1	279.0	1917	0.146		SJ	412.4	7709	0.054		315.8	10081	0.031	
BH	1	290.5	1904	0.153		SJ	462.9	8788	0.053	0.051	269.4	8107	0.033	0.032
BH	1	298.3	1989	0.150	0.150	JA	168	2937	0.057					
SJ	3	172.7	1691	0.102		JA	289.6	4531	0.064		385.7	8607	0.045	
SJ	3	234.2	2073	0.113		JA	303.3	4766	0.064	0.062	375.7	8731	0.043	0.044
SJ	3	251.8	2280	0.110	0.108	RI	390.7	8279	0.047					
JA	3	223.7	1419	0.158		RI	138.9	2840	0.049		267.0	10222	0.026	
JA	3	267.4	1629	0.164		RI	102.2	2048	0.05	0.049	152.8	6470	0.024	0.025
JA	3	254.1	1635	0.155	0.159	MO	113.4	1733	0.065					
RI	3	206.4	2115	0.098		MO	139	1959	0.071	0.068	307.9	8099	0.038	
RI	3	169.2	1648	0.103		MO	84.9	1165	ND		239.0	6685	0.036	0.037
RI	3	219.6	2140	0.103	0.101	FV	51.52	1047	ND					
MO	3	265.6	1885	0.140		FV	92.35	1677	0.055		97.9	4510	0.022	
MO	3	244.4	1565	0.156		FV	72.67	1397	0.052	0.054	165.4	6548	0.025	0.024
MO	3	281.6	1860	0.151	0.149	BC	46.9	817.4	ND					
FV	3	182.1	1492	0.122		BC	99.89	1555	0.064		76.6	3076	0.025	
FV	3	250.5	1841	0.136		BC	96.8	1467	0.066	0.065	126.0	4416	0.029	0.027
FV	3	218.0	1791	0.122	0.127	JN	ND	ND						
BC	3	297.7	1900	0.157		JN	85.34	1531	0.056		196.0	5775	0.033	
BC	3	281.0	1832	0.153		JN	97.4	1483	0.065	0.056	221.3	6521	0.034	0.034
BC	3	275.8	1844	0.150	0.153	RJ	42.24	1164	ND					
JN	3	206.5	1252	0.165		RJ	79.04	2502	0.032		104.3	6755	0.015	
JN	3	183.6	1117	ND		RJ	99.79	2434	0.039	0.036	163.6	7582	0.022	0.019
JN	3	177.8	1226	0.145	0.155									
RJ	3	154.0	1768	0.087										
RJ	3	129.3	1647	0.079										
RJ	3	133.1	1623	0.082	0.083									

Figure 3.8. Relative IRES activities of GT1 and GT3 constructs in HuH7 cells.

Mean relative firefly/renilla luciferase ratios obtained from each sample in 3 experiments are shown. GT1 and GT3 samples are shown by bold and hatched bars respectively. Results are shown in the order in which experiments were performed. Firefly/renilla luciferase ratios obtained from all samples were normalised against H77c. Data are presented as means +/- standard deviation (SD).
exp= experiment

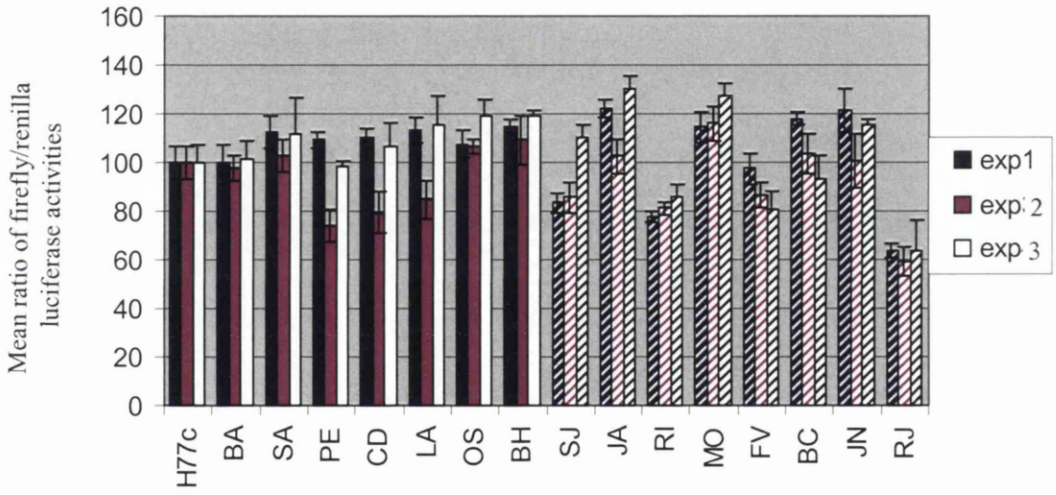
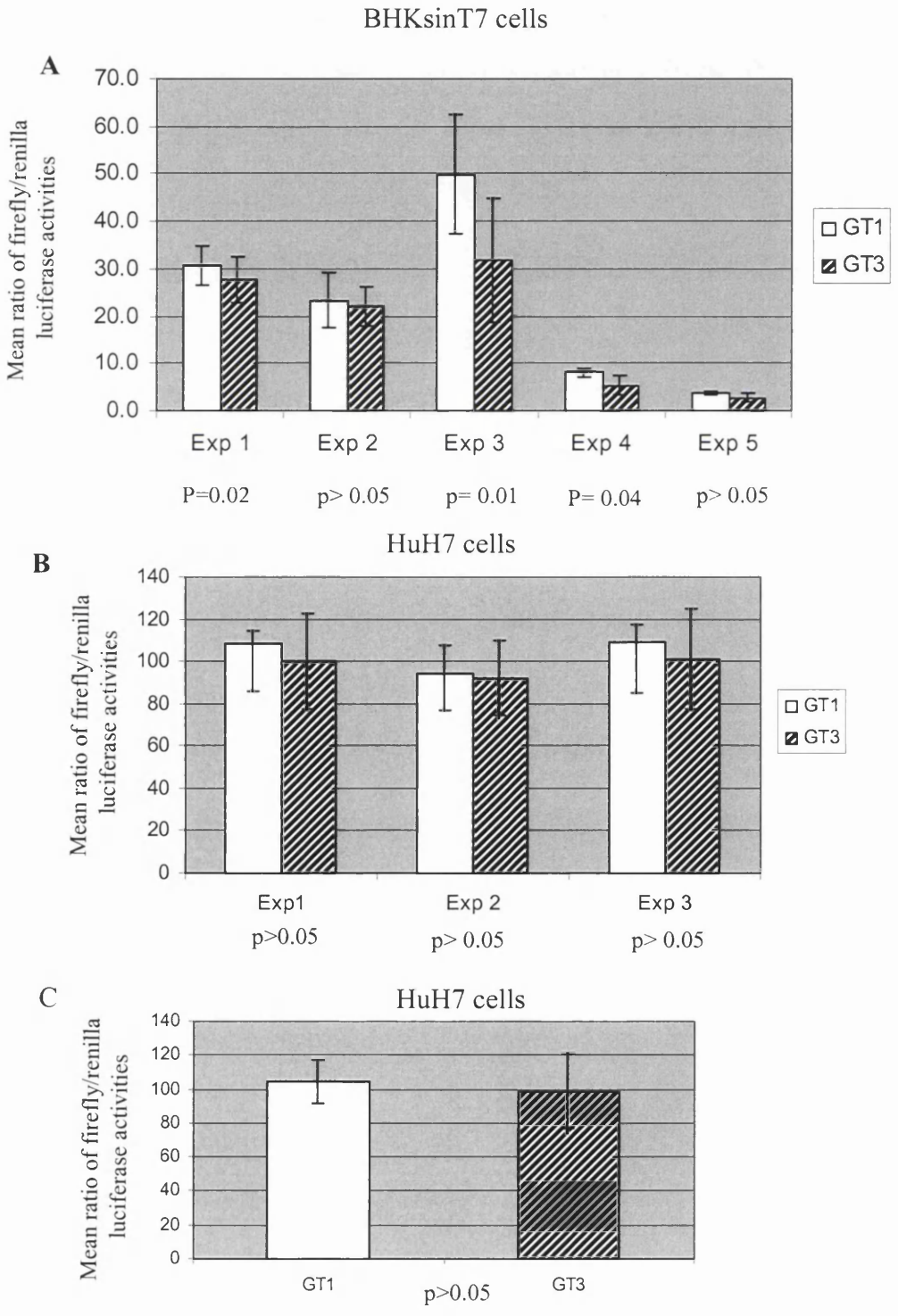


Figure 3.9. Comparison of IRES activities of GT1 and GT3 constructs.

A) Mean IRES activities of GT1 (white bars) and GT3 (hatched bars) isolates in BHKsinT7 were compared within each experiment using the Mann-Whitney test.
B) Mean normalised IRES activities of GT1 and GT3 isolates in HuH7 cells were Compared within each experiment using the Mann-Whitney test.
C) Mean relative IRES activities of combined data from three experiments were compared in HuH7 cells. Data are presented as means +/- standard deviation.
Exp= experiment



3.10. Comparison of relative translation efficiencies in BHKsinT7 and BHK cells

One report suggested that the 5'UTR cDNA has a cryptic promoter activity that can behave as a eukaryotic promoter in cells (Dumas et al., 2003). To test whether the expression of firefly luciferase in our experiments could be explained by this mechanism, it was decided to test two constructs; pRLNH77c (GT1) and that obtained from patient FV (GT3), in both BHK and BHKsinT7 cells. Unlike BHKsinT7 cells, ordinary BHK cells do not express T7 RNA polymerase, so transcription of DNA to RNA would not be expected and as a result no translation of either reporter genes would occur. In the presence of sufficient cryptic promoter activity in the 5'UTR DNA sequence to invalidate the results of our previous experiments, it would be expected that there would be firefly luciferase activity in BHK cells at a similar level to that in BHKsinT7 cells in the absence of renilla luciferase activity.

DNA constructs were transfected into both BHK and BHKsinT7 cells in duplicate wells and light outputs from firefly and renilla luciferases were obtained. In order to ensure that there was no promoter activity in the empty vector to drive the transcription of firefly luciferase, DNA from the pRL vector with no insert was also transfected. pGL3 was also transfected into both BHKsinT7 and BHK cells. In this vector, translation of firefly luciferase is controlled by strong eukaryotic SV40 promoter. As expected, no firefly activity was observed in either the BHKsinT7 or the BHK cell lines transfected with the pRL vector. Transfection of pGL3 resulted in high levels of firefly luciferase light output in both cell lines. pRLNH77c and pRLNFV transfected BHKsinT7 cells showed levels of renilla and firefly luciferase activities comparable to previous experiments. Unexpectedly, upon transfection of pRLNH77c and pRLNFV constructs into BHK cells, low levels of both renilla and firefly luciferase activities were observed (Figure 3.10.A). However, firefly luciferase activities were 80 times lower than the values seen in BHKsinT7 cells. This suggests that activity observed from constructs in BHKsinT7 cells throughout this project could not be explained simply by cryptic promoter activity.

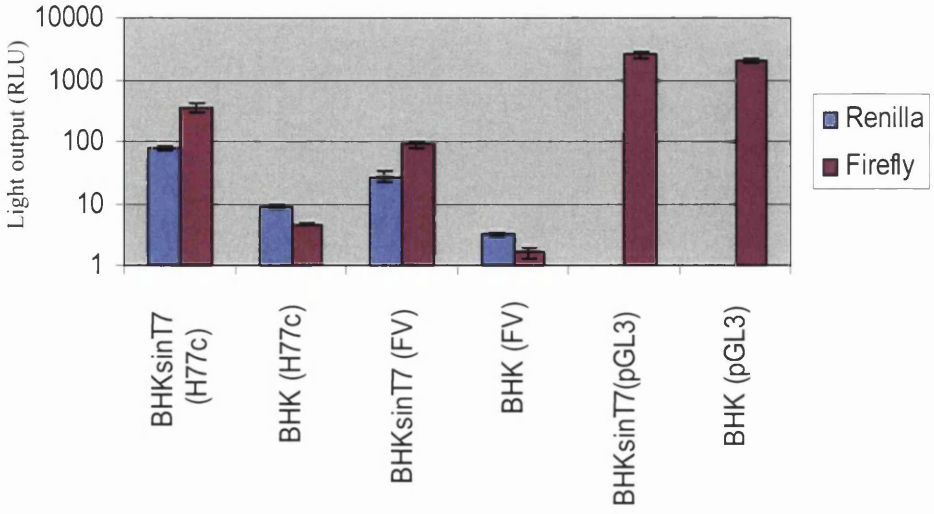
Figure 3.10. Firefly and renilla luciferase activities in BHK and BHKsinT7 cells.

A) Renilla and firefly luciferase light outputs (RLU) obtained from transfection of pRLN H77c (GT1), pRLN FV (GT3), pGL3 and the empty pRL vector into BHKsinT7 and BHK cells are shown . Lower level of firefly (1/80) and renilla luciferase (1/8) activities were detected in BHK cells in comparison to BHKsinT7 cells. pGL3 containing the firefly reporter gene driven by the SV40 promoter was used as control construct for comparison of transfection efficiency. B) Renilla and firefly luciferase light outputs obtained from pRLNH77c, pRLNFV and pGL3 in BHKsinT7 and BHK cells are shown. (logarithmic scale)

A

Sample	Cell line	Firefly activity (RLU)	Renilla activity (RLU)	Firefly/ Renilla ratio
pRLNH77c	BHKsinT7	320.20	74.83	4.27
pRLNH77c	BHKsinT7	392.40	81.24	4.83
pRLNH77c	BHK	4.76	8.43	0.56
pRLNH77c	BHK	4.65	9.34	0.49
pRLN FV	BHKsinT7	80.39	23.25	3.45
pRLN FV	BHKsinT7	96.70	31.37	3
pRLN FV	BHK	1.40	3.11	0.45
pRLN FV	BHK	1.83	3.62	0.5
pGL3	BHK	1884.00	0	
pGL3	BHK	2111.00	0	
pGL3	BHKsinT7	2372.00	0.02	86933
pGL3	BHKsinT7	2744.00	0.01	250471
pRL	BHKsinT7	0.00	58.57	
pRL	BHK	0.00	9.84	

B



3.11. Discussion

For the study of HCV IRES activities *in vivo*, bicistronic constructs were used in which 5'UTRs amplified from infected patients were inserted between the two reporter genes. Both reporter enzymes were assayed in the same cell lysate preparation and the activities were determined by the same method. In the present study, the bicistronic reporter vector, pRL, constructed by Collier *et al.* (1998) was used which contains the renilla luciferase gene act as the 5' reporter and the firefly luciferase gene as the 3' reporter. The ratio of firefly luciferase to renilla luciferase activity was used as a measure of IRES activity. In the cell cytoplasm, the T7 promoter drives transcription of the bicistronic mRNA containing the renilla luciferase, the HCV 5'UTR and the firefly luciferase.

The constructs were transfected into 2 cell lines. BHK sinT7 cells (Agapov et al., 1998) are able to express T7 RNA polymerase continuously from a noncytopathic replicon from sindbis virus (SINrep19/T7pol). One problem with using the noncytopathic replicon to support the production of T7 RNA polymerase is the limitation on cell lines which can be used. The other method used was to supply T7 polymerase by co-infecting with the recombinant vaccinia virus, vTF7-3, to allow cytoplasmic transcription of the mRNA. The advantage of this is that different cell lines can then used for the assays. However, the cytopathic effects of vaccinia infection may interfere with cellular functions. These include the induction of early cell rounding, damage to the host genome and RNA, inhibition of host protein synthesis, and death of the infected cells (Tsung et al., 1996).

In BHKsinT7 cells, the renilla luciferase light output was 20-30 times lower than that of firefly luciferase. It is possible that efficiency of translation using the sindbis replicon system was lower than that directed by the HCV 5'UTR due to inefficient capping of transcribed RNA. Unlike BHKsinT7 cells, in vaccinia infected cells higher levels of renilla luciferase activity in comparison to firefly luciferase activity was observed suggesting that transcription of T7 polymerase by vaccinia virus is more efficient than that achieved by the sindbis constructs.

We have compared the IRES activity of a number of isolates from GT 1 and 3 infected patients. It was important to ensure there were no artefactual base changes introduced by polymerase errors into the individual clones that were to be used to measure translation efficiency. Data presented in chapter 6 along with other studies suggested the presence of a quasispecies composition in the 5'UTR of the HCV genome (Soler et al., 2002). Previously, it has been shown that different 5'UTR sequences existing in an individual showed significant differences in their ability to promote translation both *in vitro* and *in vivo* (Laporte et al., 2000). Studying IRES activities from matched clones with the relevant majority sequence obtained from each patient was important because it reflects the behaviour of the major population within the quasispecies in the 5'UTR. I therefore ensured that only clones which matched exactly the sequence obtained by direct sequencing of the PCR product were examined.

In designing the constructs for this study, we included the sequences within the HCV RNA that were included in the earlier study by Collier *et al.* (1998) as representing the minimum IRES element. Our constructs contain the entire 5' UTR and 15 nts of the core coding sequence. There is controversy whether the HCV protein coding sequence downstream of the IRES is essential for IRES function. Reynolds *et al.* (1995) suggested that 14 to 32 nts of the HCV capsid-coding sequence are absolutely required for efficient IRES activity in *in vitro* and *in vivo* systems. In contrast, other studies have shown efficient IRES activity in the absence of core coding sequence (Rijnbrand et al., 1995; Wang et al., 1993). One explanation for the discrepancy between these studies was suggested by Honda *et al.* (1996a). The fusion of reporter gene sequences with the 5'UTR may form an RNA structure that is unfavorable for translation initiation. The inclusion of the 5' core sequence which constitutes part of domain IV may prevent base pair formation between the HCV sequence and the beginning of the firefly luciferase coding sequence (reviewed by Rijnbrand and Lemon, 2000). In the present study, the core coding sequences needed to form the proximal stem-loop of domain IV were included as in the study carried out by Collier (1998).

The data presented in this study showed that deletion of 17 nts from the 5' end of the 5' UTR resulted in a two fold reduction in IRES activity. This is consistent with the findings reported by others that stem loop I contains RNA elements required for optimal HCV RNA translation (Friebe et al., 2001; Fukushi et al., 1994; Luo et al., 2003). It should be noted that deletion of stem loop I in our original construct (nt 1-17) was partial whereas these studies examined the construct lacking the entire domain I including nucleotides 1-20. Our result appeared to contradict the observations made by other groups which suggest an inhibitory role for stem loop I on translation by the IRES (Honda et al., 1996b). It has been suggested that the inhibitory effect of stem loop I (nts 1-22) may be cell type specific (Kamoshita et al., 1997). They reported that IRES activity of constructs lacking stem loop I was equally efficient in HeLa cells, but less efficient in African monkey kidney cells.

Of 8 GT 3 isolates tested for IRES activity in BHKsinT7 cells, patient RJ with a substitution C121U in the pyrimidine tract (nts 120-130) with an activity of 40% had the least activity. It is interesting that mutational analysis of the single stranded poly C region between domains II and III, consisting of nucleotides 120-125, suggested that this region played an essential role in maintaining the proper spacing of specific elements within the IRES and substitutions in this region could lead to the disruption of translation initiation (Varaklioti et al., 1998). No other isolate in our cohort had changes in this region.

When the different IRES activities were compared in each separate experiment using BHKsinT7 cells, the mean IRES activity of GT 1 isolates was significantly higher than GT 3 isolates in 3 out of 5 experiments. It should be noted that the mathematics of the Mann-Whitney test make it "harder" to reach significance than with the student t test. However, in HuH7 cells which had been pre-infected with a recombinant vaccinia virus expressing the T7 RNA polymerase, the mean IRES activity of GT 1 isolates was not significantly greater than GT3 isolates. Studying IRES activity in cells of hepatic origin may be more appropriate than other cell lines, since those cells are the main site of HCV replication. However, it should be noted that HuH7 cells are derived from tumour cells, not normal hepatocytes. Buratti *et al.* (1997) reported that the GT 3a IRES was only 50% as active as 1b in COS-1 (African Green Monkey kidney cells). However, the GT 3a constructs of

Buratti *et al.* contained the first four residues of GT 1 core sequence rather than GT 3. Collier *et al.* (1998) reported similar activity for GT 1a and 3a in BHK and HuH7 cells. However, their constructs contained only nts 18 to 356 of the HCV genome thereby including 15 nt of the core coding region. The construct used in this study contained the full length 5'UTR and 15 nucleotides of the core coding region (nts 1-356). All these reports examined only one isolate as representative of each genotype. Therefore, it is not clear whether the IRES activity observed was actually isolate specific. Also, no attempt was made to ensure that the sequences tested for IRES activity were matched with the majority sequence found in the patient. It is conceivable that the observed IRES activity was that of a minor variant in the quasispecies or contained errors generated by *Taq* polymerase or other enzymes.

A basic difficulty in performing statistical analysis of data from this study was that experiments were not performed in a randomised order. In each experiment, the order of testing of isolates was the same with GT1 isolates preceding GT3 isolates. This may have led to unrecognised bias. In the second experiment using BHKsinT7 cells, the renilla luciferase activity obtained from the reference sequence H77c, which was used for normalisation of data, did not meet the criteria set down for achieving an adequate renilla activity. Therefore, comparison of IRES activities of isolates was not possible between experiments. It was decided to compare the mean IRES activity of GT1 and GT3 isolates within each experiment in BHKsinT7 cells. It was shown that the difference between mean IRES activities of GT1 and GT3 isolates was statistically significant in 3 out of 5 experiments. However, the biological significance is more uncertain. The renilla luciferase activities in the third experiment were approximately 10 fold lower than the other experiments in BHKsinT7 cells, possibly due to a low level of transfection or poor quality of the transfected cells, although the cells appeared as healthy as in the other experiments. It should be noted that all values remained clearly within the linear range of the luciferase assays except for the excluded values for OS, which, if included would have reduced the likelihood of the difference reaching statistical significance. Our feeling overall is that the low level of statistical significance of the differences in mean firefly to renilla luciferase ratios does not represent a real biological difference between GT1 and

GT3 IRES elements in BHKsinT7 cells but this could be resolved only by conducting further similar experiments.

The relative firefly/renilla ratios obtained from a single transcript differed between experiments when using the same cDNA on different days. This variation was more noticeable in BHKsinT7 cells than HuH7 cells. This causes some concern and, in the future, it would be preferable to standardise cell culture conditions more closely than was the case in this work. The exact reason for the observed variation in ratios obtained from a single transcript is not clear. During this project, different batches of BHKsinT7 cells were used and no synchronization of cells was carried out prior to transfection. It is conceivable that T7 RNA polymerase activity may differ under certain circumstances such as the confluency, quality of cells or their passage number and history, which could result in generation of different levels of transcripts. If cellular *trans*-acting factors are available in limiting amounts, the presence of more IRES-containing transcripts would exhaust the *trans*-acting factors needed for cap-dependent translation and as result the relative level of cap-independent to cap-dependent translation would increase. Another possible explanation is the position in the cell cycle of the majority of cells in the culture, which has been shown to play a role in the activity of certain IRES elements. Compared with cap-dependent translation, the activity of the IRES was greatest in actively growing cells and relatively reduced in resting synchronised HuH7 cells using bisictronic luciferase constructs similar to those used in this project (Honda et al., 2000). This suggests that HCV translation is regulated by cellular proteins that vary in abundance during the cell cycle. Differences in transfection efficiency may also influence the firefly/renilla ratios by an unknown mechanism although it appears to affect the translation of both reporters in the same DNA. In order to overcome the differences between ratios of each isolate in separate experiments, DNA from H77c was used in order to normalise the ratios. Finally, it is possible that the delay in adding one or other substrates in the luciferase assay could influence the results. This was not applicable during my experiments as the addition of substrates were carried out according to programme set up in the luminometer according to manufacturer's instructions.

Our study was not designed to address whether the response of individual patients to interferon therapy was related directly to the structure of the individual's viral 5'UTR. However, overall, GT3 infected patients have been shown repeatedly to respond to interferon-based therapies markedly better than that those with HCV GT1. If, on average, the 5'UTRs of GT3 isolates were less efficient at directing translation of the polyprotein, this could contribute to the greater effectiveness of interferon therapy. Our study results do not support this hypothesis. The lack of relation between 5'UTR sequence and response to interferon has been reported by other groups (Saiz et al., 1999; Soler et al., 2002).

Dumas *et al.* (2003) reported the presence of a strong cryptic promoter within the 5'UTR DNA sequence. They concluded that results obtained from using cells transfected with bicistronic DNA constructs containing the HCV 5'UTR should be analysed at both translational and transcriptional levels. I carried out a preliminary experiment to compare IRES activity in BHKsinT7 cells expressing T7 RNA polymerase and BHK-21 cells lacking T7 RNA polymerase. In the absence of T7 RNA polymerase, no RNA transcripts from the construct should be produced. Therefore, if there was a promoter activity in the 5'UTR cDNA in our system, we would expect to have similar firefly luciferase activities in both cell lines. The data from this experiment was not able to exclude the presence of a possible cryptic promoter in the cDNA of the 5'UTR, as the constructs tested were circular and the presence of any promoter in the 5'UTR sequence could result in transcription of the entire plasmid including the renilla and firefly luciferase genes. However, it was concluded that the presence of such weak firefly luciferase activity was insufficient to invalidate the results of translation in BHKsinT7 cells by the authentic RNA IRES in this study. However, if time had permitted, I would also have performed Northern blot analysis and/or RT-PCR analysis to determine how much RNA transcript was produced in each experiment.

Chapter 4

Comparison of IRES activity of 5'UTR plus core regions from GT 1 and 3 infected patients

4.1. Introduction

The HCV core protein is believed to form the viral nucleocapsid. The core protein has been reported to have a wide range of biological activities such as interaction with host cell proteins including members of the tumor necrosis receptor family, apolipoprotein AII and heterogeneous nuclear ribonuclear protein K (McLauchlan, 2000). It has been suggested that core protein modulates sensitivity to apoptosis (Ray et al., 1996) and induces liver steatosis (Fujie et al., 1999). It has been shown that the HCV core protein contains a unique domain which plays a vital role in the localization of core to lipid droplets (Hope and McLauchlan, 2000).

Interaction of the core protein with the genomic RNA of HCV and its effects on cap-independent translation mediated by the IRES have been studied by several groups (Kim et al., 2003; Reynolds et al., 1995; Rijnbrand et al., 2001; Shimoike et al., 1999; Wang et al., 2000; Zhang et al., 2002). The 3' boundary of the IRES is uncertain. Early studies demonstrated that the inclusion of nt 1 to 32 of the core protein coding sequences was essential for efficient IRES activity (Lu and Wimmer, 1996; Reynolds et al., 1995).

The data reported by Santolini *et al.* (1994) showed that core protein is an RNA binding protein and RNA binding domains have been localized to the N-terminal 75 amino acids. Shimioke *et al.* (1999) constructed a series of monocistronic plasmids containing HCV

5'UTR and 5'UTR plus core fused to the firefly reporter gene and found that expression of core suppressed the firefly activity in a dose dependent manner in an *in vivo* system. It was also reported that amino acids 34-44 of core interact with the IRES RNA and contribute to the inhibition of translation (Zhang et al., 2002). This suppression was eliminated by frameshift mutations introduced into this region which abolished expression of the core protein, suggesting that it is the core protein rather than the core-coding nucleotide sequence that downregulated the efficiency of the HCV IRES activity. More recently, Li *et al.* (2003) reported that a synthetic peptide representing amino acids 1–20 of the core protein inhibited IRES activity in HepG2 cells.

In contrast, Wang *et al.* (2000) reported that suppression of IRES-directed translation resulted from an RNA–RNA interaction between the core coding nt sequence and the IRES RNA. They constructed a dicistronic reporter system containing the 5'UTR plus 66 nt from core coding sequence (nt 1-407) and the 5'UTR plus 518 nt of core (nt 1-860) coding sequence fused in-frame directly with firefly luciferase along with two other constructs containing similar lengths of the core, but with frameshift mutations in the core sequence that resulted in translation of a nonsense sequence. A lower efficiency of translation was obtained from constructs containing nt 1-860 (17%) than in transcripts containing only 66 nt (nt 1-407) of core in HepG2 and HuH7 cells (60%). However, the translational analysis of the frameshift mutants showed similar results suggesting that this effect was due to the inclusion of the RNA sequence and not to the expression of the core protein. Ito *et al.* (1999) identified a region at the 3' end of the core sequence which interacted with the polyprimidine tract binding protein (PTB) resulting in the inhibition of translation activity by the IRES. This inhibition was relieved by the presence of the X region at the end of the 3' untranslated region. The presence of an RNA- RNA interaction between the IRES and the core coding sequence was further supported by study published by Kim *et al.* (2003). They demonstrated that nucleotides 428-442 of the core coding sequence annealed to nucleotides 24-38 of the 5' UTR resulting in downregulation of cap-independent translation by the HCV IRES.

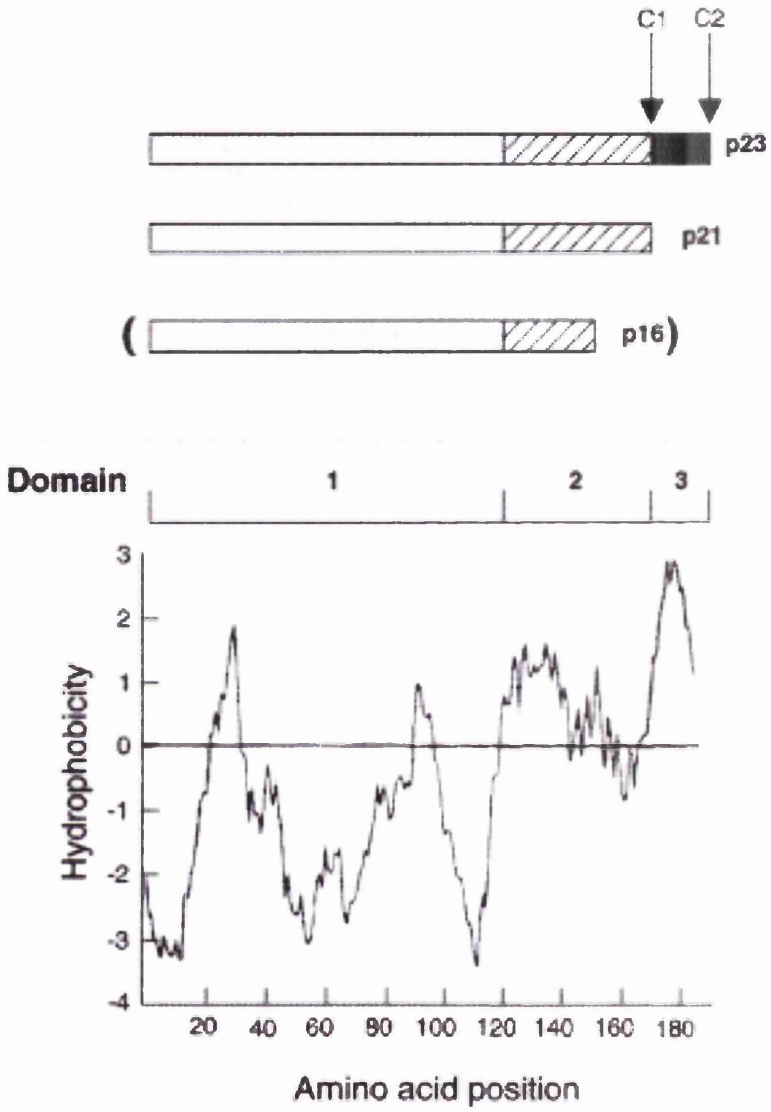


Figure 4.1. Hydrophobicity profile of the core polypeptide and the three proposed domains of core.

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).

Core protein (p23) is processed by a host signal peptidase at aa 191 to produce an immature form of the protein which is further processed by a host signal peptide peptidase (SPP) between aa 174-191 to produce what is thought to be the mature form of the protein (p21)(Figure 4.1). This latter processing event is important for conferring localization of core protein to lipid droplets (Hope and McLauchlan, 2000). Core is highly conserved between the 6 genotypes and the hydrophobicity profile of core identifies 3 domains within the protein. Domain I (aa 1-122) contains a high proportion of basic residues. Domain 2 (aa 123-174) is more hydrophobic than domain 1, while domain 3 (aa 175-191) is highly hydrophobic and acts as signal sequence for directing the viral E1 glycoprotein to the ER for further maturation prior to cleavage at aa 191 by a host signal peptidase (Santolini *et al.*, 1994). One possible explanation for downregulation of translation observed with inclusion of full-length core is that when the ribosome reaches a signal sequence the RNA and translational machinery are translocated to the ER for further processing which slows down translation.

The data presented in this chapter are preliminary. The original plan was to construct 5'UTR plus core clones matched with the majority sequence obtained from GT 1 and 3 infected patients similar to those described for the 5'UTR alone. This was in order to compare translation efficiency by the IRES in HCV isolated from GT 1 and 3 infected patients using constructs containing 5'UTR plus core to determine whether there is a consistent difference between the 2 genotypes.

4.2. Amplification of 5'UTR plus core regions

I designed oligonucleotides to amplify the entire 5'UTR of HCV and core gene (nt 1-914). The sequences of these primers and the nucleotide positions to which they bind can be seen in Table 4.1.A. *BamHI* restriction sites were incorporated into the 5' ends of the primers for second round amplification to facilitate future cloning. Samples from seven GT 1 infected patients with unique majority sequences were chosen for cloning. PCR products made by P. Preikschat as described in section 2.2.8.2 were used as template to amplify the 5'UTR plus core using sense primer MMCG1 and antisense primer MMCG2.

A

Primer	*Sequence 5' to 3'*	*nt Position*	*Genotype*
MMCG1 (s)	TGAGGATCCGCCAGCCCCTGATGGGGGCGACACTCCACCAT	1-33	1
PP5UTR (s)	TGAGGATCCACCTGCCTCTTACGAGGCGACACTCCACCAT	1-33	3
MMCG2 (as)	TGAGGATCCGGCTGAAGCGGGCACAGT	914-897	1
PPCORE-R (as)	TGAGGATCCACTAGCTGCTGGATGAAT	914-897	3

B

Primer	*Sequence 5' to 3'*	*nt Position*	*Genotype*
MS1 (s)	GCTTGCGAGTGCCCCGGGAG	301-320	1,3
MS3 (s)	CGCGGATCCAGATCTGGGAGGTCTCGTAGACCGTG	316-335	3
MM10 (as)	CCTC (AG) TACACAATACT	972-988	1, 3
PP1 (as)	AGCAAGGATCCTCATACACAATACT	972-988	3
PP3 (as)	AGCAAGGATCCGCCTCGTACACAATACT	972-988	1

Table 4.1. Oligonucleotide primers for amplification and sequencing of core region.

Primers used for PCR of GT 1a and 3a 5'UTR pluscore regions (A). Primers for sequencing of genotype 1 and 3 core regions (B).
Nucleotide numbering according to H77c (AF 011751) for GT 1 and NZL1(D17763) for GT 3 samples."s"= sense, "as"= antisense. *BamHI* restriction sites are underlined.

From GT 3 infected samples, 7 unique sequences were identified. Amplification of GT 3 samples was carried out using primers PP5'UTR and PPCORE-R. High fidelity DNA polymerase (HF2, Clontech) was used for amplification to reduce errors resulting from misincorporation of bases by *Taq* DNA polymerase. PCR products from two identical reactions were pooled and gel purified as described in section 2.2.9.2

4.3. Sequencing of core coding sequences from patients

In order to find clones matched with majority sequences from patients, the first step was to obtain the majority sequence from those patients. Since only incomplete or no core PCR product majority sequences were available, sequencing of the core coding region from studied patients was carried out using the sense and antisense primers shown in Table 4.1.B. All sequences were edited using "Sequence Navigator" (Applied Biosystems) and then exported into GCG (Wisconsin Package, version 10.2-Unix). Several sequences from the 5'UTR and core regions were assembled, using "GelMerge" and "GelAssemble" programs to obtain the majority sequence for each patient. Alignments of the full core majority sequences from 7 GT1 and 7 GT 3 isolates are shown in Figures 4.2. and 4.3. respectively. Figure 4.4 shows the amino acid sequences of core protein observed in the patient samples analysed. In GT1 samples, only 5 out of 73 nucleotide differences were nonsynonymous. Of 72 nucleotide differences observed in GT3 samples, 4 were nonsynonymous.

4.4. Construction of plasmids containing 5'UTR plus core (pRLNC)

Amplified PCR products containing the 5'UTR plus entire core region (nt 1-914) obtained from serum samples were cloned into the pRL vector (Collier et al., 1998) using the unique *BamHI* restriction site as shown in Figure 4.5.A. Individual clones were cultured overnight and plasmid DNAs were extracted using an in-house miniprep method as described in chapter 2.2.1.2. Plasmids were digested with *NcoI* and *HindIII* restriction

Figure 4.2. Alignment of the majority nucleotide sequence of the core region (nt 341-914) obtained from genotype 1a infected patients. Sequences derived from serum from each patient. H77c (accession no: AF 011751) was used as reference sequence for comparison. "." represents those bases which are identical to the consensus sequence.

```
          341                                                                                                  441
   BA     .......... .......... .......... .......... .......... .......... .....t.... .......... .......... ..........
   OS     .......... .......... .......... .......... .......... .......... .....t.... .......... .......... ..........
   PE     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   SA     .......... .......... .......... .......... .......... c......... .....t.... .......... .......... ..........
   CD     ........a. .......... .......... .......... .......... .......... .......... .......... .......... ..........
   LA     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   BH     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   h77    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Consensus ATGAGCACGA ATCCTAAACC TCAAAGAAAA ACCAAACGTA ACACCAACCG TCGCCCACAG GACGTCAAGT TCCCGGGTGG CGGTCAGATC GTTGGTGGAG

          442                                                                                                  541
   BA     .......... .......... .......... .......... .......... .......... .......... g........c .......... ..........
   OS     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   PE     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   SA     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   CD     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   LA     .......... .......... .......... .......... .......... .......... .......a.. ...c...... .......... ..........
   BH     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   h77    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Consensus TTTACTTGTT GCCGCGCAGG GGCCCTAGAT TGGGTGTGCG CGCGACGAGG AAGACTTCCG AGCGGTCGCA ACCTCGAGGT AGACGTCAGC CTATCCCCAA


          542                                                                                                  641
   BA     ...g...... .......... .......... .......... ..c....... .......... .......... .......... .......... ......c...
   OS     ...g...... .......... .......... .......... ..c....... .......... .......... .......... .......... ......c...
   PE     ...a...... .......... .......... .......... ..t....... .......... .......... .......... .......... ..........
   SA     ...a.....c t......... .......... .......... ..t....... .......... .......... .......... .......... ..........
   CD     ...g....a. .......... .......... ......t... ..t....... ....t..c.. .........t .......... .......... ..........
   LA     ...a...... .t........ .......... .......... ..t....... .......... .......... .......... .......... ..........
   BH     ...g...... .......... ..t....... .......... ..c....... .......... .......... .......... .......... ......c...
   h77    ...a...... .......... .......... .......... ..c....... .......... .........t .......... .......... ..........
Consensus GGC-CGTCGG CCCGAGGGCA GGACCTGGGC TCAGCCCGGG TA-CCTTGGC CCCTCTATGG CAATGAGGGC TGCGGGTGGG CGGGATGGCT CCTGTCTCCC
```

Figure 4.2. part 1

```
          642                                                                                                741
   BA     .......... .......... ......t... .......... .......... .......... .......... ....t..... .......... ..........
   OS     .......... .......... ......t... .......... .......... .......... .......... .......... .......... ..........
   PE     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   SA     ..c....... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   CD     .....t.... ....c.a... .......... .......... .......... .......... .......... .......t.. .......... ..........
   LA     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   BH     .......... .......... ......a... .......... .......... .......... .......... .......... .......... ..........
   h77    .......... .......... .......... .......... .......... .......... .......... .......t.. .......... ..........
Consensus CGTGGCTCTC GGCCTAGCTG GGGCCCCACA GACCCCCGGC GTAGGTCGCG CAATTTGGGT AAGGTCATCG ATACCCTCAC GTGCGGCTTC GCCGACCTCA

          742                                                                                                841
   BA     .......... t......... .......... ....g..... .......... .......... .......... .......... .......... ..........
   OS     .......... t......... .......... ....g..... .......... .......... .......... .......... .......... ..........
   PE     .......... .......... ........c. .......... .......... .......... .......... .......... ........t. ..........
   SA     .......... .......... ........c. .......... .......... .......... .......... .......... .......... ..........
   CD     .......... .......... .......... .........t .......... .......... .......... .......... .......... .c........
   LA     .......... .......... ........c. .......... a......... .......... .......... .......... .......... ..........
   BH     .......... t......... .......... .......... .......... .......... .......... .......... .......... ..........
   h77    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Consensus TGGGGTACAT ACCGCTCGTC GGCGCCCCTC TTGGAGGCGC TGCCAGGGCC CTGGCGCATG GCGTCCGGGT TCTGGAAGAC GGCGTGAACT ATGCAACAGG

          842                                                                                    914
   BA     ...t...... .......... .......... ......g... ..a....... .......... ......g... ...
   OS     .......... .......... .......... ......a... ..c....... .......... ......g..g ...
   PE     .......... .......... .......... ...ct.a... .......... .tt....c.. .......... ...
   SA     .......... .......... .......... ...c..a... .......... .......... .......... ...
   CD     .......... .......... .......... ......g... .......... .......c.. .......... ...
   LA     ...t.....c .......... .......... ...c..a... .....t.... .......... .......... ...
   BH     .......... .......... .......... ......g... .......... .......... ......g... ...
   h77    .......... .......... .......... ......g... .......... .......... .......... ...
Consensus GAACCTTCCT GGTTGCTCTT TCTCTATCTT CCTTCT-GCC CTGCTCTCTT GCCTGACTGT GCCCGCTTCA GCC
```

Figure 4.2. part 2

Figure 4.3. Alignment of the majority nucleotide sequences from the core region (nt 339-912) obtained from GT 3a infected patients. Sequences derived from serum from each patient. NZL1 (accession no: D17763) was used as reference sequence for comparison. "." represents those bases which are identical to the consensus sequence.

```
          339                                                                                                      439
FV        .......... .......... .......... .......... .......... .......... .....t.... .......... .......... ..........
MO        .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
ME        .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
BC        .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SJ        .......... .......... .......... c......... .......... .......... .......... .......... .......... ..........
JA        .......... ....a..... .......... .......... ....a..... .......... .......... .......... .......... ..........
RI        .......... .......... a......... .......... .......... .......... .......... .......... .......... ..........
NZL1      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Consensus ATGAGCACAC TTCCTAAACC TCAAAGAAAA ACCAAAAGAA ACACCATCCG TCGCCCACAG GACGTCAAGT TCCCGGGTGG CGGACAGATC GTTGGTGGAG

          440                                                                                                      539
FV        .......... .......... .......... .......... .......... .......... .......... .......... .....g.... ..........
MO        .......... .......... .......... .......... .......... ..g....... .......... .......... ..a....... ..........
ME        .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
BC        .......... .......... .......... .......... ...a...... .......... .......... .......... .......... ..........
SJ        .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
JA        .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
RI        .......... a......... .......... .......... .......... .......... .......... .......... .....g.... ..........
NZL1      .......... .......... .......... .......... .......... .......... .......... .......... ..a....... ..........
Consensus TATACGTGTT GCCGCGCAGG GGCCCACGAT TGGGTGTGCG CGCGACGCGT AAAACTTCTG AACGGTCACA GCCTCGCGGA CGGCGACAGC CTATCCCCAA

          540                                                                                                      639
FV        .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
MO        ...a...... .......... .......... .......... .......... .......... .......... .......... .......... ......t..g
ME        .......... .......... .......... .......... .......... .......... .......... ..t....... ....a..... ..........
BC        .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SJ        .......... .......... .......... .......... .......... .......... .......... .......... ....a..... ..........
JA        .......... .......... .......... .......... .......... ....t..... .......... .......... .......... .........t
RI        .......... .......... .......... .......... .......... .......... .......... .......... .......... .........g
NZL1      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Consensus GGCGCGTCGG AGCGAAGGCC GGTCCTGGGC TCAGCCCGGG TACCCTTGGC CCCTCTATGG TAACGAGGGC TGCGGGTGGG CAGGGTGGCT CCTGTCCCCA
```

Figure 4.3. part 1

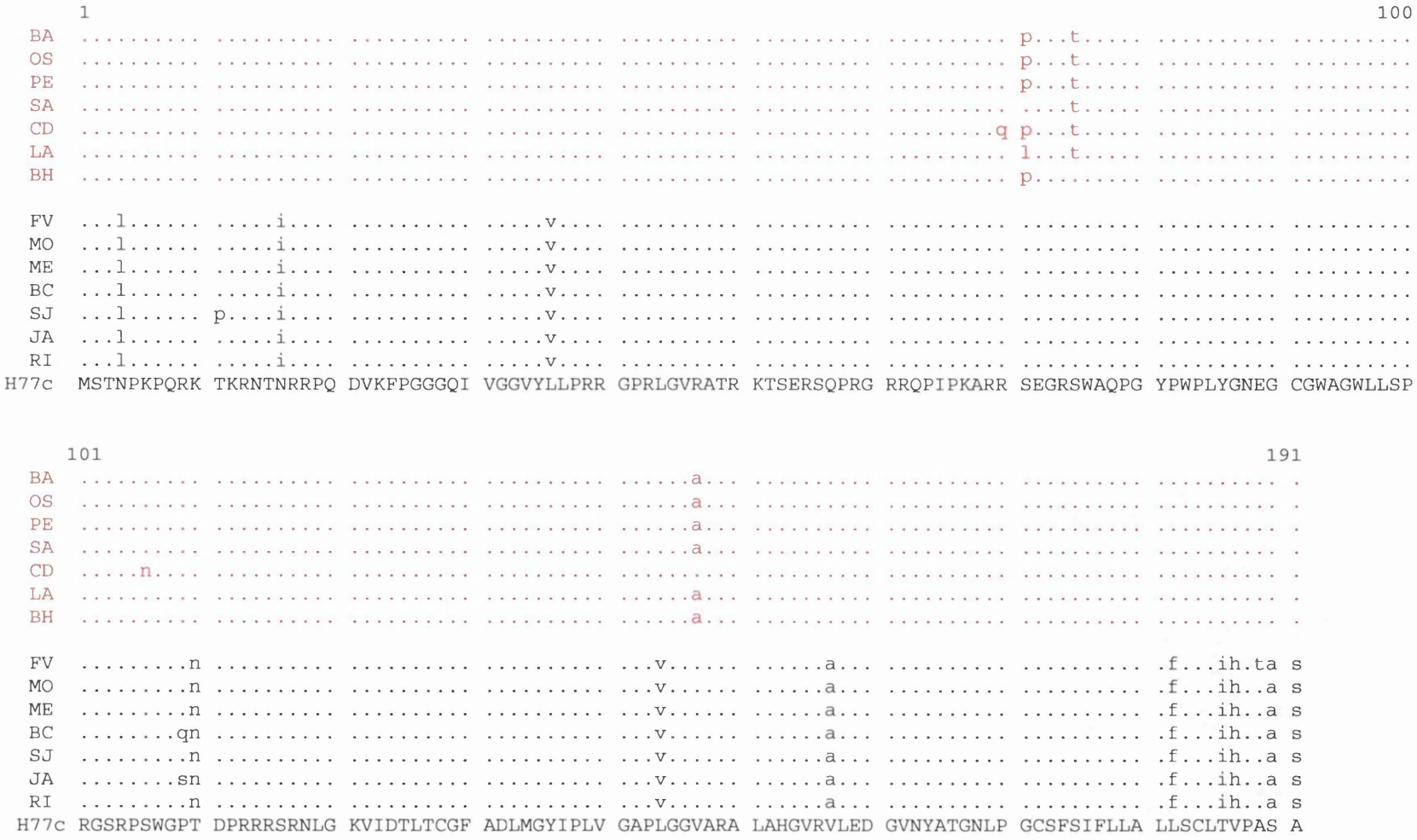
```
          640                                                                                                  739
   FV     .......... .......... .........c .......... .a........ ...cc..... .......... .......... .......... ..........
   MO     .......... .......... .........c .......... .g........ .......... .......... .......t.. .......... ..........
   ME     ..a....... .......... .........t .......... .g........ ...c...... .......... .......... .......... ..........
   BC     .......... .......... .....a...t .......... .a.....t.. .......... ..g....... .......... .......... ..........
   SJ     .....t.... .......... .........t .......... .a..a..... ...c...... ..g....... .c.....t.. .......... ..........
   JA     .......... .......... ....t....c .......... .a........ .......... .......... ....a..... .......... ..........
   RI     ........t. .......... .........c .......... .g........ .......... .......... .......... .......... .....t....
   NZL1   .......... .......c.. .........t .......... .g........ .......... .......... .......a.. .......... ..........
Consensus CGCGGCTCCC GTCCATCTTG GGGCCCAAA- GACCCCCGGC G-AGGTCCCG CAATTTGGGT AAAGTCATCG ATACCCTCAC GTGCGGATTC GCCGACCTCA

          740                                                                                                  839
   FV     .......... ......t... .......... .......... .......... .......... .......... .......... ...g...... ..........
   MO     .......t.. .......... .....c.... ....g..... .......... .......... .......... .......... .......... ..........
   ME     .......... .......... .......... .......... .......... ........c. .......... .......... .......... ..........
   BC     .......... .......... .....c.... ....g..... .......... .......... .......... .......... .......... ..........
   SJ     .......... .......... .......... .......... .......... ........c. .......... .......... .......... ..........
   JA     .......... .......... .......... .......... .......... ..t....... ....a..... .......... .......... .c........
   RI     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
   NZL1   .......... .......... ........t. .......... .......... .......... .......... .......... .......... .c........
Consensus TGGGGTACAT CCCGCTCGTC GGCGCTCCCG TAGGAGGCGT CGCAAGAGCC CTCGCGCATG GCGTGAGGGC CCTTGAAGAC GGGATAAATT TTGCAACAGG

          840                                                                                  912
   FV     .......... .......... .......... ......c... .......... .......... ....a..... ...
   MO     ...t...... .......... .......... .......... .......... .......... .......... ...
   ME     .......... .......... .......... .......... .......... .......... ......t... ...
   BC     .......... .......... .......... .......... .......... ....g..... c........c ...
   SJ     .......... .......... .......... .......... .......... .......... ......t... ...
   JA     .......... .......... .......... t......... .......... .......... .......... ...
   RI     .......... .......... .......... .......... .......... .......c.. .......... ...
   NZL1   .......... .......... .......... .......... .......... .......... .........c ...
Consensus GAACTTGCCC GGTTGCTCCT TTTCTATCTT CCTTCTTGCT CTGTTCTCTT GCTTAATTCA TCCAGCAGCT AGT
```

Figure 4.3. part 2

Figure 4.4. Alignment of the amino acid sequence of the core region (aa 1-191) obtained from 7 GT 1a (indicated by red) and 7 GT 3a (black) infected patients.


1
100
BA
OS
PE
SA
CD
LA
BH
FV
MO
ME
BC
SJ
JA
RI
H77c MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR KTSERSQPRG RRQPIPKARR SEGRSWAQPG YPWPLYGNEG CGWAGWLLSP
101
191
BA
OS
PE
SA
CD
LA
BH
FV
MO
ME
BC
SJ
JA
RI
H77c RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGVARA LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTVPAS A

enzymes to identify the orientation of the insert. In total, 337 clones were examined to find 5 correctly orientated clones for each sample. Extraction of DNA for sequencing was carried out using Miniprep kit (Qiagen) as described in chapter 2.2.1.1. At least 2 clones for each sample were sequenced using 2 sense and 2 antisense primers to cover the entire 5'UTR plus core region (Figure 4.5.B). Sequences obtained from different primers were assembled and compared to the majority sequence from each patient. In total, 28 clones were sequenced but no complete clone matched with the majority sequence was found. Clones differed by from 1 to 8 nucleotides from the majority sequence. Sequencing of more clones was not attempted because of the high cost of sequencing and time constraints.

4.5. Relative translation efficiency of 5'UTR plus Core constructs (pRLNC) to equivalent 5'UTR constructs (pRLN)

As mentioned previously, construction of clones containing the 5'UTR and full length core (pRLNC) which matched exactly the majority sequence obtained from patients was not successful therefore, as a compromise, it was decided to measure the IRES activity of 3 clones containing 5'UTR and core from each patient. It was reasoned that if there was a consistent and significant difference in translation between genotypes, it would be evident with such an approach. I attempted to use constructs from three patients in each group (GT1 and GT3 patients) including those showing lowest and highest IRES activity as found in the previous study. Therefore, pRLNC plasmids obtained from patients OS, LA and PE from GT1 and RJ, FA and JA from GT3 were used for transfection. In total, DNA from nineteen pRLNC constructs including pRLNC of H77c was transfected into BHKsinT7 cells and the relative IRES activity measured. This experiment was carried out with duplicate wells and on two different occasions.

Table 4.2 shows the renilla and firefly luciferase light outputs obtained from two experiments. Comparison of IRES activity between constructs containing the 5'UTR and 15 nts of core (pRLN) and 5'UTR plus full length core (pRLNC) was carried out as shown

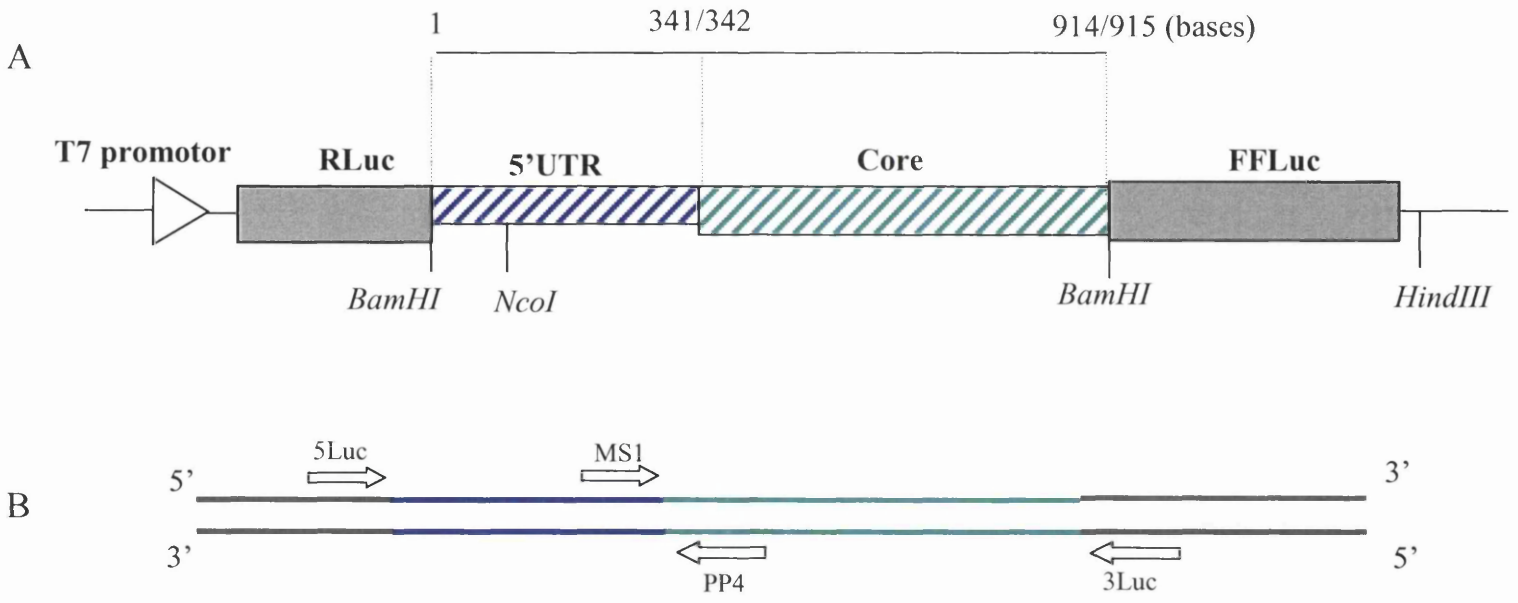


Figure 4.5. Structure of the bicistronic, dual luciferase reporter plasmid containing 5'UTR and entire core coding sequence (pRLNC).

Open triangle represents T7 promotor. The two luciferase reporter genes are indicated by gray boxes; Rluc, renilla luciferase and FFluc, firefly luciferase. The HCV cDNA corresponding to 5'UTR and core sequences are shown by blue and green hatched boxes respectively. Restriction enzyme cleavage sites used for DNA manipulation are indicated under the cDNA. B) shows the sequencing of pRLNC clones in both direction using 5Luc and MS1 as sense and PP4 and 3Luc as antisense primers. Double lines with the same color as Fig.A represent the same region.

in Figure 4.6. All renilla luciferase light output values met the criteria for acceptability (section 2.6.3). The IRES activity of the pRLNC clones is presented as the ratio of firefly to renilla activities, relative to that obtained from the corresponding pRLN construct which was taken as 100%. Constructs containing 5'UTR and full length core showed a reduction in relative translation efficiency to only 3-7% of the level seen with the corresponding pRLN constructs. No firefly activity was observed in cell extracts transfected with the construct JA-2. Sequencing of this clone in both directions revealed the presence of a deletion at nt 490 within the core coding sequence causing a frameshift which generated 4 stop codons in the downstream core sequence. Therefore, no functional firefly protein would be predicted to be produced.

4.6. IRES activity in GT1 and GT3 pRLNC constructs

Figure 4.7 shows translation levels of pRLNC constructs in two experiments. There was a variation in translation efficiencies of the 3 different clones obtained from each patient. In order to examine genotype difference in IRES activity, the mean relative translation efficiency of all clones obtained from GT1 patients was compared with those obtained from GT3 in each separate experiment. No significant difference was observed between the IRES activities of GT1 and GT3 constructs in two experiments analysed separately using both student t and Mann-Whitney tests (Figure 4.7.C). Clone JA-2 was excluded from analysis.

The ratio of firefly to renilla obtained from pRLNC H77c was arbitrarily taken as 100% in each experiment and other ratios were normalised against that. No significant difference between IRES activity of GT1 and GT3 constructs was observed when combined normalised data from both experiments were analysed using student t or Mann-Whitney tests (Figure 4.8).

Figure 4.6. Comparison of relative IRES activity of pRLNC (5'UTR/core) and relevant pRLN (5'UTR) clones from 6 patients in BHKsinT7 cells.

All constructs were tested in two separate experiments in duplicate wells. IRES activity of matched pRLN (5'UTR) clone from each patient was taken as 100%. Data are presented as +/- standard deviation (SD). IRES activity of GT1 (A) and GT3 (B) constructs are shown. Exp: experiment

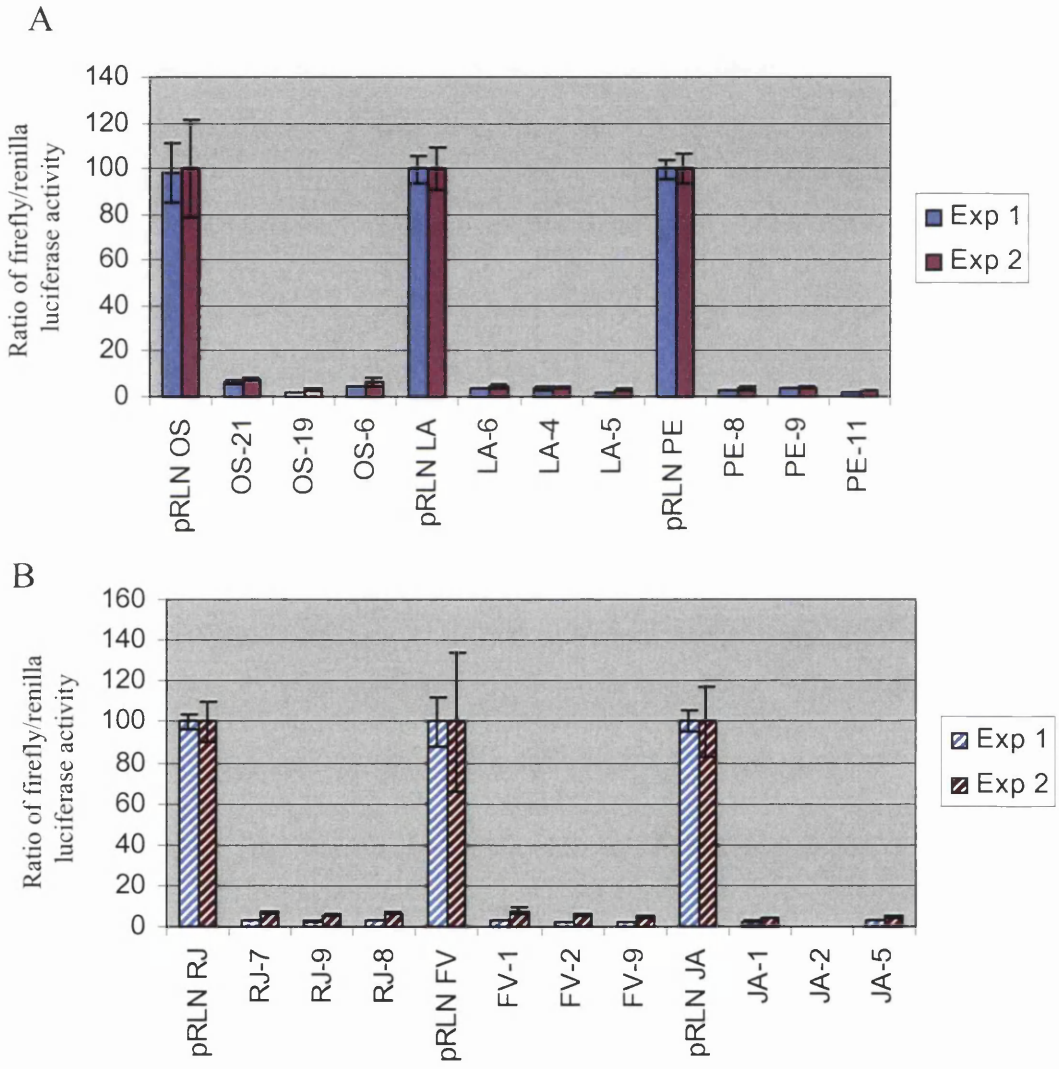


Figure 4.7. IRES activity of GT1 and GT3 pRLNC constructs in BHKsinT7 cells.

Firefly/renilla luciferase ratios obtained from constructs containing 5'UTR and core sequences tested in BHKsinT7 cells in two separate experiments (A and B). IRES activity of GT 1 and GT3 constructs are shown by white and hatched bars respectively. Mean IRES activity of GT1 and GT3 constructs were compared in each experiment separately (C). Data are presented as +/- standard deviation.

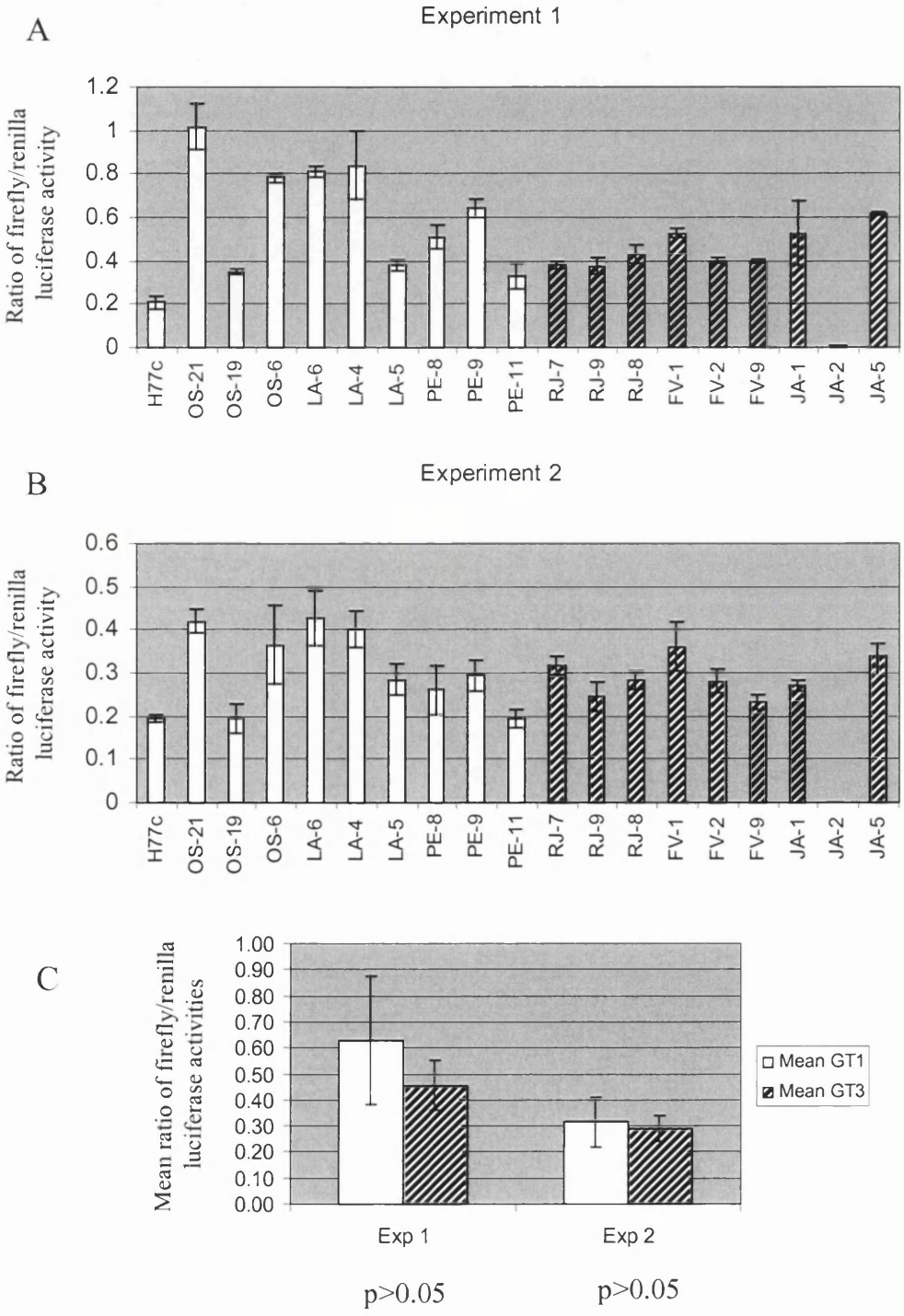
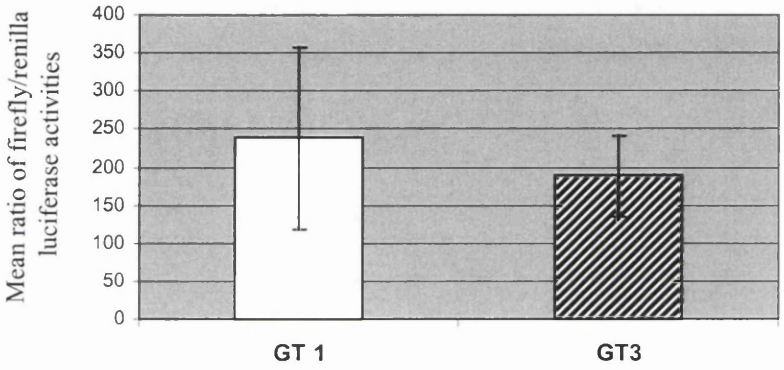


Figure 4.8. Comparison of mean IRES activity of GT1 and GT3 pRLNC constructs in BHKsinT7 cells.

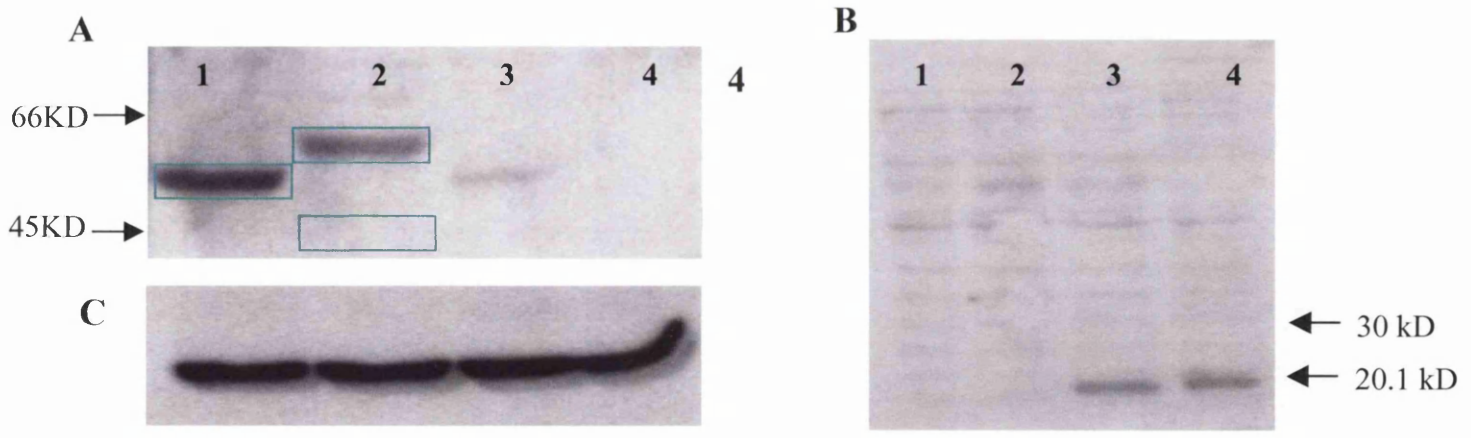
Firefly/renilla luciferase ratios obtained from each constructs were normalised against that obtained from pRLNC H77c in each experiment. The data from two experiments were combined and analysed using both student t and Mann-Whitney tests.
The data are presented as +/- standard deviation.



$p>0.05$

4.7. The effect of core on IRES activity

The results of luciferase assays revealed that the inclusion of the additional core coding sequence substantially reduced firefly luciferase light output. It is possible that fusion of the core protein to the firefly protein reduced the enzymatic activity of this protein resulting in lower production of light. In this case, the inhibitory effect of core would be overestimated. To assess better the extent to which the inclusion of the core-coding sequence affected translation, the next experiments were carried out. BHK-21 cells were infected with vTF7-3, a vaccinia virus expressing T7 RNA polymerase. This system was used in preference to the BHKsinT7 system because the proteins of interest are expressed at a higher level. The cells were then transfected with plasmid DNAs containing the 5'UTR alone (pRLN) and 5'UTR plus entire core (pRLNC) of H77c and after 16 hours, the cells were lysed and cell extracts obtained. The pGL3 vector which contains the firefly gene driven by the SV40 promoter was also transfected as a control for the correct band size of the firefly protein. Luciferase assays were carried out using extracts of cells which had been transfected with pRLN and pRLNC constructs. The ratio of firefly to renilla observed from pRLNC was 5% of that obtained from the pRLN construct (Figure 4.9.D). However, the renilla luciferase activity, which indicated the transfection level, was 4 fold lower in pRLNC than that of pRLN constructs. In order to verify these results, western blot analysis was performed to assess the amount of downstream firefly protein expression using the same cell extracts tested by the luciferase assay. Figure 4.9.A shows the SDS-PAGE analysis of the translation products which clearly demonstrated the presence of each expected protein. pGL3, pRLN and pRLNC all showed expression of the firefly protein. In cells transfected with pGL3 and pRLN the protein size (61 kD) was as expected, but pRLNC gave expression of a protein with larger size in comparison to pRLN. Figure 4.9.B shows the expression of core protein from the pRLNC construct which is similar in size to the control full length core provided by Dr. J. McLauchlan. The presence of full length core indicated that cleavage of core protein by signal peptidase at the core/firefly junction has occurred. A lower abundance of the firefly protein was produced from pRLNC compared to pRLN. Densitometric analysis of bands showed that the firefly protein obtained from pRLNC was only 60% of that from pRLN (Figure 4.9.A).



D

Construct	Firefly luciferase (RLU)	Renilla luciferase (RLU)	Firefly/Renilla ratio
pRLN	839	8150	0.1
pRLNC	9.9	2169	0.005

Figure 4.9. Analysis of the expression of firefly and core genes by Western blot analysis.

A) pRLN (lane 1), pRLNC (lane 2), pGL3 (lane 3) were transfected into BHK-21 cells pre-infected with vaccinia virus. Lane 4 contains extract from mock transfected cells. Proteins were separated by SDS-PAGE and analysed by Western blot probed with goat anti luciferase antibody. Boxes indicate regions submitted to densitometric analysis of bands obtained from pRLN, pRLNC and background. Values obtained were 4388 (pRLN), 2698 (pRLNC) and 1934 (background) i.e.1.6 fold more protein was present in band obtained from pRLN clone in comparison with that from pRLNC.B) expression of core protein in BHK-21 cell extracts. pRLN (lane 2), pRLNC (lane 3) were transfected into cells. Lane 1 contains mock transfected cells and 4, full length core protein. Arrows indicate the position of molecular weight markers. C) The blot was stripped and reprobed with the anti beta actin used as an internal control to measure the amount of cell extract loaded D) Firefly and renilla luciferase light outputs observed from pRLN and pRLNC constructs and their ratio were shown.

Anti β-Actin antibody was used as an internal control to normalize the amount of total cell lysate loaded in each lane (Figure 4.9.C). Taking into account the 4 fold less transfection level of pRLNC in comparison to pRLN construct, it would be expected that the abundance of firefly protein in pRLNC construct would be only 25% of that from pRLN. From these data, it would appear that no downregulation of firefly luciferase expression occurred in the presence of core but rather, increased amount of downstream reporter protein has been produced. It should be noted that western blot analysis was carried out twice whereas, transfection and preparation of cell extract for western blot was carried only once.

4.7.1 Glycosylation of firefly luciferase protein

As described earlier, a very low level of firefly activity was detected by luciferase analysis of pRLNC constructs in comparison with equivalent pRLN constructs and, as a result, the firefly to renilla ratio was only 3-7% of that observed from the pRLN construct suggesting that inclusion of core into the 5'UTR reduced IRES activity sharply. In contrast, more of the firefly protein was detected by western blot analysis of pRLNC in comparison with pRLN once the transfection level was adjusted for. One important difference between the translation of the two constructs is that translation of pRLNC would be expected to occur in the ER due to presence of the signal sequence at the 3' end of the core protein, whereas, pRLN would be translated in the cytoplasm. Therefore, pRLNC may be subjected to post translation modifications such as glycosylation and formation of disulfide bonds. Therefore, the amino acid sequence of the firefly protein was further examined. It was noticed that 3 predicted N-glycosylation sites were present at positions 21, 90 and 161 (Figure 4.10). This suggests that firefly protein would be further modified by N-glycosylation in the ER resulting in the production of a heavier protein. This would explain the presence of bigger firefly band obtained from pRLNC constructs in western blot (Figure 4.9.A). Unfortunately, I was not being able to perform the experiment of removing the carbohydrate portion from the protein (deglycosylation) using glycosidases. It was also shown that 4 cysteine residues were present in the firefly amino acid sequence

MKRYALVPGTIAFTDAHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVV**C**SENSLQFFMPVLGALFIGVAVAPANDIY 80

NERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRD 160

KTIALIMNSSGSTGLPKGVALPHRTA**C**VRFSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLI**C**GFRVVLMYRFE 240

EELFLRSLQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGIRQGYGLTETTSAI 320

LITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGEL**C**VRGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDED 400

EHFFIVDRLKSLIKYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTTA 480

KKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL

Figure 4.10. Amino acid sequence of firefly luciferase protein.
Showing asparagines predicted to be N-glycosylated highlighted in red and cysteine residues in blue.

suggesting the possibility of formation of disulfide bonds between different regions of the firefly protein in the ER.

4.8. Discussion

The work presented in this chapter was undertaken in the final part of my PhD. The data presented are preliminary and I was not able to answer some questions due to time limitation. The original plan was to compare the IRES activities of GT 1 and 3 using constructs containing 5'UTR plus entire core sequence matched with majority sequence obtained from each patient similar to the work carried out for 5'UTR. However, when 28 different clones were partially sequenced, no clone matching the majority sequence was found even in the limited regions of core sequenced. Therefore no further sequencing was performed due to the high cost and time constraints.

The data presented here showed that the inclusion of full length core coding sequence significantly reduced the relative firefly to renilla activity in our bicistronic expression system to 3-7% of that observed for constructs containing 5'UTR plus 15 nt of core coding sequence in BHKsinT7 cells. It has been reported that inclusion of core coding sequence (nt 1-860) in a bicistronic reporter construct, similar to what was used in our study, reduced the relative luciferase activity to 17% and 61% in HepG2 and HuH7 cells respectively with the relative luciferase activity of a construct containing nts 1-404 (i.e. 48 nt more than our pRLN constructs) normalized as 100% (Wang et al., 2000). Western blot analysis of translation products from an *in vitro* system suggested 70% reduction in IRES activity in constructs containing core coding sequence (1-860) in comparison to constructs containing 5'UTR alone (1-404). However, no such analysis was carried out on *in vivo* translation products. Another study by Zhang *et al.* (2002) reported that inclusion of full length core (nts 1-914) reduced IRES activity to 10% of that of constructs consisting of 5'UTR (nts 1-341) alone in an *in vivo* system.

It was possible that fusion of core protein to firefly luciferase altered the conformation of the expressed protein such that the enzymatic activity of the fusion protein was reduced. In this case, the degree of suppression of IRES activity might be overestimated by the

luciferase assay as previously reported (Wang et al., 2000). In order to test this, western blot analysis of cell extracts containing 5'UTR plus core was carried out. In addition, the same cell extracts were subjected to luciferase assays. The firefly/renilla ratio obtained from pRLNC constructs was only 5% of that obtained from pRLN. Western blot analysis revealed the presence of core protein which was apparently similar in size to positive control core and a heavier firefly protein. One possible explanation could be incomplete cleavage of core resulting in generation of bigger firefly protein. However, the presence of full length cleaved core did not support this scenario. Further analysis of the firefly luciferase amino acid sequence showed the presence of 3 predicted glycosylation sites. It is likely that the inclusion of the core coding sequence containing signal sequences leads to translocation of the translation machinery to the ER which would be expected to result in glycosylation of the firefly protein. It was noted that the apparent size of firefly protein was increased consistent with glycosylation occurring at the predicted sites. Unfortunately, there was no time to carry out electrophoresis following deglycosylation prior to SDS-PAGE to confirm this speculation. A lower abundance (60%) of firefly protein was produced from pRLNC compared with pRLN. However, taking into account the lower transfection level (25%) of pRLNC construct in comparison with pRLN, there would appear to be no downregulation of expression of firefly protein in pRLNC. In fact, it seemed that inclusion of core enhanced the production of firefly protein by 2 fold when the transfection level of pRLNC was taken into account. However, this experiment needs to be repeated. Another factor to be considered if repeating the work presented here, would be to measure the abundance of renilla protein in western blot analysis at the same time which would allow the quantitation of the level of transfection more accurately.

In addition to 3 glycosylation sites, 4 cysteine residues were identified in the firefly luciferase amino acid sequence. Disulfide bonds are formed by oxidation of thiol (-SH) groups in cysteine residues. The formation of disulfide bonds occurs during the folding of protein in the lumen of ER. Although disulfide bonds help to maintain the tertiary structure of protein, proper pairing of cysteine residues is very important for normal structure and activity. It is conceivable that misfolding of at least part of the firefly protein in the oxidative environment of ER lumen in addition to its glycosylation could interfere

with its enzymatic function. It is also possible that formation of misfolded protein in the ER had a negative impact on the translational machinery. It has been shown that the accumulation of misfolded proteins in the ER triggers the unfolded protein response (UPR)-signalling pathway resulting in downregulation of protein synthesis by increased phosphorylation of eIF2-α (Brewer and Diehl, 2000).

The above concerns regarding the discrepancy between results obtained from western blot analysis and luciferase assay results regarding apparent downregulation of IRES activity seen in core coding constructs has not been addressed in previous reports looking at the effect of core on translation efficiency. Our data emphasize that assessment of the inhibitory effect of core on IRES activity in bicistronic constructs cannot be made by interpretation of results obtained only from luciferase assays.

Different IRES activities were observed among clones tested from each patient. This suggests that changes in nucleotide sequence influence the IRES activity. We were not able to analyse the effect of individual substitutions since complete sequencing data was not available. No firefly activity was detected in clone JA-2. Further sequencing of this clone revealed the presence of a deletion at nt 490 within the core coding sequence. Because the -1 reading frame of the core gene contains 4 stop codons, no functional firefly would be produced due to premature termination of translation. This has been reported elsewhere (Varaklioti et al., 2002). However, the presence of a +1 ribosomal frameshift within core coding region resulting in the production of a so called "F" protein has been reported in GT 1 (Xu et al., 2001).

As described earlier, generating constructs consisting of 5'UTR and full length core coding sequence matched with majority sequence from each patient was not possible, so it was decided to test 3 individual clones from each patient. It was thought that the mean IRES activity obtained from 3 clones could be used as representative of IRES activity from patient if the differences in their activities were negligible. However, there was difference in translation efficiencies of clones obtained from one patient. Therefore, it was decided to compare the mean IRES activity obtained from all GT1 clones with those

of GT3. No significant difference was observed between translation efficiencies of constructs obtained from GT1 and GT3 samples.

It should be noted that clone JA-2 was not included in analysis as no translation of firefly luciferase gene resulted from this clone. The nonfunctional IRES sequence could result from error produced by *Taq* polymerase during PCR or by HCV RdRp during virus replication.

In summary, taken together, the preliminary results I obtained suggest that the inclusion of core into the 5'UTR did not suppress the IRES activity. Rather, it seemed to enhance its activity in our system. In addition, no significant difference was observed between translation efficiencies of the 5'UTR plus core constructs obtained from GT1 and GT3 patients.

Chapter 5

Studying the relation between viral load, liver histology and translation efficiency of HCV in patients

5.1. Introduction

Most reports on quantitative detection of HCV RNA have used either signal amplification with the branched DNA assay (Urdea et al., 1991) or target sequence amplification by RT-PCR (Levis et al., 2001; Young et al., 1995). The HCV branched DNA assay "Quantiplex Bayer"(bDNA) is a commonly used automated method but it is expensive. It is based upon specific hybridisation of virus RNA by two specific probes (capture and extender probes) to the 5'UTR and core regions of HCV RNA. The detection limit of the bDNA second generation assay was 33,000 IU/ml. The detection limit of the bDNA 3rd generation assay has improved to 800 IU/ml which still limits its use for detection of low level viremia (Hawkins et al., 1997; Hofgartner et al., 2000). The Roche Amplicor Monitor assay, which is a standardised quantitative kit based on RT-PCR, is used widely for monitoring HCV replication during therapy. The dynamic range for Roche assay extends from 300 IU/ml to 200,000 IU/ml, an approximately 3 log linear range (Martinot-Peignoux et al., 2000). Therefore, samples with higher viral load should be diluted prior to quantitative PCR.

HCV RNA can be quantitated using TaqMan technology using a dual labelled fluorogenic probe (Takeuchi et al., 1999). In this system, a reporter and a quencher are attached to 5' and 3' ends of a probe as shown in Figure 5.1. When both dyes are attached to the probe, reporter dye emission is quenched. During each extension cycle the *Taq* DNA polymerase

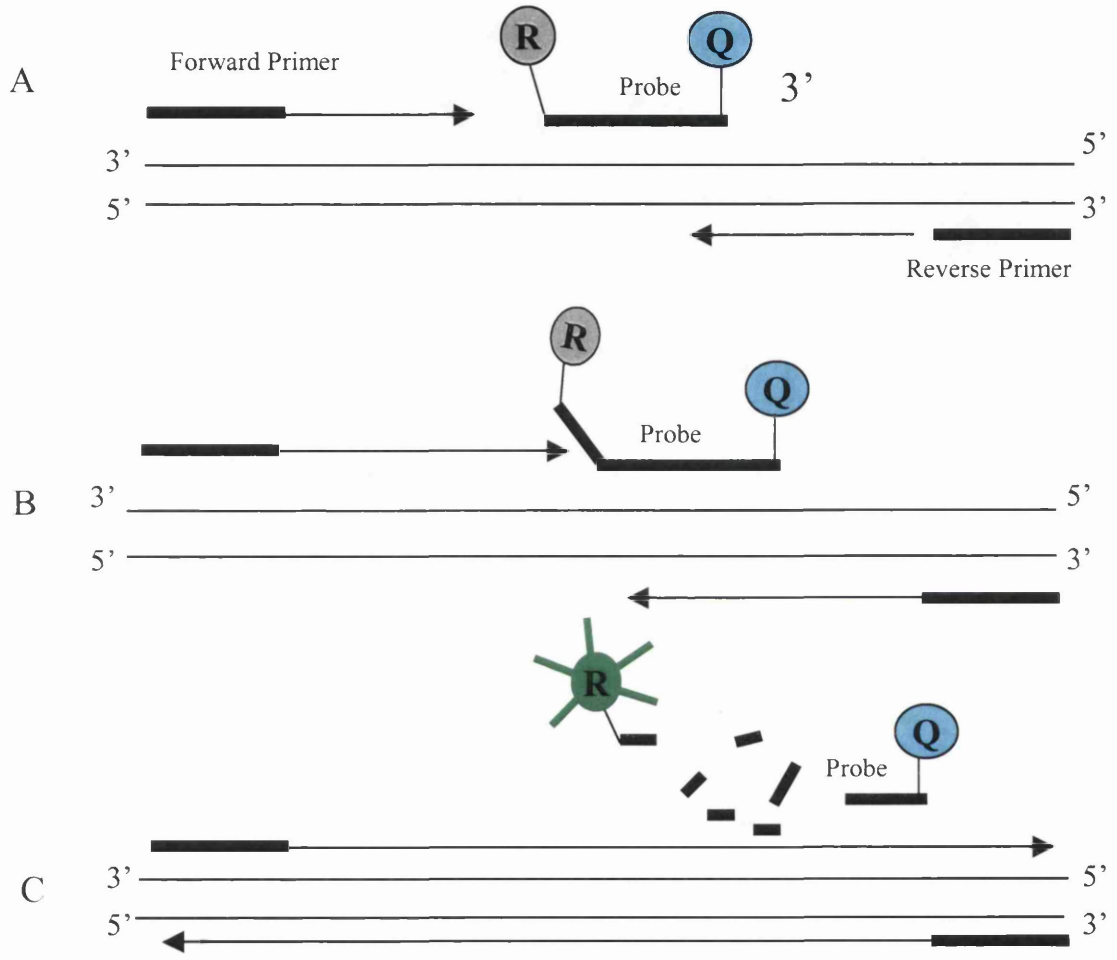


Figure 5.1. Schematic representation of TaqMan real-time PCR.

A) Primers and probe hybridise to a specific sequence of the target DNA.
B) the *Taq* polymerase enzyme cleaves the probe with its 5' to 3' nuclease activity.
C) The reporter dye and quencher dye are separated upon cleavage, resulting in fluorescent light emission by the reporter.The figure was modified from the TaqMan Universal PCR mastermix catalogue (Applied Biosystems,1998).

cleaves the reporter dye from the probe by its 5' exonuclease activity (Holland et al., 1991). This ends the activity of quencher and the reporter dye starts to emit fluorescence which increases at each cycle proportional to the rate of probe cleavage. The fluorescence signal is generated only if the target sequence for the probe is amplified during PCR. Because both primers and probe must hybridise to the target for amplification and cleavage to occur, non-specific amplification is not detected (Livak et al., 1995). The C_T value is used for quantification of the samples. The parameter C_T (threshold cycle) is defined as the cycle number at which the fluorescence emission exceeds the fixed threshold. During the early cycles of PCR, amplification of the target sequence is in the exponential phase. As the reaction components in reaction tubes are limiting, the rate of target amplification increases until a plateau is reached. Real-time PCR monitors the fluorescence emitted during the reaction as an indicator of amplicon production during each PCR cycle.

In the past, a problem with measurement of viral load was that most laboratories worldwide used different assays with no standardized HCV RNA quantification units. For example, 1 copy/ml in "Amplicor HCV Monitor", and 1 genome equivalent (Eq)/ml in "Quantiplex Bayer" did not represent the same amount of RNA in a clinical sample. Each unit was defined with standards of different lengths and sequences. This meant that these assays could not be used to decide the duration of treatment because the exact correspondence between the different units could not be evaluated. Recently, the World Health Organization (WHO) International Standard for HCV RNA quantification has been established (Saldanha, 1999). An international standard obtained from a GT 1 serum sample was designated and aliquotted into a batch of vials, each containing 50000 IU/vial in lyophilised form.

The mechanisms responsible for the onset and progression of hepatic damage during chronic HCV infection are not well known. Liver damage in chronic hepatitis C is seen initially as inflammation. Fibrosis progression, which is associated with chronic destruction of liver cells, results from chronic inflammation of the liver. Fibrogenesis is characterized by increases in collagen and other extracellular matrix constituents, such as

fibronectin and proteoglycans (Rockey, 2000). Ultimately, cirrhosis develops when annular fibrosis surrounds nodules of liver cells. The lack of any relationship between HCV viral load and the severity of liver disease suggests that HCV itself is not cytopathic. It has been suggested that the presence of liver lesions is due to the local immune response. The only lesion that can be related to a direct pathogenic effect of HCV is steatosis, which is the accumulation of intracytoplasmic lipid droplets in hepatocytes. Steatosis is a feature of hepatitis C particularly in association with GT 3 infection. Indeed, it has been reported that HCV GT 3 is the only genotype responsible for virus-induced steatosis, whereas, in patients infected with other HCV genotypes, the presence of steatosis is related mainly to other exogenous metabolic factors (Poynard et al., 2003). It has been shown that transfected HCV core protein can cause lipid accumulation by cells *in vitro* (McLauchlan et al., 2002; Perlemuter et al., 2002).

Liver biopsy is an essential investigation to evaluate the grade and stage of liver disease in HCV infected patients. Grading is used to describe the severity of inflammatory activity in chronic hepatitis. Staging is a measure of fibrosis and architectural alteration. Several scoring systems have been introduced. In the Ishak scoring system (1995), grading includes assessment of portal, periportal and intra-acinar inflammatory cell infiltration, and various forms of liver cell damage and necrosis (grading 0 to 16) and staging includes fibrosis, architectural disturbance and cirrhosis (staging from 0 to 6).

Several investigators have tried to identify non-invasive markers which correlate with the histologic changes in liver. It has been suggested that higher aspartate aminotransferase (AST) levels may be associated with more severe necroinflammatory activity on liver biopsy (Assy and Minuk, 2000). Others have found no significant correlation (Luo et al., 1998). Several groups reported that higher RNA levels in serum were associated with the presence of severe liver histopatological changes (Adinolfi et al., 2001b; Gretch et al., 1994; Kumar et al., 1994). In contrast, no correlation was reported by other studies (Anand and Velez, 2004; Lee et al., 2001; Zeuzem et al., 1996).

Changes in the secondary or tertiary structure as well as of the primary nucleotide sequence of the IRES can result in a decrease in the efficiency of protein translation. Therefore, nucleotide substitutions in the IRES could correlate with clinically significant findings such as serum HCV RNA concentration or response to interferon. Translation is one of the processes in viral replication. It is possible that a more efficient IRES may be correlated with either a higher viral load (proven to be correlated with lower response to interferon) or to a greater ability to resist the inhibitory effects of interferon. Yamamoto *et al.* (1997) studied 25 patients infected with GT 1b and found no association between sequence variation in the IRES and serum viral loads. Similar results were reported by Thelu *et al.* (2004) examining 14 patients. Several groups have attempted to correlate the variability in the 5'UTR with the response to treatment (Laporte et al., 2000; Nakazawa et al., 1994; Soler et al., 2002) but no such correlation has yet been found.

To date, the significance of translational efficiency has not been fully understood in the clinical context. HCV replicates in the liver, yet no study has attempted to relate translation efficiency to serum viral loads and histological changes observed in the liver of patients infected with HCV. In the present study, I analysed the relationship between the translation efficiency of constructs based on the nucleotide sequence obtained from each patient, serum HCV RNA concentrations and their relationship to the degree of liver damage using histological scoring of liver biopsies.

5.2. Clinical samples

Of 25 patients with chronic HCV infection studied, 13 were male and 12 female and 11 were infected with GT 1 and 14 with GT 3. No serum aliquot was available from one of the patients. The mean age was 41 years ranging from 28 to 51 years. All of these patients were referred to the Liver Clinic at Gartnaval General Hospital, Glasgow for assessment and possible treatment of chronic hepatitis C. None was on interferon-based therapy at the time when samples were collected. Serum samples were stored at -70 °C. The formalin fixed liver biopsies were all reviewed and graded by one pathologist (Dr. K. Oien, Department of Pathology, Royal Infirmary). Biopsies were graded for intensity of

necroinflammatory activity and scored for severity of fibrosis using the Ishak histology activity index (Ishak et al., 1995).

5.3. HCV RNA quantification

Serum aliquots which had not been thawed previously were used for RNA extraction. Total RNA was extracted using the QIAmp Viral Minikit (Qiagen) according to the manufacturer's instructions as described in chapter 2.2.6. RNAs were reverse transcribed in a separate reaction using Omniscript Reverse Transcription Kit (Qiagen) and antisense primer VtagRT as described in chapter 2.2.7.1. Real time PCR was carried out after a reverse transcription step. The primers EMCHCVF (sense) and EMCHCVR (antisense), which had been designed previously in our laboratory to amplify a 166 bp fragment within the 5'UTR of the HCV genome, were used (Table 5.1). These primers match completely conserved sequences between GT 1 and 3 of HCV. The probe binding site is also completely conserved in all known isolates. DNA probes with conjugated minor groove binder (MGB) groups form extremely stable duplexes with single-stranded DNA targets, allowing shorter probes with higher melting temperature to be used for hybridisation based assays. PCR was performed using cDNAs, TaqMan Universal PCR Master Mix (Applied Biosystems) and the GeneAmp 5700 sequence detector system (Perkin Elmer, Foster City, CA) as described in chapter 2.2.8.3. (Figure 5.2). The fluorescence released from the TaqMan probe during PCR was proportional over the exponential phase of accumulation to the amount of PCR product, and thus to the amount of original HCV RNA in the sample. The real-time data of increasing fluorescence during PCR were plotted and the point at which the amplification plot crossed the threshold was defined as C_T. The amount of HCV RNA in the sample was calculated from C_T value using the standard curve which was generated from HCV positive sample 2000/187 previously calibrated to the World Health Organization (WHO) First International Standard for HCV RNA (NIBSC) (Saldanha, 1999) and also to synthetic transcribed RNA (Figure 5.3). Lyophilised WHO material was reconstituted in deionised water to a concentration of 100,000 IU/ml.

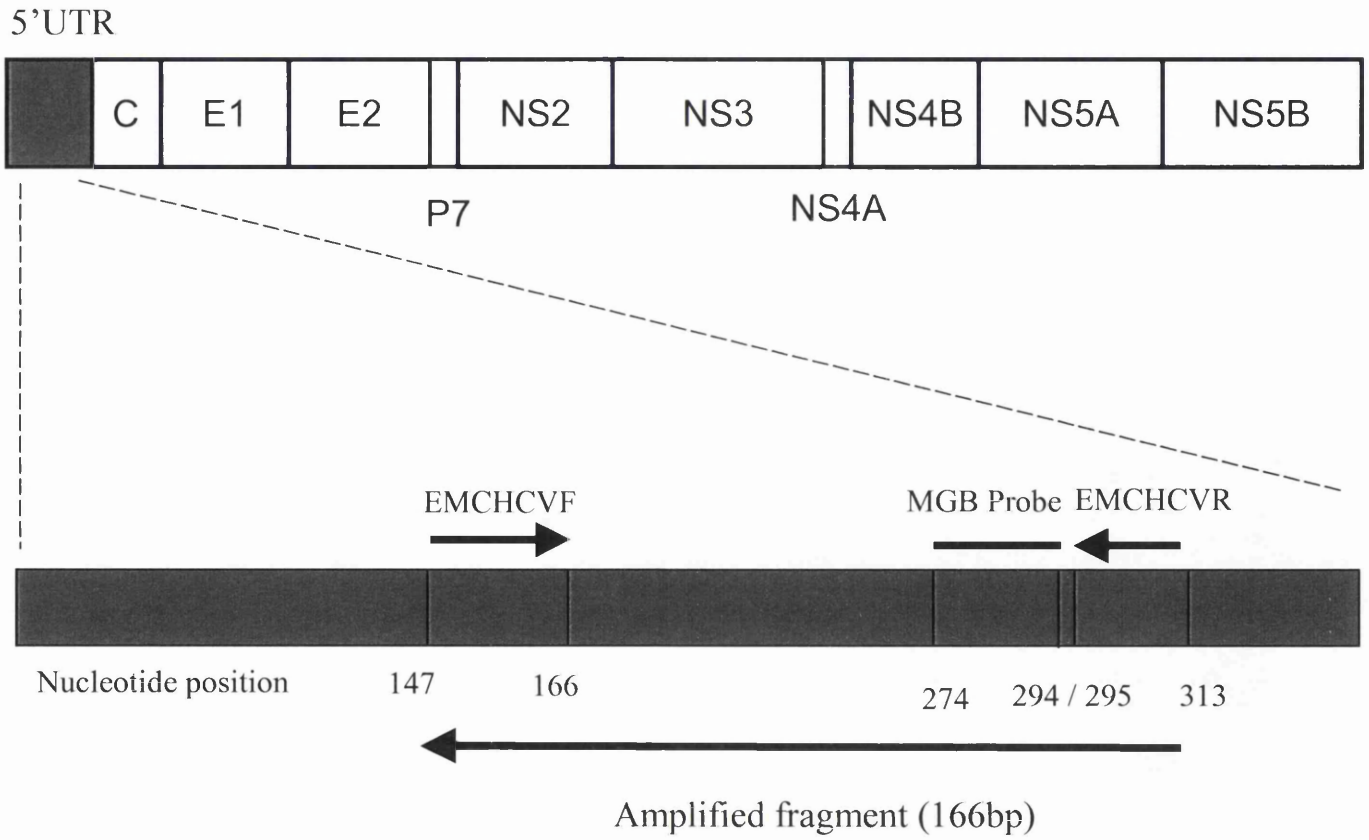


Figure 5.2. Schematic representation of the primers and probe (HCV MGB) used in the 5'NCR of HCV genome.

The primers EMCHCVF as sense and EMCHCVR as antisense at the positions as indicated were used for the amplification of HCV RNA in serum samples. The length of amplified fragment is 166 bp as indicated. HCV MGB probe at the position as indicated was used to detect product accumulation.

Figure 5.3. Standard curve of 3 replicates of serial dilutions of transcribed RNA, plotting quantity versus threshold cycle (C_T). The vertical scale is the cycle where PCR product crosses the threshold. The horizontal scale is the log of the starting copy number. The linear relationship between C_T value and initial transcribed RNA copy number is shown over five orders of magnitude.

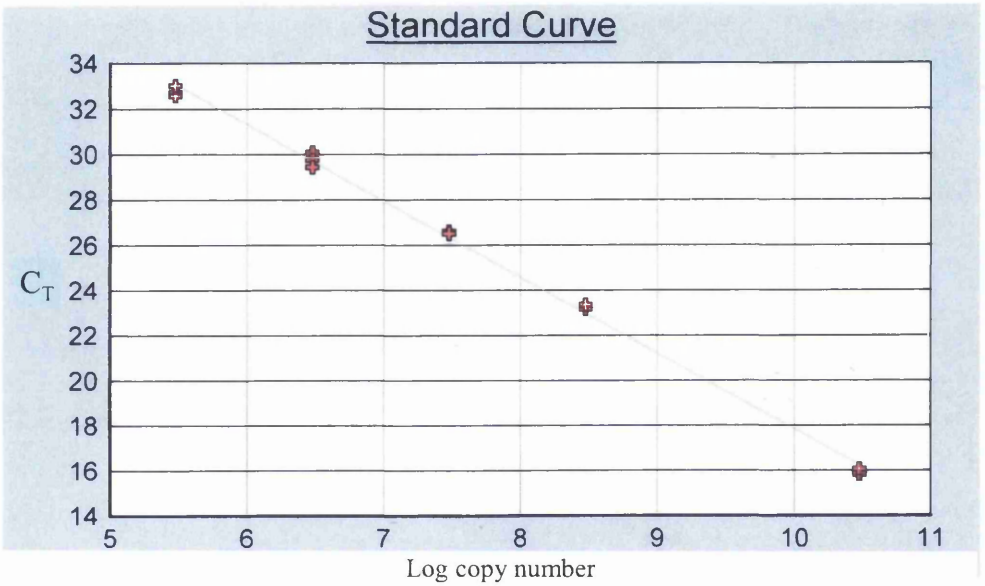


Table 5.1 Oligonucleotides primers for quantitation of RT-PCR

Primer	*Sequence 5' to 3'*	*Position*	*Use*
VtagRT (as)	TTTTTCTTTGAGGTTTAGGA	353-372	RT
EMCHCVF (s)	TCTGCGGAACCGGTGAGTAC	147-166	PCR
EMCHCVR (as)	TCTGCGGAACCGGTGAGTAC	295-313	PCR
HCV MGB (Probe)	TCTGCGGAACCGGTGAGTAC	272-294	PCR

Nucleotide numbering according to H77c (AF 011751).

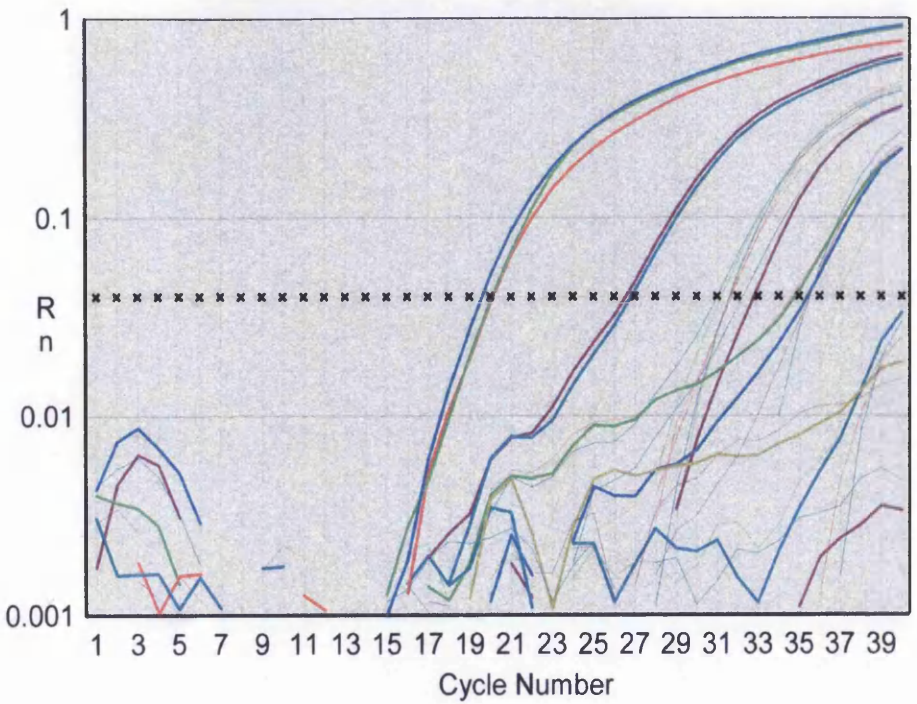
s = sense

as= antisense

5.3.1. Calibration of the assay

All reactions were quantified against an internal HCV positive serum (2000/178) of known titre ($3x10^6$ IU/ml) which had been previously measured by C.A. Smith in our laboratory. The sample was used neat and at dilutions of 1:3, 1:30 and 1:300 to generate a standard curve. To increase the dynamic range of the assay, standard RNA dilutions which had been previously synthesised by C.A. Smith using one of the constructs made by C. Nairn in R. Elliott's laboratory, were used in parallel with the samples. This RNA contains a sequence from the 5'UTR of HCV. In order to recalibrate the transcribed RNA against serum 2000/178, real time PCR was carried out using serum 2000/178 and transcribed RNA in triplicate wells. C_T values obtained from $1:10^6$ dilution of transcribed RNA was equivalent to that obtained from neat 2000/178 (C_T~31) (Figure 5.4). Dilutions of $1:10^7$ to $1:10^2$ containing 10^5 to 10^{10} copy/ml from transcribed RNA were used in the same reaction to ensure that samples with higher viral loads were measured accurately.

A



B

Sample	Dilution	CT	Quantity	Mean quantity
2000/178	neat	31.58	3.00E+06	3.00E+06
2000/178	neat	31.5	3.00E+06	3.00E+06
2000/178	neat	31.54	3.00E+06	3.00E+06
2000/178	1/3	32.75	1.00E+06	1.00E+06
2000/178	1/3	32.21	1.00E+06	1.00E+06
2000/178	1/3	32.21	1.00E+06	1.00E+06
2000/178	1/30	34.59	1.00E+05	1.00E+05
2000/178	1/30	34.78	1.00E+05	1.00E+05
2000/178	1/30	35.35	1.00E+05	1.00E+05
2000/178	1/300	40	1.00E+04	1.00E+04
2000/178	1/300	40	1.00E+04	1.00E+04
2000/178	1/300	40	1.00E+04	1.00E+04
transRNA	1/1000	20.09	4.95E+09	5.76E+09
transRNA	1/1000	20.06	5.05E+09	5.76E+09
transRNA	1/1000	19.53	7.27E+09	5.76E+09
transRNA	1/100000	26.51	6.26E+07	4.95E+07
transRNA	1/100000	26.79	5.14E+07	4.95E+07
transRNA	1/100000	27.38	3.45E+07	4.95E+07
transRNA	1/1000000	31.64	1.89E+06	2.48E+06
transRNA	1/1000000	31.17	2.60E+06	2.48E+06
transRNA	1/1000000	30.99	2.93E+06	2.48E+06
transRNA	1/10000000	35.42	1.44E+05	1.13E+05
transRNA	1/10000000	35.44	1.44E+05	1.13E+05
transRNA	1/10000000	36.91	5.21E+04	1.13E+05

Figure 5.4. Representative results of quantitative real time reverse transcription PCR assays

A) The amplification plot window showing fluorescent values plotted vs. the PCR cycle number.
The threshold cycle (dotted line) was set in the middle of the linear region of amplification curves. B) The report window displays the analysed data in triplicate wells. Serial dilutions of serum 2000/178 was used as standard control in order to recalibrate transcribed RNA. Similar C_T values obtained from neat 2000/178 and transcribed RNA (transRNA)1/10 ^{6} dilution.

5.3.2. Comparison of TaqMan real time PCR and COBAS Amplicor Monitor Assay for the quantitation of HCV PCR

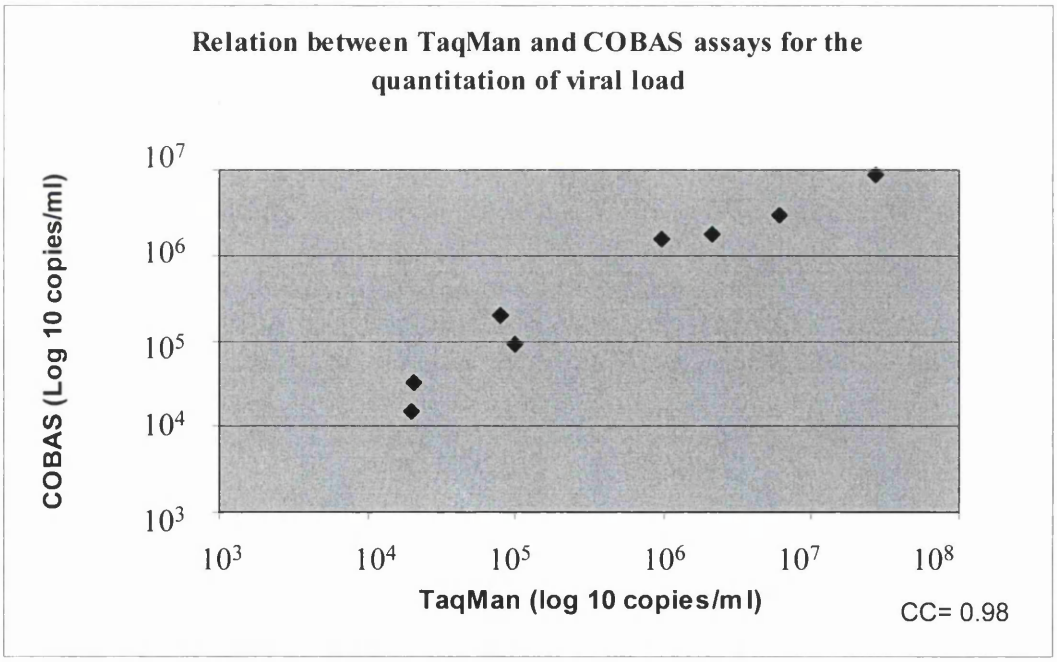
In order to confirm that the "in house" TaqMan HCV quantitative test accurately measured viral load from different samples, it was decided to measure viral load in a number of samples using the COBAS Amplicor HCV system (Roche) which is routinely used for quantitation of clinical samples. Therefore, 8 serum samples including 2 from patients with the highest and 2 from patients with lowest viral loads quantified by TaqMan from each group of GT1 and GT3 patients were chosen. Fresh previously unthawed serum aliquots from these samples were sent to Regional Virus Laboratory at Gartnaval General Hospital and quantification was carried out by G. Gunning. Although the upper limit of the linear range claimed by the manufacturer for the Amplicor Monitor test is 850,000 IU/ml, other studies indicate that the test reaches a plateau at concentrations above 500,000 IU/ml (Konnick et al., 2002). In the first experiment using COBAS Amplicor, all 4 samples were reported as higher than the upper limit of detection of the assay. Therefore, a second experiment was carried out using 1:10 or 1:100 dilutions of these samples and the results were corrected for the initial dilutions. As shown in Figure 5.5.A and graphically in figure 5.5.B, the HCV RNA levels measured by COBAS Amplicor assay correlated significantly with those measured by TaqMan PCR assay in patients infected with GT1 and GT3 (correlation coefficient (CC) = 0.98, Spearman test).

Figure 5.5. Comparison of viral load quantitated by TaqMan and COBAS Amplicor.

A

Patient	Genotype	TaqMan	COBAS Amplicor
RJ	3	6×10^6	3×10^6
ME	3	2.7×10^7	9×10^6
JN	3	1×10^5	9.2×10^4
CP	3	7.9×10^4	2×10^5
LA	1	9.7×10^5	1.5×10^6
TW	1	2.1×10^6	1.8×10^6
BH	1	2.1×10^4	3.3×10^4
BA	1	2×10^4	1.5×10^4

B



5.3.3. Quantitation of viral load in serum samples

HCV RNA in samples from 25 patients (11 GT 1, 14 GT 3) was quantified using the TaqMan real time PCR assay. The copy number was determined based on a standard curve drawn using the known amount of the internal control serum 2000/178 as described above. The measurement of samples was performed in triplicate. The amount of HCV RNA in the serum samples was defined as the mean of the triplicate data. HCV RNA was detected in all samples tested as shown in Figure 5.6. Serum viral loads in samples varied from 4.8×10^4 to 3.3×10^7 IU/ml. Among GT 1 samples, a mean viral load of 2.5×10^6 IU/ml was observed, whereas in GT 3 samples, the mean viral load was 8.2×10^6 IU/ml. When using an unpaired Student's *t* test for comparisons between the two genotypes, no significant difference in viral load was found ($p = 0.07$).

5.4. Liver histology scoring

Liver biopsies from 21 of the 25 patients from whom sections could be found in the pathology department were assessed by one liver pathologist (Dr. Dr. K. Oien, Department of Pathology, Royal Infirmary) without knowledge of the patients' characteristics including genotype or 5'UTR sequence. Biopsies were scored according to the Ishak scoring system (Ishak et al., 1995) and recorded as inflammation grade and fibrosis score respectively as shown in Table 5.2. The fibrosis score reflected different stages of fibrosis, including none (0 points), fibrous expansion of some portal tracts (1 point), fibrous expansion of most portal tracts (2 points), occasional portal-to-portal bridging (3 points), marked bridging (4 points), bridging with occasional nodules (5 points), and cirrhosis (6 points). The inflammation grade reflected composite score of grades for interface hepatitis, confluent necrosis, focal inflammation, and portal inflammation. According to this scoring system, the maximum was 18 points. Of 21 patients only 4 had fibrosis stage 1 to 3 and no patient had cirrhosis. Inflammatory grades ranged from 2 to 5. The Ishak scores were then analysed to check for presence of any correlation with viral load and translation efficiency.

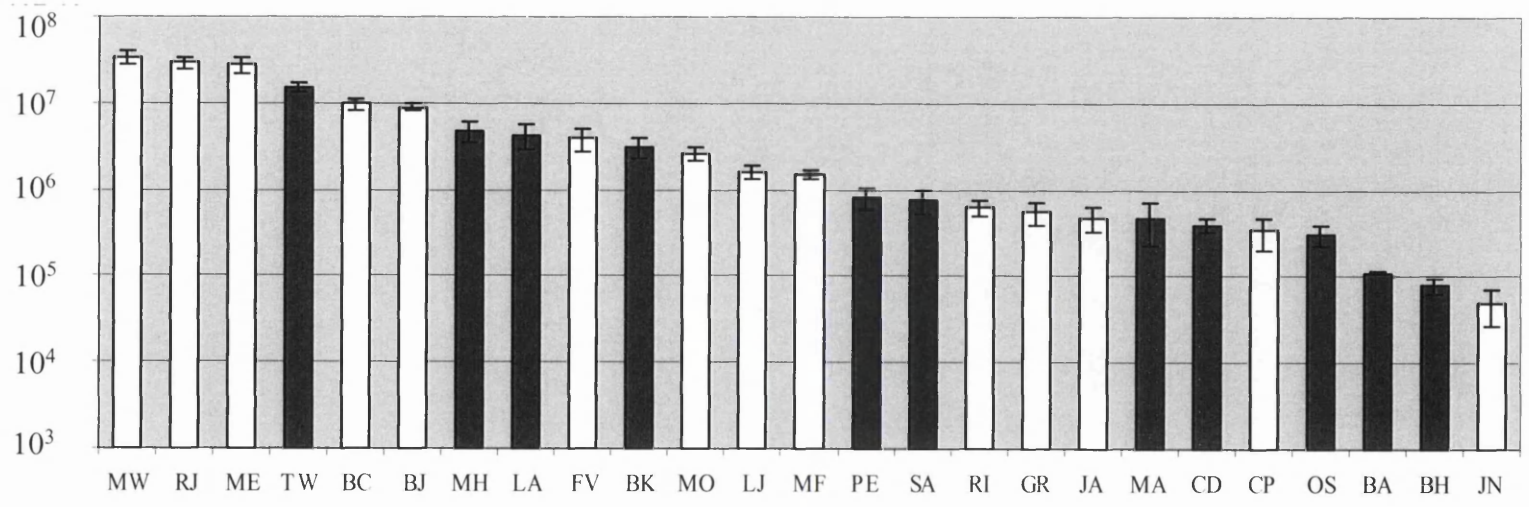


Figure 5.6. Quantitation of the HCV-RNA serum levels by TaqMan real time PCR assay.

Serum samples from 25 patients were assayed in triplicate and the mean values and standard deviations were calculated. GT1 and GT3 samples are indicated by black and white columns respectively in decreasing order.

Table 5.2. Ishak and steatosis scores in liver biopsies from patients infected with GT1 and GT3.
NF=No biopsies available

Patient	Genotype	Grading	Staging	Steatosis
		Inflammation	Fibrosis	
MA	1	3	0	0
TW	1	2	1	2
BA	1	5	0	0
OS	1	5	0	0
BK	1	2	0	0
MH	1	3	0	0
PE	1	NF	NF	NF
SA	1	4	0	0
CD	1	5	3	1
LA	1	3	0	1
BH	1	3	0	0
FV	3	3	1	2
MO	3	2	0	0
GR	3	5	0	1
RJ	3	4	0	3
ME	3	5	0	2
MF	3	5	1	3
LJ	3	4	0	2
BC	3	NF	NF	NF
JN	3	5	0	1
MW	3	NF	NF	NF
SJ	3	3	0	1
BJ	3	4	0	1
RI	3	4	0	1
JA	3	NF	NF	NF
CP	3	NF	NF	NF

In addition, biopsy sections were evaluated for the presence of steatosis. Steatosis was graded as absent or minimal (less than 1% of hepatocytes) (score 0), mild (<30% hepatocytes involved) (score 1), moderate (between 30 and 60% of hepatocytes involved) (score 2) or severe (>60% of hepatocytes involved) (score 3). Steatosis of the liver was observed in all GT3 infected patients except one, with scores ranging from mild (5 cases) to moderate (3 cases) to severe (2 cases). In GT 1 infected patients, only 3 out of 10 patients had steatosis (2 mild, 1 moderate).

5.5. Analysis of the correlation between IRES activity, viral load and histological features

Table 5.3 summarises the data used for statistical analysis. Data analyses were performed using SPSS version 11.0 (SPSS, Inc., Chicago IL). The correlations between translation efficiencies, Ishak scores and HCV RNA titres were analysed by the nonparametric Spearman rank-order correlation coefficient. A p value less than 0.05 was considered statistically significant.

The relationship between the translation efficiency of the 5'UTR obtained from patients in HuH7 cells as described in chapter 3 and HCV RNA concentrations in serum of patients was investigated. In order to assess correlations, all patients having an identical 5'UTR sequence were considered to have the same relative IRES activity as the representative 5'UTR isolate in HuH7 cells. The mean IRES activity (normalised) obtained from 3 separate experiments from each isolate was used for analysis. The IRES activity did not correlate with serum HCV RNA levels in the patients studied (Correlation Coefficient (CC) = -0.34, $p > 0.05$).

The relationship between the IRES activity of isolates obtained from the translational study and the liver histology from patients containing the identical 5'UTR to those isolates were analysed. Similarly, no association was found between the grade of liver inflammation and the corresponding translation efficiency (CC= 0.17, $p > 0.05$). There was

also no significant correlation between circulating HCV RNA titres and the degree of inflammation (CC= -0.3, p> 0.05).

When we compared the steatosis scores with the HCV genotype, we found a significant correlation (p=0.008) with the presence of GT 3.

Table 5.3. Patient data used for statistical analysis

Patient	GT	Inflammation grade	Fibrosis stage	Steatosis grading	Viral load (IU/ml)	RTE in HuH7 cell %
MA	1	3	0	0	$4.7x10^5$	100
TW	1	2	1	2	$1.51x10^7$	110.6
BA	1	6	0	0	$1.3x10^4$	100.7
OS	1	6	0	0	$3.1x10^5$	110.6
BK	1	2	0	0	$3.1x10^6$	110.6
MH	1	2	0	0	$4.8x10^6$	100
PE	1	NF	NF	NF	$8.3 x10^5$	97
SA	1	4	0	0	$7.6 x10^5$	107
CD	1	5	3	1	$4 x 10^5$	110
LA	1	2	0	1	$4.3 x10^6$	97.4
BH	1	2	0	0	$7.8 x10^4$	115.3
FV	3	3	1	2	$3.9 x10^6$	90.4
MO	3	1	0	0	$2.6 x10^6$	120
GR	3	6	0	1	$5.5 x10^5$	120
RJ	3	4	0	3	$2.9 x10^7$	62.8
ME	3	6	0	2	$2.7 x10^7$	Not done
MF	3	5	1	3	$1.5 x10^6$	120
LJ	3	4	0	2	$1.6 x10^6$	90.4
BC	3	NF	NF	NF	$9.6 x10^6$	107.8
JN	3	6	0	1	$4.8 x10^4$	114.9
MW	3	NF	NF	NF	$3.3x 10^7$	120
SJ	3	4	0	1	NF	92
BJ	3	4	0	1	$8.8 x10^6$	120
RI	3	4	0	1	$6.3 x10^5$	82.3
JA	3	NF	NF	NF	$4.7 x10^5$	123.6
CP	3	NF	NF	NF	$3.4 x10^5$	107.8

RTE= relative translation efficiency
NF= not available
For details of scoring system, see text

5.6. Discussion

The aim of this part of the study was to address whether there was a correlation between the degree of histological damage, the translation efficiency of the IRES and the HCV RNA titres in patients infected with GT1 and GT3. Using the TaqMan real time PCR assay, the HCV RNA levels in serum samples from 25 patients in our cohort were measured. All liver biopsies in this study were assessed by a single histopathologist to prevent inter-observer error.

In this study, using synthetic RNA, the linearity of the assay was conserved over a wide range of HCV copy numbers ranging from 10^4 to 10^{10}. Therefore, it was ensured that no sample with a higher viral load than our internal HCV positive control was missed. None of our samples was measured at less than 10^4 IU/ml. In comparison to other quantitation methods, real-time PCR measures a much wider range of viral loads. The dynamic range of any assay is determined as how much target concentration can vary and still be quantified with equal sensitivity and specificity. Using TaqMan system, a wide range of detection has been reported ranging from 10^1 to 10^{10} (Enomoto et al., 2001).

The result of this study concurs with other studies in finding no significant differences in virus load among patients infected with GT1 and GT3 (Lau et al., 1996; Smith et al., 1996). The presence of a higher viral load in serum of patients infected with GT 1 in comparison to GT2 and GT3 has been reported by several studies (Mahaney et al., 1994; Orito et al., 1994) and it has been argued that this difference might be related to the increased probability of achieving response to treatment observed in GT 2 and 3. These analyses require that the assays used to determine virus load are equally sensitive for each genotype. There is evidence that this was not the case for the original branched-DNA (bDNA-1) assay with a reported 2 fold reduction in the efficiency of detection of GT3 compared with those of GT1 (Hawkins et al., 1997). The current version of the Amplicor assay has been shown to amplify all known genotypes with equal efficiency. The

quantitative PCR assays used in this study ought to measure accurately viral load in all samples with the same efficiency because the 5'UTR sequences which were used in design of primers and probe are completely conserved in all isolates of all genotypes in GenBank. In order to confirm this, 4 serum samples including 2 patients with highest and 2 patients with least viral load quantified by TaqMan from each group of GT1 and GT3 patients were quantified by COBAS Amplicor HCV system by an independent worker blindly. Similar results to those obtained from TaqMan were reported using COBAS Amplicor test with a significant correlation between the results obtained from the two assays.

Sequence variability in the 5'UTR may cause variations in the efficiency of viral translation. It is conceivable that variation in translation efficiency may also correlate with clinical features such as response to interferon, serum viral load or extent of the liver damage. It had been previously been shown that the sequence variability of the IRES did not appear to correlate with any difference in serum HCV RNA concentration (Yamamoto et al., 1997). They examined pre-treatment sera from 25 patients with GT1b infection and found no correlation between the number of nucleotide changes in the IRES with viral load or interferon response. We believe this is the first study in which correlations between translation efficiency, serum viral RNA and liver histology has been studied. The data presented here did not support such correlations. However, an association may have been masked in our study, as, by chance, most patients had very mild disease on liver biopsy. Against this, the range of viral loads seen was similar to that seen in patients in general with hepatitis C including those with more severe liver disease (Personal communication, Dr. E.A.B, McCruden).

Studies assessing the relationship between serum viral titres and the severity of histological abnormalities have reported conflicting results. Some found no correlation between HCV viral loads and the extent of histological damage (Lee et al., 2001; McCormick et al., 1996; Zeuzem et al., 1996). On the other hand, Fanning *et al.* (1999) in a study on 77 Irish women who acquired their HCV infection through the administration of contaminated anti-D immunoglobulin, observed a significant correlation between

serum HCV viral loads and the degree of hepatic inflammation in liver biopsy samples using the Ishak scoring system. In the present study, no correlation was observed between serum HCV RNA titres and the degree of hepatic inflammation supporting the idea that the severity of liver disease was independent of serum levels of hepatitis C virus. It is also possible that the number of cases studied in this project was not sufficient to show a correlation. The precise mechanism by which hepatitis C virus damages the liver is not known. It has been reported that a direct cytopathic effect of the virus is responsible for the primary form of liver injury. However, it is widely accepted that the degree of liver damage is more likely to be the result of an interaction between the virus and immune response of the host (Rehermann, 2000). The results of the present study along with others argue against a direct cytopathic effect of HCV.

Patients infected with GT 3 showed a higher prevalence of steatosis than those infected with GT1 in our study. It has already been shown that the prevalence of steatosis among the HCV genotypes is significantly different; patients infected with GT 3 showing the highest prevalence followed by type 2 infection and lastly type 1 infection (Adinolfi *et al.*, 2001a; Rubbia-Brandt *et al.*, 2004). The mechanisms responsible for development of steatosis in HCV infection are not understood. In transgenic mice, it has been shown that the HCV core protein induces hepatic steatosis (Moriya et al., 1998). Core protein expression within the mitochondria alters the double membrane structure and causes an impairment of lipid oxidation, which produces steatosis (Moriya et al., 1998). It has been reported that steatosis may influence liver fibrosis progression in GT 3 infected patients. In patients with HCV genotypes other than 3, progression of liver disease may depend on other factors, such as prolonged alcohol abuse or being overweight (Rubbia-Brandt et al., 2004).

Our data demonstrate that the observed differences in circulating viral loads and biopsy scores could not be explained by the translational activity by the IRES measured in our system in Huh7 cells.

Chapter 6

Quasispecies composition of 5'UTR in serum and liver

6.1. Introduction

A characteristic of the HCV genome, like those of other RNA viruses, is sequence heterogeneity. The HCV population in each individual consists of mixture of a genetically different but closely related variants called a quasispecies (Martell et al., 1992). New variants are generated during virus replication as a result of errors made by the viral RNA-dependent RNA polymerase, which lacks proofreading activity. A quasispecies distribution within a population can be studied using different methods. The sequencing of cloned RT PCR products was the methodology first used to analyse HCV quasispecies (Martell et al., 1992). However, both cloning and sequencing are time-consuming and may not accurately reflect the true nature of the quasispecies makeup. Therefore, some other techniques based on differential gel electrophoresis mobility were developed such as single-strand conformation polymorphism analysis (McKechnie and McCruden, 2001), and heteroduplex tracking analysis (Gretch et al., 1996).

In the use of cloning and sequencing for determination of quasispecies composition, nucleotide sequence information is obtained through PCR amplification of virus-specific cDNA produced by reverse transcription of viral RNA. The enzymes used through these two steps, a reverse transcriptase (RT) and a thermostable DNA polymerase, exhibit relatively high error rates. Therefore, such error rates have to be taken into account together when considering the actual heterogeneity within a viral population (Smith et al., 1997b).

The 5' UTR, even though the most highly conserved part of the virus genome, shows a quasispecies distribution (Laporte et al., 2000; Lu et al., 2000; Malet et al., 2003; Soler et al., 2002; van Leeuwen et al., 2004). Soler *et al* (2002) sequenced 360 clones from 6 patients infected with different HCV genotypes and showed that most substitutions were in unpaired regions of 5'UTR or clustered such that base pairing was maintained. Substitutions in paired regions of the 5'UTR would be expected to alter the stem loop structure of the IRES if there were not compensating substitutions such that base pairing was maintained. Alteration of the stem-loop structure would be expected to result in effects on the translation efficiency of the IRES. No nucleotide sequence changes were observed after interferon-α therapy suggesting that the quasispecies distribution of IRES sequences does not play a role in HCV resistance to interferon. Laporte *et al.* (2000) examined the sequence of 5' UTRs from the quasispecies characterized in the serum of a chronic HCV GT 1a infected patient and its corresponding translational activity. They showed that sequence heterogeneity between IRES elements led to important changes in their translation efficiency both *in vitro* and in different cell lines.

Differences in the composition of the quasispecies between that circulating in the blood and that in the liver within an infected patient have been reported in several studies. However, most of these studies were conducted on the E2 HVR1 region, in which the number of variants within quasispecies is expected to be especially high (Cabot et al., 2001; Cabot et al., 2000; Navas et al., 1998). Only one report (Cabot et al., 1997) was found in which the 5'UTR quasispecies was compared in paired serum and liver samples from 1 patient. They cloned and sequenced the 5'UTR from paired serum (9 clones) and three different parts of liver explant (33 clones) of a patient undergoing liver transplantation and showed the presence of the same master sequence in both tissues. The proportion of master sequence in the three liver samples was variable, while the frequency of any other variants never surpassed 5% of total clones sequenced. In serum, the Pn ratio (the ratio of number of polymorphic sites to number of nucleotides sequenced) was lower than those in the liver samples. As can be seen, the comparison between 5'UTR in serum and liver has not been investigated properly.

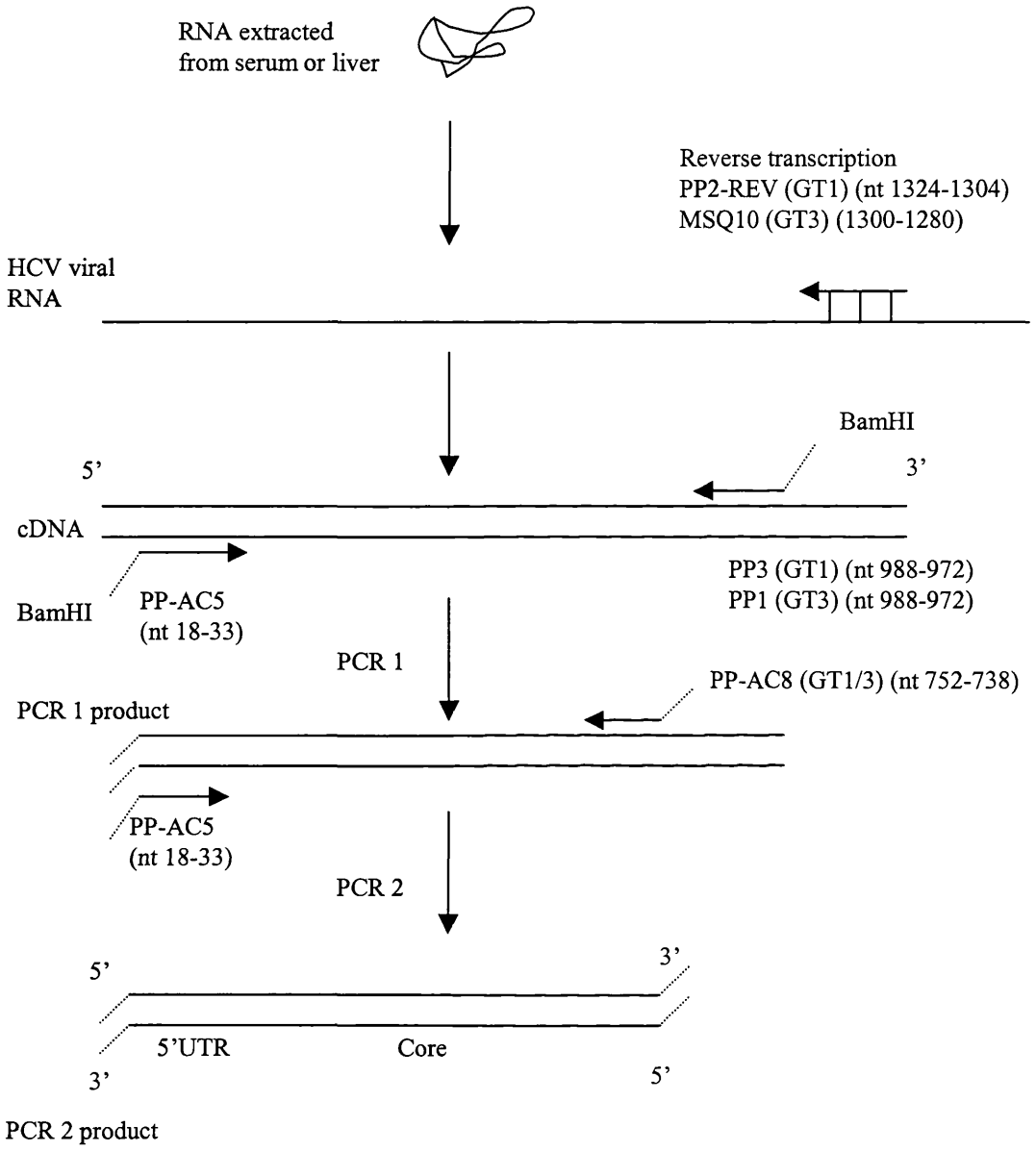


Figure 6.1. Schematic representation of 5'UTR/core RT-PCR.

This shows the amplification of the 5' UTR and core region from serum and liver fron HCV GT 1 and 3 infected patients by reverse transcription and amplification of cDNA by seminested PCR. Bases corresponding to a *BamHI* digestion site were added to the 3' ends of the relevant primers to permit easy cloning and are represented by dotted lines. PCR 2 (seminested) was carried out only on serum and liver samples from patient LA and the liver sample from patient BH.

The 5'UTR is the most conserved region of the HCV genome and contains an internal ribosomal entry site which directs translation of viral proteins (Honda et al., 1996b). Therefore, any change in the nucleotide sequence of 5'UTR may influence the amount of protein that is expressed. Primary work carried out by P.Preikschat in our laboratory showed that majority sequences obtained from serum and matched liver samples were identical in 26 HCV infected patients studied. It was hypothesized that the liver might contain minor variants with lower translational activity which allows sufficient protein expression for viral replication but insufficient levels to induce immune recognition of infected cells resulting in persistent infection. However, the majority sequencing data did not support or refute this hypothesis. The work described in this and the next chapter were carried out in order to examine the above hypothesis using two techniques. In this chapter, using a cloning and sequencing strategy, I have compared the circulating and intrahepatic quasispecies composition of the HCV 5'UTR derived from serum and liver tissues of six chronically infected patients to answer whether there are any differences in quasispecies makeup of the 5'UTR region between serum and liver.

6.2. Study patients and samples

We chose, at random, paired serum and liver samples from 3 patients infected with GT1 and 3 with GT3 from our cohort. None of the patients had received interferon treatment before the samples were obtained. Patients SA, LA and BH were infected with GT 1 and patients MO, RJ and ME with GT 3.

6.3. Amplification of 5'UTR/core from patients and comparison of resulting "majority" sequence

All the following steps were carried out by P. Preikschat in our laboratory. As shown in Figure 6.1, extracted RNA from non-thawed paired serum and liver samples was reversed

transcribed to generate cDNA using Superscript II RT (Invitrogen). Single rounds of PCR involving 38 cycles were carried out for patients SA, MO, RJ, ME and cDNA from serum from patient BH (Table 6.1.A). In order to increase the yield of PCR product, for serum and liver samples from patient LA and the liver sample from patient BH, seminested PCR involving 58 cycles was carried out using primers PPAC5 and PPAC8 (Table 6.1.B).

Comparison of sequences obtained by direct sequencing of PCR products obtained from GT 1 infected patients with reference sequence H77c (EMBL accession number AF 011751) revealed that all had a difference with G at nt 107 instead of A (G107A). Patient LA had a mix of G and A at nt 233 indicated by (R) and patient BH had A204C (Figure 6.2A).

Differences in sequence in GT 3 patients were G203A in patient MO, C121U in patient RJ and A340C in patient ME when compared with reference sequence NZL1 (EMBL accession number D17763). (Figure 6.2B)

6.4. Cloning of 5'UTR/core into pCR 2.1 vector

Gel purified PCR products generated by P. Preikschat as described in the previous section containing the 5'UTR and core of HCV were ligated by me into the pCR2.1 TOPO vector (Invitrogen), as described in chapter 2.2.10.2. Fourteen to 27 independently isolated clones from PCR products were selected using X gal selection. Plasmid DNAs containing the 5'UTR sequences were purified using columns (Miniprep Kit, Qiagen). Clones were selected for the presence of the correct size insert by digestion with *EcoRI*, which is present on both sides of the cloning site. The orientation of the insert within the vector was determined by cleavage with *Hind III* and *ClaI* restriction enzymes as can be seen in Figure 6.3. DNA from 202 clones was quantified by spectophotometry and nucleotide sequenced.

Table 6.1. Amplification of 5'UTR and core region from six patients.

(A) The number of PCR cycles for each sample is shown. Seminested PCR was carried out for serum and liver samples from patient LA and the liver sample from BH.
B) Oligonucleotide primers for RT-PCR of genotype 1 and 3 5'UTR and core regions.

A

Patient	*GT*	*Sample*	*PCR cycles*	*Comment*
S.A.	1	Serum	38	
	1	Liver	38	
L.A.	1	Serum	58	Seminested
	1	Liver	58	Seminested
B.H.	1	Serum	38	
	1	Liver	58	Seminested
M.O	3	Serum	38	
	3	Liver	38	
R.J.	3	Serum	38	
	3	Liver	38	
M.E	3	Serum	38	
	3	Liver	38	

B

Primer	*Sequence (5' to 3')*	*Genotype*	*Position*	*Use*
PP2-REV (as)	GACCAGTTCATCATCATATCC	1	1304-1324	RT
MSQ10 (as)	GCCATTCGGTGTCCTGAGAG	3	1280-1300	RT
PP-AC5 (s)	TTGCTGGATCCGCGACACTCCACCAT	1+3	18-33	PCR 1/2
PP-AC8 (as)	CCGACGCTGCAGATGTACCCCATGAG	3	752-738	PCR 2
PP1-REV (as)	AGCAAGGATCCTCATACACAATACT	3	988-972	PCR 1
PP3-REV (as)	AGCAAGGATCCGCCTCGTACACAATACT	1	988-971	PCR 1
PP4 (as)	CCGGGAACTTGACGTCCT	1+3	400-417	Sequencing

Nucleotide numbering according to H77c (AF 011751) for genotype 1 and NZL1(D17763) for genotype 3 samples.
s= sense, as= antisense
RT= reverse transcription
BamHI restriction sites are underlined.

A

Patient	GT	Sample	nt-Position		
			107	204	233
SA	1	Serum	G→A		
	1	Liver	G→A		
LA	1	Serum	G→A		G→R
	1	Liver	G→A		G→R
BH	1	Serum	G→A	A→C	
	1	Liver	G→A	A→C	

B

Patient	GT	Sample	nt-Position		
			121	203	340
MO	3	Serum		G→A	
	3	Liver		G→A	
RJ	3	Serum	C→U		
	3	Liver	C→U		
ME	3	Serum			A→C
	3	Liver			A→C

Figure 6.2. Nucleotide differences between majority sequences of patients and reference sequences.

Tables A and B show changes in 5'UTR region in genotype 1 and 3 patients studied respectively. H77c (AF 011751) and NZL1 (D17763) were chosen as reference sequences for GT1 and GT3 respectively.

6.5. Sequencing of clones

All clones containing insert were sequenced in both directions by the Microarray and DNA Analysis Unit of the Sir Henry Wellcome Functional Genomics Facility (IBLS), University of Glasgow, using an ABI 377 sequencer (Applied Biosystems) based on the dideoxy method developed by Sanger (Sanger et al., 1992). In this system, each dideoxynucleotide contained a specific fluorescent dye which could be excited by a laser. Dideoxynucleotides (ddNTP) contain a hydrogen group (H) on the 3' carbon instead of a hydroxyl group (OH) which prevents the addition of further nucleotides after integration into DNA chain. The reactions were performed as linear amplification on a PCR machine by Dr. G. Riboldi-Tunnicliffe, using BigDye Terminators v 1.1 (Applied Biosystems), in a single tube containing all four ddNTPs which are each labelled with dyes which fluoresce at different wavelengths. The contents of tubes were subjected to clean up by ethanol precipitation and electrophoresis on a slab polyacrylamide gel in order to separate the different sized bands. The signals were then collected and analysed by the Sequencing Analysis 3.4.1 programme (Applied Biosystems). A coloured electropherogram was produced. In order to sequence only the 5'UTR region in the insert containing 5'UTR plus core, clones with correct orientation (5' to 3') were sequenced using M13 Reverse primer (5'CAGGAAACAGCTATGAC3') and PP4. For clones in the opposite orientation, M13 Forward (5'GTAAAACGACGGCCAG 3') and PP4 were used (Figure 6.3). A fragment (nt. 33-336) was examined from the 5'UTR of HCV including stem loop II and III and part of IV. Reference sequences H77c and NZL1 were used to align the sequences obtained for GT1 and 3 respectively.

6.6. Comparison of liver and serum quasispecies composition in patients

The populations of 5'UTR sequences in paired serum and liver samples from 6 patients (3 GT 1 and 3 GT 3) were analysed. Figures 6.4 to 6.9 summarise the results obtained. The same majority sequence was present in both the liver and serum in each patient and this

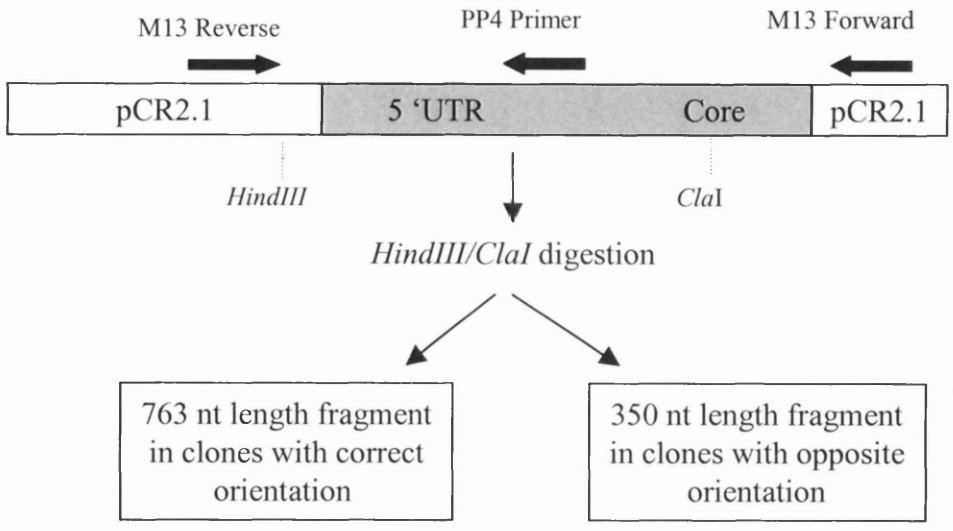


Figure 6.3. Schematic representation of pCR2.1-TOPO and insert.

The restriction sites *HindIII* cleaving at nt position 235 in the vector and *ClaI* at position 709 in the HCV core sequence are indicated. Primers used for sequencing indicated by thick arrows.

sequence was the same as that observed by direct sequencing of the PCR product. The proportion of the predominant sequence varied from 26% to 73% of clones in serum samples and 25% to 60% in liver samples and in most patients, no other clone was detected more than once. The exception to this was patient LA in whom a mixed population of G and A was observed at nt position 233 (Figure 6.4). Forty six percent of clones from serum and 35% clones from liver had G at position 233. A change from U to C was found at positions 56 and 63 in two serum clones (d and f) but clone f had two other changes which were not present in clone d. A change of A to G was present in clones *n* and *q* from the liver but clone *n* had three other changes therefore the variants were not the same. One clone (clone k) had an insertion of C at position 119/120.

In patient BH, (Figure 6.5), a change from U to C at position 37 was observed in clones d and e, however, clone d had an additional change of A to G at position 66. Clone f from the serum sample had a deletion of C at position 120.

In patient MO, (Figure 6.8), with GT 3, from 37 clones analysed from serum and liver, clone e from serum and clone *n* from liver had an insertion of C at position 117/118. Another insertion of U was found in clone q from the liver sample at position 279/280. Clones *r* and *n* had deletion of G at position 118 and 227 in comparison with the majority sequence for this patient.

6.7. Artefactual sources of variation

One difficulty with interpreting this kind of nucleotide sequence data is that both reverse transcriptase and *Taq* polymerase have relatively high error rates, and for the latter this is compounded by multiple cycles of PCR. Errors produced during RT PCR amplification of the virus genome will be present in clones and therefore be interpreted as virus heterogeneity.

Serum: 15 clones																													
nt Position		37	56	63	73	74	101	107	116	119/120	128	136	173	182	185	194	196	205	212	233	238	243	259	269	281	291	310	332
clone a (4)	26%	T	T	T	A	A	T	A	A		C	A	A	A	A	T	T	A	T	G	C	A	T	T	T	G	A	A
clone b (2)	13%																			A								
clone c (1)							C								G				C	A								G
clone d (1)			C	C																G								
clone e (1)								G												G	T							
clone f (1)			C	C												C	C			G								
clone g (1)																				A						A		
clone h (1)																		G		A								
clone j (1)													G							A								
clone k (1)										C*										A			C					
clone m (1)																				A								
Liver: 14 clones																													
clone b (4)	*28%*																			*A*								
clone a (2)	*14%*																			*G*								
clone n (1)											*T*	*G*								*G*		*G*			*C*			
clone p (1)		*C*																		*G*								
clone q (1)																				*G*		*G*						
clone r (1)																				*G*				*C*				
clone s (1)						*G*														*A*								
clone t (1)					*G*															*A*								*G*
clone v (1)																				*A*							*G*	
clone w (1)									*G*					*G*						*A*								

Figure 6.4. Nucleotide differences observed in clones derived from patient LA (GT1)

Fifteen clones were analysed from serum and 14 from liver. The number of each clone found is shown in brackets and the proportion of the quasispecies made up by each variant is shown as a percentage. The grey row shows the nucleotides present in the majority sequence. Nucleotide numbers refer to the H77c sequence. "*" denotes the nucleotide position between which the insertion occurred.

Serum: 15 clones																
nt Position		36	37	66	99	111	119	120	147	185	205	217	276	294	305	309
clone a (9)	60%	A	T	A	A	C	A	C	T	A	A	G	A	T	T	G
clone b (1)					G											
clone c (1)										G						C
clone d (1)			C	G												
clone e (1)			C													
clone f (1)								Del								
clone g (1)									C							
Liver: 14 clones																
clone a (6)	*42%*															
clone h (1)													G			
clone j (1)																
clone k (1)							C							C		
clone m (1)						A					C					
clone n (1)							G									
clone p (1)												A				
clone q (1)		T														
clone r (1)															C	

Figure 6.5. Nucleotide differences observed in clones derived from patient BH (GT1)

Fifteen clones were analysed from serum and 14 from liver. The number of each clone found is shown in brackets and the proportion of the quasispecies made up by each variant is shown as a percentage. The grey row shows the nucleotides present in the majority sequence. Nucleotide numbers refer to H77c sequence. "Del" denotes deletion of nucleotide in position indicated.

Serum: 15 clone														
nt Position		74	96	99	109	139	142	164	189	194	228	253	306	311
clone a (11)	73%	A	A	A	A	C	A	T	G	T	T	G	T	G
clone b (1)				G										
clone c (1)					G									
Clone d (1)							T							
clone e (1)								C						
Liver: 15 clone														
clone a (6)	*40%*													
clone f (1)		G												
clone g (1)			G											
clone h (1)						T								
clone I (1)									A					
clone i (1)										C				
clone k (1)											C			
clone l (1)												C		
clone m (1)													C	
clone n (1)														A

Figure 6.6. Nucleotide differences observed in clones derived from patient SA (GT1)

Fifteen clones were analysed from serum and 15 from liver. The number of each clone found is shown in brackets and the proportion of the quasispecies made up by each variant is shown as a percentage. The grey row shows the nucleotides present in the majority sequence. Nucleotide numbers refer to H77c sequence. No deletion or insertion was observed in clones from this patient.

Serum: 19 clones																								
nt Position		40	62	85	101	132	144	156	157	168	173	183	184	201	202	212	223	225	238	244	279	300	312	333
clone a (10)	52%	C	T	G	T	A	G	G	G	G	C	A	C	G	C	A	A	T	C	A	T	T	C	G
clone b (1)		T										G												
clone c (1)			C																					
clone d (1)					C																		T	
clone e (1)										A														
clone f (1)											T													
clone g (1)													T											
clone h (1)																G								
clone j (1)																	G			G				
clone k (1)																						C		A
Liver: 17 clone																								
clone a (10)	*58%*																							
clone m (1)				*A*																				
clone n (1)						*G*	*A*																	
clone p (1)								*A*																
clone q (1)									*A*						*T*									
clone r (1)														*A*										
clone s (1)																			*T*		*C*			
clone t (1)																		*C*						

Figure 6.7. Nucleotide differences observed in clones derived from patient ME (GT3)

Nineteen clones were analysed from serum and 17 from liver. The number of each clone found is shown in brackets and the proportion of the quasispecies made up by each variant is shown as a percentage. The grey row shows the nucleotides present in the majority sequence. Nucleotide numbers refer to NZL1 sequence. No deletion or insertion was observed in clones from this patient.

Serum: 22 clones																						
nt Position		46	50	61	72	78	89	117/118	118	147	158	169	183	202	203	221	227	230	279	282	296	333
clone a (13)	59%	T	G	T	A	T	T		C	T	T	G	A	C	A	A	G	C		T	A	G
clone b (1)			A																			
clone c (1)					G																	A
clone d (1)						C									G							
clone e (1)								C*														
clone f (1)											C											
clone g (1)														T								
clone h (1)												A										
clone j (1)																G						
clone k (1)																		T				
Liver: 15 clones																						
clone a (9)	*60%*																					
clone l (1)		C		C																		
clone m (1)							G													C	G	
clone n (1)								C*		C							Del					
clone p (1)													G									
clone q (1)																			T*			
clone r (1)									Del													

Figure 6.8. Nucleotide differences observed in clones derived from patient MO (GT3)

Twenty two clones were analysed from serum and 15 from liver. The number of each clone found is shown in brackets and the proportion of the quasispecies made up by each variant is shown as a percentage. The grey row shows the nucleotides present in the majority sequence. Nucleotide numbers refer to NZL1 sequence. "*" denotes the nucleotide position between which the insertion occurred. "Del" denotes the nucleotide deletion.

Serum: 27 clones																									
nt Position		44	51	52	64	65	79	88	99	101	114	132	136	147	152	160	166	181	204	231	241	247	257	269	309
clone a (18)	66%	T	A	A	A	C	A	G	T	T	A	A	C	T	A	A	C	T	A	G	G	A	T	A	G
clone b (1)			G																						
clone c (1)					G																				
clone d (1)								A						C	G										
clone e (1)										C										A					
clone f (1)											G														
clone g (1)													T												
clone h (1)																			G				C		
clone j (1)																					A				
clone k (1)																								G	
clone m (1)																									A
Liver: 14 clones																									
clone a (7)	*50%*																								
clone n (1)		*C*																							
clone p (1)				*G*			*G*																		
clone q (1)						*T*																*G*			
clone r (1)												*G*													
clone s (1)																*G*		*C*							
clone t (1)									*C*																
clone v (1)																	*T*								

Figure 6.9. Nucleotide differences observed in clones derived from patient RJ (GT3)

Eighteen clones were analysed from serum and 14 from liver. The number of each clone found is shown in brackets and the proportion of the quasispecies made up by each variant is shown as a percentage. The grey row shows the nucleotides present in the majority sequence. Nucleotide numbers refer to the NZL1 sequence. No deletion or insertion was observed in clones from this patient.

The frequency of differences from the majority sequence observed in the HCV 5' UTR target region was calculated as:

$$\frac{\text{No. of differences in each group of sequences}}{\text{No. of clones x sequence length in bp x No of PCR cycles}}$$

and expressed as the number of unique polymorphism sites per nucleotide per cycle (Smith et al., 1997b)

In total, 131 differences including 124 substitutions, 4 insertions and 3 deletions were observed amongst 202 clones sequenced. According to the manufacturer's information, the error rate of Superscript II was 1/15000 bases sequenced. Again, according to the manufacturer, the error rate of AdvanTaq Plus was 0.66 errors/1000 bp/25 cycles or 2.6 $x10^{-5}$ per nt sequenced per cycle of PCR. The expected number of substitutions in the PCR reaction was calculated as the error rate of Taq ($2.6 x10^{-5)}$ x length of fragment (303) x number of PCR cycles x number of clones sequenced. As a result, the expected number of changes would be higher for those samples (serum and liver from patient LA and liver from Patient BH) involving more PCR cycles. This was confirmed by the number of observed changes in those two patients. As shown in Table 6.2, the predicted number of errors generated by enzymes for sequences obtained after 1 round of PCR was 50 bases and that for those generated using 2 rounds of PCR including serum and liver samples from patient LA and liver sample from patient BH was 19.7 bases. Instead, 84 substitutions were observed in 125 clones involving 1 round of PCR and 40 substitutions amongst 43 clones generated by 2 rounds of PCR. In total, the expected number of changes were 69.6 nt, whereas, 124 nt changes were observed. Therefore, the number of observed substitutions obtained from cloning in this study was higher than errors predicted by RT and DNA polymerase used for amplification. In only one sample (SA, serum) the observed number of substitutions was lower than predicted.

Table 6.2. Frequency of nucleotide substitutions in the 5'UTR region derived from patient samples.

The expected number of substitutions in each sample was calculated as the sum of errors produced by reverse transcriptase and *Taq* polymerase. Errors resulting from *Taq* polymerase were calculated as the error rate of *Taq* (2.6×10^{-5}) x length of 5'UTR (303) x number of PCR cycles x number of clones.

Patient	*Tissue*	*Number of clones per sample*	*PCR cycles*	*Expected number of substitutions*	*Observed number of substitutions*	*Observed substitution rate x(10^{-5})*
MO	Serum	22	38	7	10	3.9
	Liver	15	38	4.7	7	4
RJ	Serum	27	38	8.5	14	4.5
	Liver	14	38	4.3	10	6.2
ME	Serum	19	38	6	13	5.9
	Liver	17	38	5.3	10	5.1
SA	Serum	15	38	4.7	4	2.3
	Liver	15	38	4.7	9	5.2
LA	Serum	15	58	7.1	17	6.4
	Liver	14	58	6.3	13	5.2
BH	Serum	15	38	4.7	7	4
	Liver	14	58	6.3	10	4
Total		202		69.6	124	

6.8. Characteristics of observed 5' UTR changes

Altogether, compared to the 5'UTR sequences obtained from majority sequences from each individual, 124 substitutions, 4 insertions and 3 deletions were found in 202 cloned sequences. Only 8 (6%) substitutions were transversions (changes from purine to pyrimidine or *vice versa*): A to C (2), C to A (1), G to C (2), A to T (2), T to G (1) and the reminder were transitions. These were substitutions of purine to purine: T to C (39), C to T (14), or pyrimidine to pyrimidine: A to G (45), G to A (18). Three out of 4 insertions were observed at position 119/120. The fourth insertion was at position 270/271. Three insertions and 2 deletions were observed in one patient (MO). All insertion and deletions were in regions predicted to be unpaired except one (nt 227).

6.9. Comparison of heterogeneity of sequences in serum and liver

In order to determine whether the liver samples contained more complex populations than serum samples, quasispecies complexity was calculated for all patients. Quasispecies complexity can be quantified using Pn (the ratio of the number of polymorphic sites to the number of nucleotides sequenced (Cabot et al., 2000).

To correct for differences in numbers of cycles of amplification in the calculation of Pn, the number of observed substitutions in serum and liver samples from patient LA and liver sample from patient BH which involved 58 cycles of PCR, were recalculated as:

$$\frac{\text{No. of observed substitutions} \times 38}{58}$$

Table 6.3. summarises the results obtained for the 6 patients. Clones obtained from liver samples from patients RJ and SA were more complex than clones obtained from paired serum samples. For patients MO and BH the complexity of clones obtained from serum samples was very similar to those from liver samples. In patients ME and LA, the

Table 6.3. Sequence complexity of the 5'UTR in the liver and serum quasispecies of patients.

Appearances refers to the number of times each sequence was observed in the clones analyzed. Pn: the ratio of number of polymorphic sites to number of nucleotides sequenced.

Patient	Tissue	Number of clones analyzed	Number of changes	Number of identical sequences	Appearances (no of clones)	Pn
M.O	Serum	22	10	13 (59%)	1 (13), 9 (1)	1/666
	Liver	15	7	9 (60%)	1 (9), 6 (1)	1/649
R.J.	Serum	27	14	18 (66%)	1 (17), 8 (1)	1/584
	Liver	14	10	7 (50%)	1 (7), 4 (1)	1/303
M.E	Serum	19	13	10 (52%)	1(10), 9 (1)	1/442
	Liver	17	10	10 (58%)	1(10), 7 (1)	1/515
S.A.	Serum	15	4	11 (73%)	1(11), 4 (1)	1/1136
	Liver	15	9	6 (40%)	1(6), 9 (1)	1/505
L.A.	Serum	15	11	4 (26%)	1 (4), 1(2), 9(1	1/413
	Liver	14	8.5	4 (28%)	1(4), 1(2), 8 (1	1/499
B.H	Serum	15	7	9 (60%)	1(9), 6 (1)	1/649
	Liver	14	6.5	6 (42%)	1 (6), 8 (1)	1/652

quasispecies complexity of clones derived from serum was higher than those obtained from liver samples. Figure 6.10.A shows the number of changes, which occurred per base in sequences obtained from paired serum and liver from each patient. No significant difference was found between complexities of sequences studied pooling all the data for liver and serum derived clones. ($p > 0.05$ Student's unpaired *t* test). (Figure 6.10.B)

6.10. Comparison of variation between paired and unpaired regions

All substitutions were positioned onto the predicted 5'UTR secondary structure. Seventy-five (58%) were found in paired regions whereas the total number of nucleotides in paired regions predicted by the 5'UTR secondary structure was 197 (65%). Fifty six (42%) changes were positioned in unpaired regions whereas 106 (35%) of the total nucleotides were predicted to be unpaired. (Figure 6.11 and 6.12)

The number of changes occurring in nucleotides predicted to be unpaired was higher than in paired regions. Differences in Pn values were analysed by unpaired Student's *t* test. The rate of change (Pn) in paired and unpaired region was calculated. As shown in (Figure 6.10.C) the rate of change per nt in predicted unpaired regions was 0.0028 against 0.0018 for paired regions. The difference was not significant ($p>0.05$).

6.11. Discussion

By cloning and sequencing PCR products spanning the 5'UTR, we demonstrated the existence of a quasispecies with the majority of clones corresponding exactly to the sequence obtained by sequencing the PCR product directly. Additionally, in each patient sample, there appeared to be a number of minor variants suggesting that the quasispecies nature of HCV includes variants in the 5' UTR.

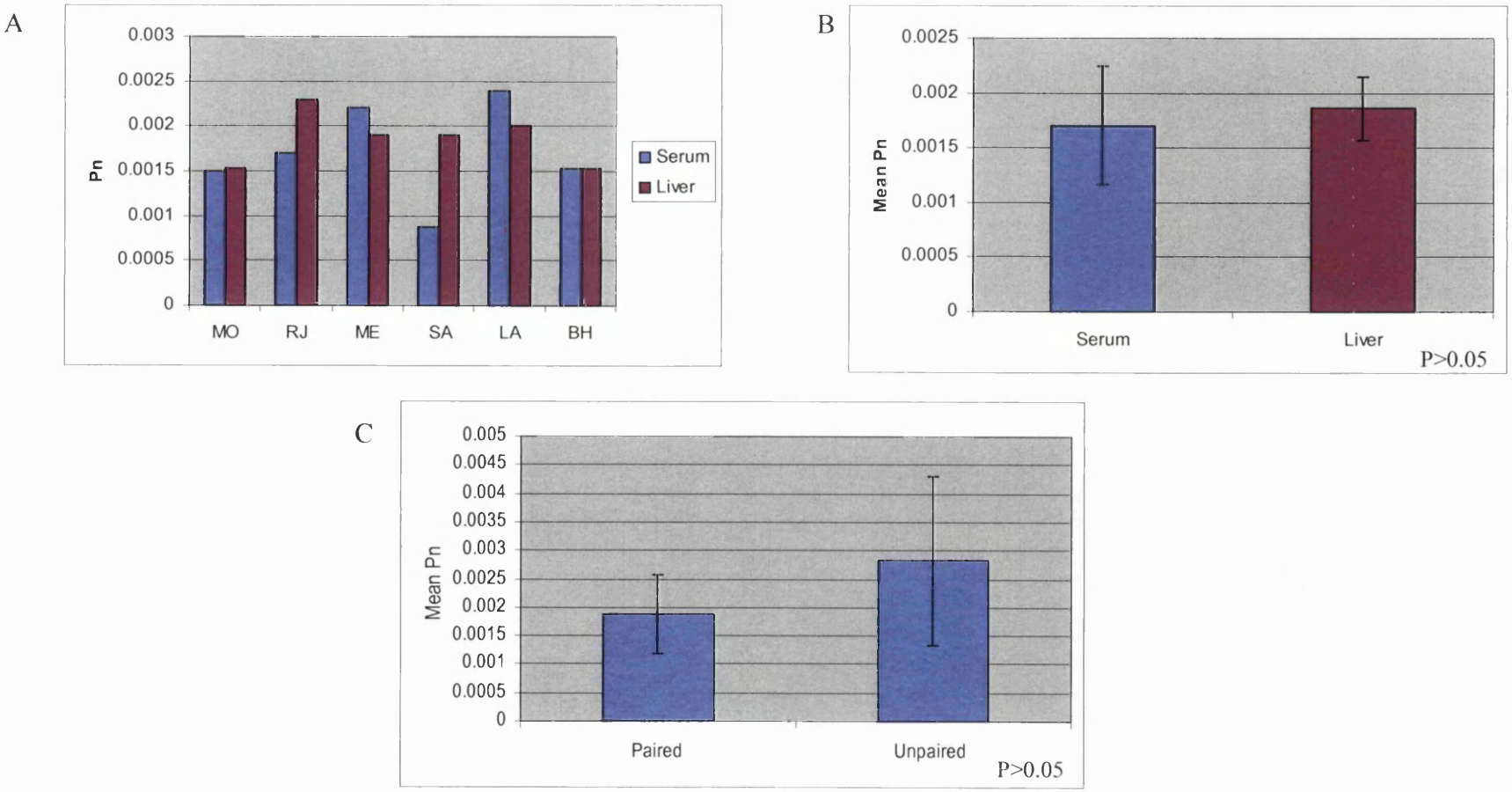


Figure 6.10. Complexity of sequences obtained from serum and liver of six patients.

Comparison of Pn between sequences obtain from serum and liver in each individual (A). Comparison of mean Pn between serum and liver from all patients is shown in (B) and comparison of Pn value in predicted paired and unpaired region of 5UTR in (C). Mean Pn values expressed as mean± standard deviation.

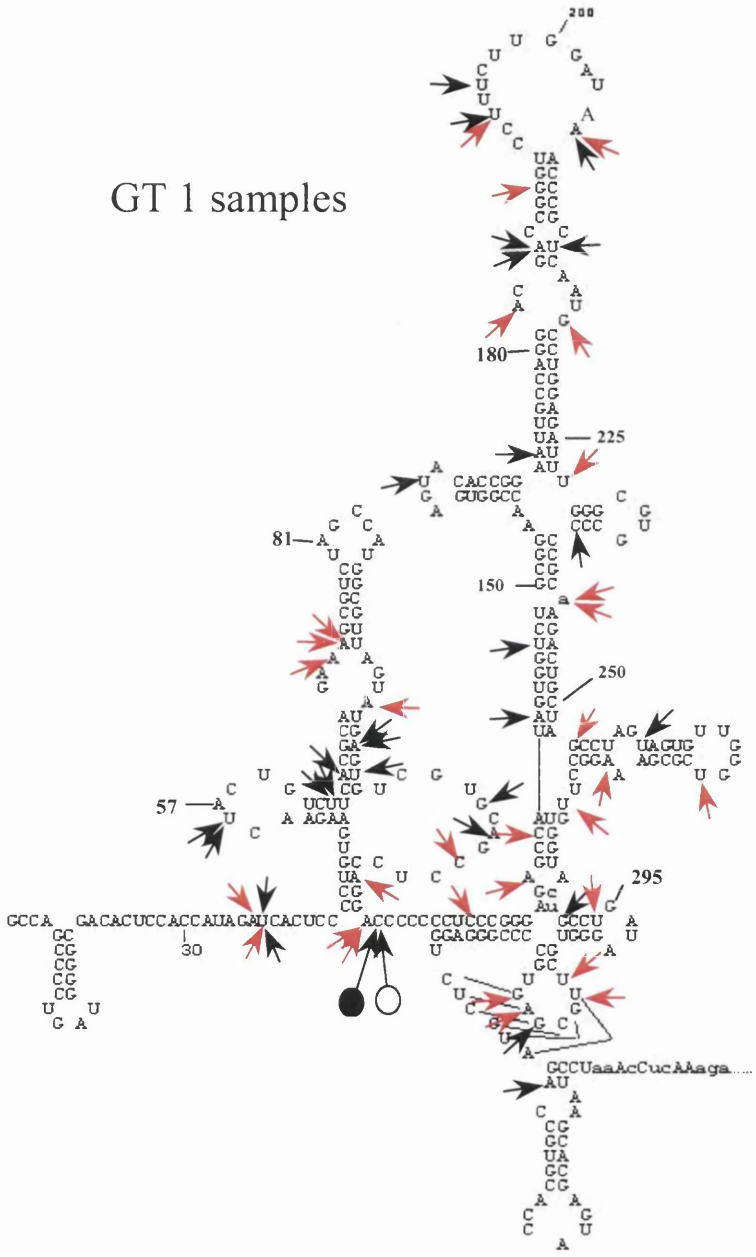


Figure 6.11. The location of nt differences in sequences studied from 3 GT 1 infected patients on the secondary structure of the 5'UTR All differences are indicated by arrows. Arrows preceeded by "●" for insertion and "❍" for deletion. Black and red arrows indicate differences from consensus observed in clones from serum and liver samples respectively. (modified from Honda *et al* 1996 by P. Preikschat)

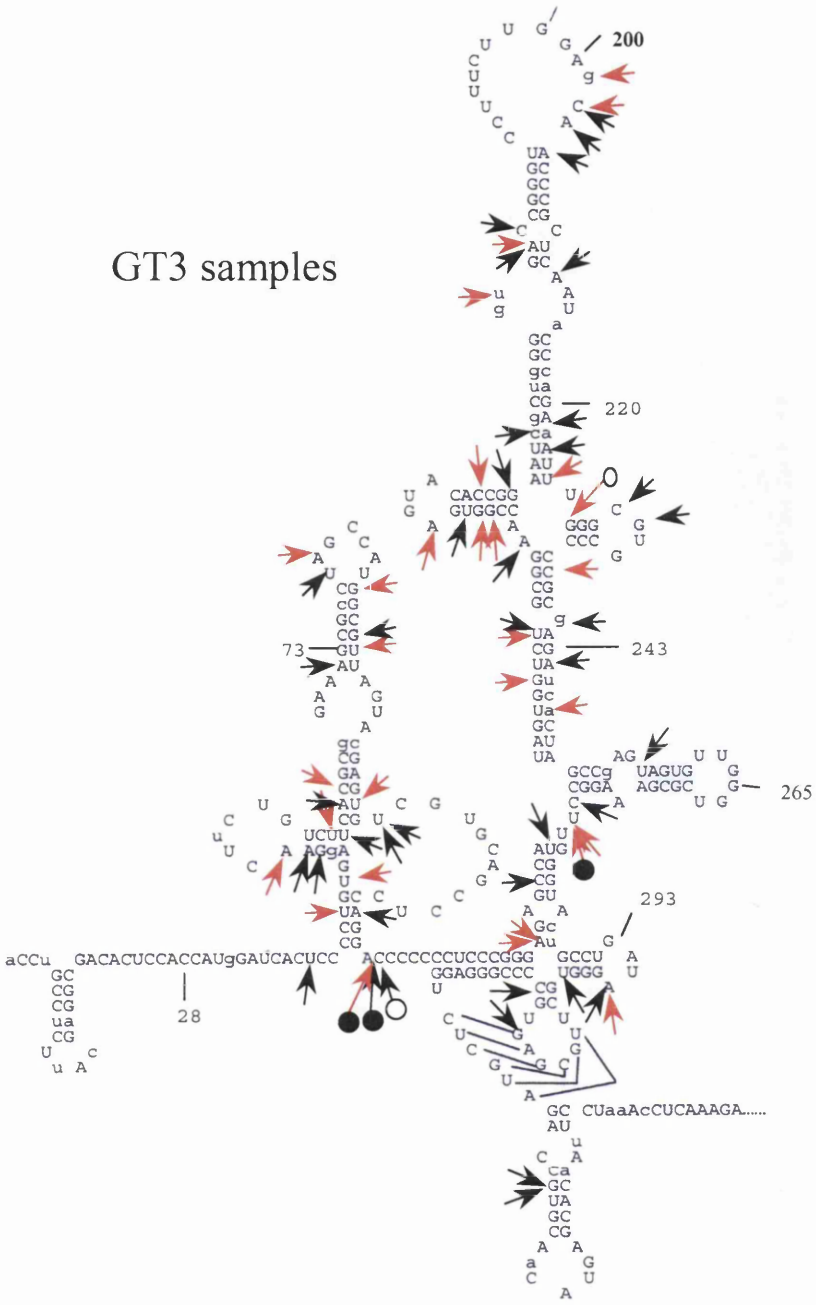


Figure 6.12. The location of nt changes in sequences studied from 3 GT 3 infected patients on predicted secondary structure of the 5'UTR.
All differences indicated by arrow. Arrows preceded by "●" for insertion and "❍"for deletion. Black and red arrows indicate changes observed in serum and liver samples respectively. (modified from Honda *et al* 1996 by P. Preikschat).

The study of HCV quasispecies composition is problematic. Sequencing of cloned PCR products is the most commonly used technique for quasispecies analysis. However, it is prone to artificial polymorphism introduced during RT-PCR. Both reverse transcriptase and *Taq* polymerase have high error rates. Therefore any errors resulting from RT PCR could be interpreted as virus heterogeneity. One approach to minimise the rate of errors introduced by DNA polymerase is to use proofreading DNA polymerases. It has been shown that the error rate for *Taq* polymerase is seven times higher than for proofreading DNA polymerases (Malet et al., 2003). In this study we used "Advan Taq Plus" polymerase which according to manufacturer's instruction has an error rate of 2.6×10^{-5} per nt sequenced per cycle of PCR.

It is possible that all variants present in quasispecies in each sample do not amplify with equal efficiency during RT PCR. One problem could be differential hybridisation of primers due to mismatches at the primer binding site. This could serve as a negative selection mechanism contributing to under-representation of variants containing those mismatching sites. Primers used during RT PCR in this study were designed from a conserved regions of the genome found by aligning as many sequences of the appropriate genotype that were available at the start of the project. At that time, there were many more complete and partial GT1a than GT3a sequences so bias may be more problematic for GT3a. I would have liked to use a different set of primers external to those used for amplification of the 5'UTR and to sequence a number of clones containing the new PCR products in order to find out whether the same variants would be detected as in the original experiment. However, this would be costly.

There is evidence which suggests at least part of the observed differences in our clones reflected real variants within the quasispecies rather than errors produced during the amplification process. Firstly, the substitution rate in the 5'UTR was found to vary between the 6 patients and between serum and liver samples within the same patient in 4 cases. Secondly, the observed substitution rates (4.5×10^{-5} for sequences obtained from 1 round and 5.3×10^{-5} for sequences obtained from 2 rounds of PCR) were higher than expected errors produced by reverse transcriptase and *Taq* polymerase (2.6×10^{-5}) used

for amplification. Thirdly, although the data presented in this chapter showed that most changes were unique to each clone, there were altered bases which were shared between more than one clone e.g. positions 56 and 63 in clones d and f in patient LA and also the mixed position at nt 233. In addition, there were some nucleotide positions which were substituted more than once among different clones sequenced (e.g. position 37, 56, 63, 74, 99 and 243 in clones obtained from GT 1 patients and positions 101, 132 and 147 in clones derived from GT 3 patients). Most artefactual substitutions would be expected to be found in only one of the multiple clones sequenced (Smith et al., 1997b). Fourthly, the rate of substitution observed in predicted unpaired regions was higher than those observed in predicted paired regions. Similar findings were reported by Soler *et al.* (2002). These suggest that substitutions in unpaired regions are tolerated more than those in paired regions. Substitutions located in paired regions change the stem loop structure of the IRES which may cause translational and replicative disadvantage. Any change in a paired region requires a covariant substitution in the complementary base to preserve the secondary structure of 5'UTR. Only one such substitution was observed, in clone c from serum sample of patient LA, which had two complementary changes of A to G at position 185 and T to C at position 212. The non-random distribution of observed changes suggests that errors introduced by DNA polymerase during the PCR reaction were unlikely to account for all the observed changes.

The data from our study support the presence of quasispecies in HCV 5'UTR which has been reported by other groups (Laporte et al., 2003; Martell et al., 1992; Soler et al., 2002). The first report of the quasispecies structure of HCV 5'UTR was by Martell *et al.* (1992). They cloned and sequenced 20 5'UTR clones from a serum sample from a patient infected with GT1 and found that 60% of the sequences were identical to the master sequence obtained from direct sequencing of PCR product and only 40% contained single, nonrepetitive base substitutions. In our study, between 26-73% of sequences obtained from serum were identical to their corresponding majority sequence. Three out of 4 insertions were observed at position 119/120 which was also observed in other 5'UTR quasispecies studies (Laporte et al., 2000; Lu et al., 2000; van Leeuwen et al., 2004).

Differences in the composition of the quasispecies between serum and liver have been reported mostly in the HVR1 region (Navas et al., 1998; Shimizu et al., 1997). Greater complexity in the liver could be due to the existence of different functional compartments in infected liver cells. It may also explained by the existence of sequences in the liver which are not able to become mature virions to be released into the circulating pool. Greater complexity in the serum could be explained by the contribution of minor variants replicating in extrahepatic sites (Laskus et al., 1998).

Only two studies were found to analyse the differences between quasispecies composition in the serum and liver in the 5'UTR (Cabot et al., 1997; Jang et al., 1999). However, the study by Jang *et al.* was carried out using SSCP analysis which will be discussed in the next chapter. Cabot *et al,* using cloning, examined the 5'UTR from three different biopsy samples from a liver explant from one patient with that from matched serum and found that the major sequence was present in a similar proportion in the 3 biopsies suggesting that quasispecies complexity in serum is not due to anatomical compartmentalization of virus replication. In their study, complexity (Pn ratio) of clones derived from serum was lower than those from liver. In our study, no significant difference was observed when the mean Pn ratio obtained from serum samples was compared with those obtained from liver from 6 patients.

The data from this study showed a higher rate of transitions (94%) than transversions (6%) in the 5'UTR fragment studied. The higher ratio of transitions in 5'UTR compared to the coding regions has been reported by other groups (Tanaka et al., 1993; Vizmanos et al., 1998). However, these studies examined the variability of 5'UTR in different isolates and not in quasispecies. The 5'UTR is not a protein coding region therefore it cannot be under immune pressure as is the case for the HVR1 region. However, the influence of other selection processes on the 5'UTR in which the fittest variant for the environment would be selected cannot be excluded.

It has been reported that the GGG triplet between nucleotides 260 and 268 of the apical loop of stem loop IIId was very conserved (Soler et al., 2002). Assuming no selection

against mutants, the expected number of mutations in this segment among clones studied would be 3.8 nts. Interestingly, this triplet was completely conserved among the clones studied here. It has been shown that this triplet is essential for IRES activity *in vitro* and *in vivo* in HCV as well in flaviviruses, pestiviruses, and GB virus B (Jubin et al., 2000) .

We have shown that a proportion of the observed substitutions in the 5'UTR derived by cloning and sequencing must represent variants present in the patient providing evidence for a quasispecies distribution in the 5'UTR. The data presented in this chapter did not support the original hypothesis that there is a subpopulation of virus in liver cells which has a lower IRES activity than the predominant HCV population circulating in the blood.

Chapter 7

SSCP analysis of quasispecies composition of 5'UTR in serum and liver

7.1. Introduction

Single-strand conformation polymorphism (SSCP) is one of the techniques widely used to identify a mutant sequence or a polymorphism in a known gene. In this technique (SSCP), the migration of single stranded DNA in non denaturing PAGE is analysed. The mobility of DNA fragment is characterised by the tertiary structure which it adopts. The PCR products are heat denatured and three dimensional structures emerge during cooling which depend on the primary nucleotide sequence. The technique allows the detection of single base changes which affect the folding of the DNA. Therefore, mixtures of DNA molecules of the same size, such as would be found when RT PCR is performed on material containing related RNA genomes as with hepatitis C, may be separated into bands of different mobility. The sensitivity of SSCP tends to decrease as fragment length increases.

The success of any particular SSCP experiment depends on the optimisation of conditions to maximise differential migration among fragments. This is of particular importance when analysing viral quasispecies as the number of expected bands cannot be determined beforehand. The original SSCP protocol used the incorporation of radioactive label and polyacrylamide gel electrophoresis on sequencing gels for detection (Orita et al., 1989) which was labour intensive and time-consuming. Simpler methods using polyacrylamide gel electrophoresis and non-radioactive staining such as silver staining have been

proposed. To improve the resolving power of SSCP, a variety of methods have been suggested, for example adding glycerol, reducing temperature, increasing the length of the gels or the duration of the gel runs (Sentinelli et al., 2000).

According to the manufacturer, MDE (Mutation Detection Enhancement Gel, Cambrex) gel solution is a polyacrylamide-like matrix that has a high sensitivity for DNA conformational differences. Since its introduction, MDE has been used in other studies including the work carried out in our laboratory in order to study variation in the NS5A gene in response to interferon treatment (McKechnie and McCruden, 2001).

This technique has been used in many studies to examine the quasispecies nature of HCV, but in most, the target region was the HVR1 (Enomoto et al., 1994; Moribe et al., 1995). The 5' UTR is the most conserved region of the HCV genome. Only a few studies have compared the quasispecies composition between matched liver and serum samples from patients infected with HCV. Laskus *et al* (1998), using SSCP analysis, reported the presence of different 5'UTR variants in peripheral blood mononuclear cells (PBMCs) from those found in serum in 5 HCV infected patients co-infected with HIV-1 suggesting the presence of extrahepatic HCV replication in PBMCs. However, no difference in quasispecies makeup was observed between liver and serum samples. The work carried out by Jang *et al.* (1999) examined the differences between the quasispecies composition of 5'UTR sequences in matched serum and liver samples obtained from 6 chronically HCV infected patients using SSCP. They reported that liver samples from 4 patients contained variants which were not found in serum samples from the same individuals. SSCP analysis of PCR product amplified from serum and liver revealed consistent major band patterns, indicating the presence of an identical master sequence in four of the six patients studied. However, in each of these patients, additional bands specific to liver derived RNA were identified. In the remaining 2 patients, however, differences in master sequence were observed and it was concluded that the observed inconsistency was due to lower viral loads in those patients.

The work presented in this chapter was undertaken to analyse the composition of the HCV quasispecies in the 5'UTR region using SSCP to answer the following questions:
1) Is there a difference in the composition of the 5'UTR quasispecies detectable between serum and liver from individual patients?
2) Are the results of quasispecies analysis by cloning and sequencing comparable with SSCP regarding variation in the 5'UTR of HCV? Specifically, are any of the variants detected by cloning present in sufficient quantity to be visible by SSCP?

.

7.2. Optimisation of SSCP

In order to determine the reliability and reproducibility of this technique, primary experiments were carried out using clones generated by V. McKechnie in our laboratory. The clones were obtained by blunt end ligation of the PCR product from the HVR1 region from a patient infected with HCV GT 1b into the vector pUC119 (McKechnie and McCruden, 2001).

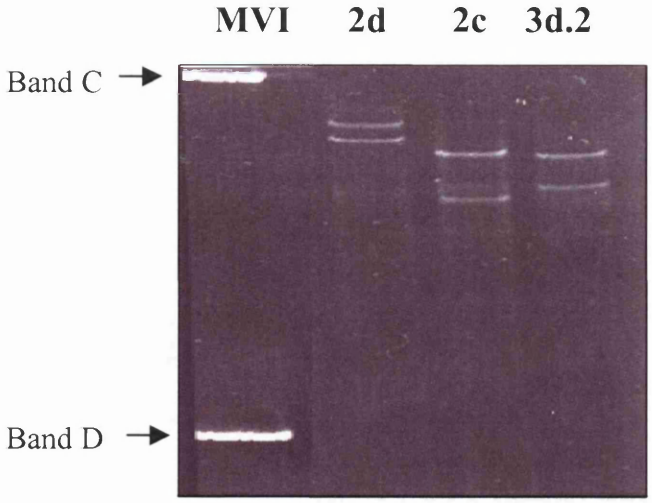
Plasmids 2c, 2d and 3d.2 containing the HVR1 inserts were amplified with primers EN3 and EN4 (Table 7.1.A) using the PCR conditions shown in Table 7.1.B resulting in the amplification of a 352 nt fragment. The numbers of nt differences between clones are shown in Figure 7.1.A. The PCR products were analysed by SSCP as described in section 2.2.11.3 to determine the number of variants with detectable differences in mobility. As expected, amplification using sense and antisense primers resulted in 2 separate bands consistent with each strand of DNA adopting a different conformation (Figure 7.1.B).

Figure 7.1. SSCP analysis of HVR clones.

A) The number of nucleotide differences between clones 2d, 2c and 2d, 3d.2 and 2c, 3d.2 are 33, 33 and 4 respectively. B) Individual clones 2d, 2c and 3d.2 were analysed by SSCP. Detectable mobility differences were seen for the three clones. Gel stained with ethidium bromide and viewed under UV light. MVI, molecular weight marker (Boehringer Mannheim) was loaded alongside the individual clones to show the relative mobility.
As DNA does not migrate according to size in SSCP analysis, the bands of MVI were assigned letters only, not sizes (C,D)

A

	3d.2	2c	2d
3d.2	.....	4	33
2c	4	.....	33
2d	33	33	.....



B

Table 7.1.A. Oligonucleotides primers for PCR of HVR1 from clones.

Primer	*Sequence 5' to 3'*	*Position**
EN3 sense	GCTTGGGATATGATGATGAACTGGTC	1296 to 1321
EN4 antisense	GGTGTGGAGGGAGTCATTGCAGTT	1623 to 1646

* Nucleotide positions are numbered according to the sequence HCV-1 (Choo et al., 1991). This resulted in the amplification of a 352 nt fragment.

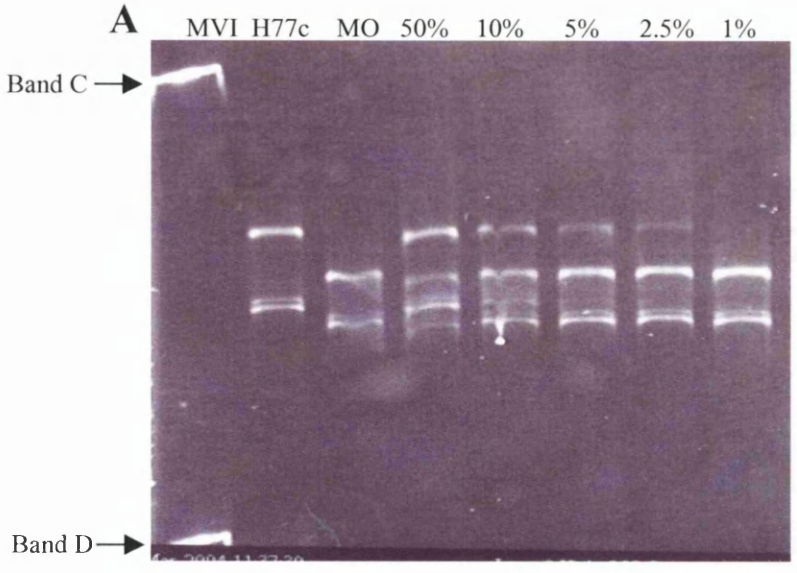
Table 7.1.B. Amplification conditions used for HVR1 region PCR.

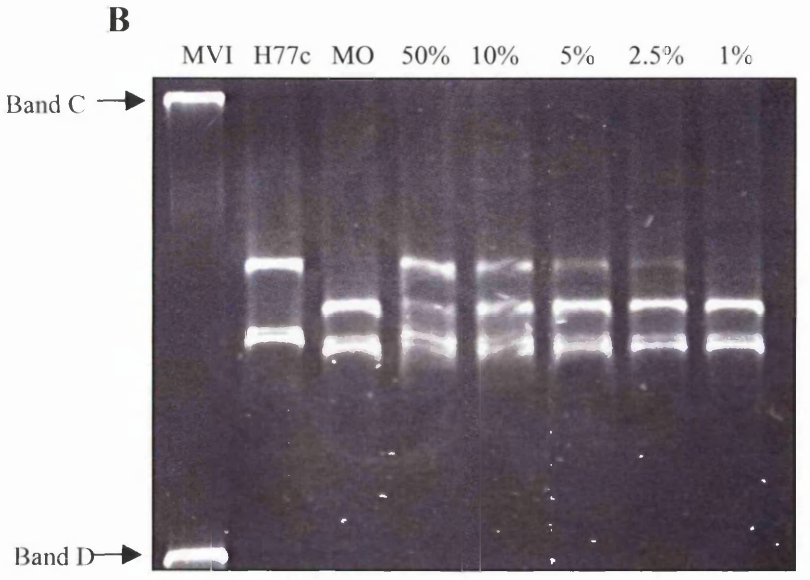
Primer	*Number of cycles*	*PCR Conditions*
EN3/EN4	25	94°C/ 1min 60°C/ 1min 72°C/ 1min

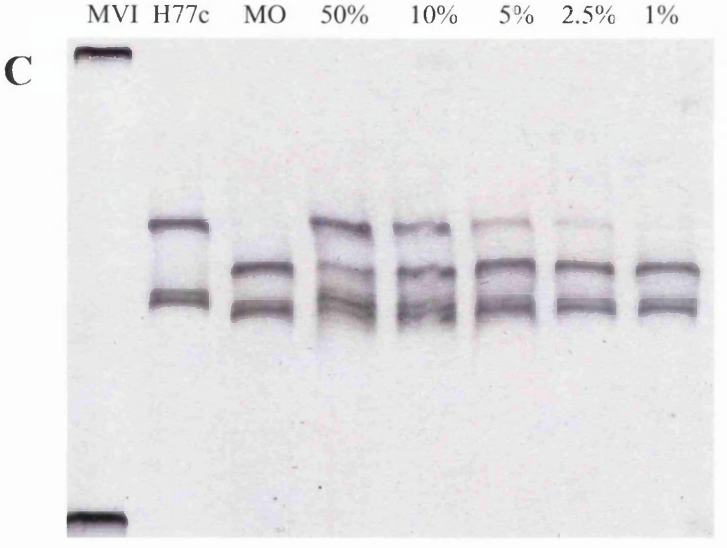
7.2.1. Sensitivity of SSCP when two variants present.

In order to determine the limit of detection of rare subpopulations, an experiment described by Enomoto *et al.* (1994) was performed with some modification. Two plasmids, pRLN MO and pRLN H77c, containing different 5'UTR sequences with different SSCP mobilities were mixed at various ratios from 500pg:500pg down to 990 pg:10pg per µl. Double stranded DNA was amplified by PCR using PPAC5 and NCR4 primers and amplified products were gel purified and subjected to SSCP analysis. This experiment was carried out twice. In the first experiment, as shown in Figure 7.2, the upper band from H77c (minor variant) in both experiments disappeared as expected as the concentration decreased. However, both H77c and MO give faster migrating bands which seem to split and this made it difficult to see the disappearance of the lower band of H77c, particularly in the second experiment. The minor subpopulation (H77c) was visible at clone mixtures of 500pg:500 pg, 900pg:100pg, 950pg:50pg and 975pg: 25pg. Bands

Figure 7.2. Sensitivity of SSCP when two variants present.
Clones 220 and H77c were mixed at various proportions prior to SSCP analysis to determine sensitivity. The concentration of clone MO remained constant (1 ng) while that of clone H77c varied. The proportion of the minor variant (H77c) compared with the major variant (clone MO) is represented as a percentage. This experiment was carried out twice. A) shows ethidium bromide staining of the bands in the first experiment. In the second experiment, bands were stained with both ethidium bromide (B) and silver staining (C). The upper of the 2 bands representing clone H77c was visible when present as 50%, 10%, 5% and 2.5% of the major variant. The corresponding lower band can be seen in the first experiment but it is obscured in the second experiment due to splitting when templates were mixed. MVI, molecular weight marker (Boehringer Mannheim) bands were assigned the letters C and D.







representing clone H77c were not visible at 990pg:10pg. Therefore, the technique was able to detect a minor variant representing only 2.5% of total population.

7.2.2. Sensitivity of SSCP analysis using Silver staining

In an attempt to increase the sensitivity, silver staining was performed using the Silver Xpress® silver staining Kit (Invitrogen) as described in chapter 2.2.11.4. It was claimed that this technique was 20 times more sensitive than ethidium bromide for the staining of double-stranded DNA in polyacrylamide gels (Boulikas and Hancock, 1981). In our case, comparison of SSCP gels stained with both techniques revealed that bands were clearer using silver staining but that, in any given gel, no additional bands were detected (Figure 7.2).

7.3. Study patients

The same samples that were studied by cloning and sequencing as described in chapter 6 were studied by SSCP. Patients SA, LA and BH, were infected with GT1 and patients MO, RJ and ME with GT3.

Comparison of direct sequences obtained from GT 1 patients with reference sequence H77c (EMBL accession number AF 011751) revealed that all had a difference with G at nt 107 instead of A (G107A). Patient LA had a mix of G and A at nt 233 and patient BH had A204C. (Figure 6.2.A)

Differences in sequence in GT 3 patients were at position G203A in patient MO, C121U in patient RJ and A340C in patient ME when compared with the reference sequence NZL1 (EMBL accession number D17763). (Figure 6.2.B)

Extraction of RNA from the samples followed by PCR had been carried out by P. Preikschat in our laboratory as described in the chapter 6. Single rounds of PCR involving 38 cycles were carried out for patients SA, MO, RJ, ME and the serum sample from patient BH. For serum and liver samples from patient LA and the liver sample from patient BH, seminested PCR involving 58 cycles was carried out using primers PPAC5 and PPAC8.

7.4. Amplification of 5'UTR region for SSCP analysis

7.4.1. Control clones:

Plasmids, containing the majority sequence of the 5'UTR determined by direct sequencing of PCR products obtained by amplifying the serum from individual patients, were used as templates to generate control PCR product for SSCP. pRLN clones derived from samples H77c, SA, BH, LA (clone1 and 6) from GT 1 and ME, MO and RJ from GT 3 were selected as templates for amplification of the 5'UTR using PPAC5 and NCR4 primers (Table 7.2.A). The band pattern of each product amplified from the relevant clone therefore represents the major variant within each sample in the absence of other variants in the quasispecies.

7.4.2. Patients samples:

For SSCP analysis of the PCR product from patient samples, I performed PCR using PCR products made by P. Preikschat as templates using primers PPAC5 and NCR4 (Table 7.2.A and 7.2.B). This resulted in the amplification of a 306 nt fragment which covered all of changes in patient majority sequences except the difference of A340C in nt 340 of patient ME. Therefore, amplified product for SSCP from this patient had an identical nt sequence with the NZL1 sequence over the region studied. PCR products were purified by gel electrophoresis and subjected to SSCP analysis

Table 7.2.A. Oligonucleotides primers for PCR of 5'UTR.

Primer	*Sequence 5' to 3'*	*Position**
PP-AC5 sense	TTGCTGGATCCGGCGACACTCCACCAT	18 - 33
NCR4 antisense	CACTCTCGAGCACCCTATCAG	293 - 313
VtagRT (RT)	TTTTTCTTTGAGGTTTAGGA	372- 353

* Nucleotide positions are numbered according to the sequence HCV-1 (Choo et al., 1991). *BamHI* site is underlined.

Table 7.2.B. PCR conditions used for amplification of 5'UTR region from clones and patient samples.

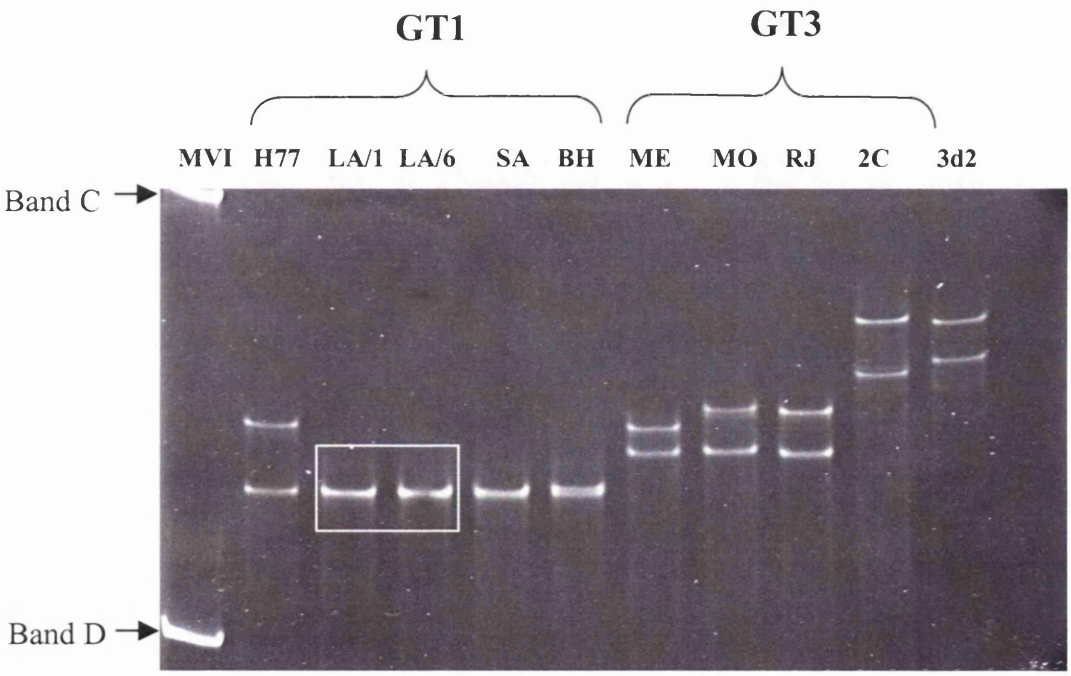
Primer	*Number of Cycles*	*PCR*
PP-AC5/NCR4	30	94°C/ 1min 60°C/ 1min 72°C/ 1min

7.5. SSCP profile of representative clones

Clones obtained from patients within one GT differ from each other by only 1-2 nt. There are 20 nt differences between the two genotypes in the amplified region. Figure 7.4 indicates that clones LA and SA sense and antisense strands had the same mobility and in clone BH, the two sense and antisense strand separated only slightly. To investigate whether SSCP detected the presence of mixed nts of A or G at nt 233 in patient LA, clones 1(A in nt 233) and 6 (G in nt 233) were examined. No differences were found between mobility of the two clones (Figure 7.3). Clone H77c containing the reference sequence for

Figure 7.3. SSCP analysis of clones matched with majority sequences of patients.

Individual clones matched with majority sequence were subjected to SSCP analysis to determine if clones could be distinguished. MVI, molecular weight marker (Boehringer Mannheim) bands have been assigned the letters C and D. Clones from the two genotypes have different migration patterns. No detectable mobility difference was found between clones LA/1 with adenosine at position 233 and LA/6 with guanosine at that position (boxed).



comparison of direct sequences obtained from GT 1 patients was included. It is notable that two bands were seen with H77c.

Detectable mobility differences were seen between clones ME, MO and RJ derived from GT 3 samples and sense and antisense strands separated.

7.6. Effect of nucleotide 107 on SSCP pattern

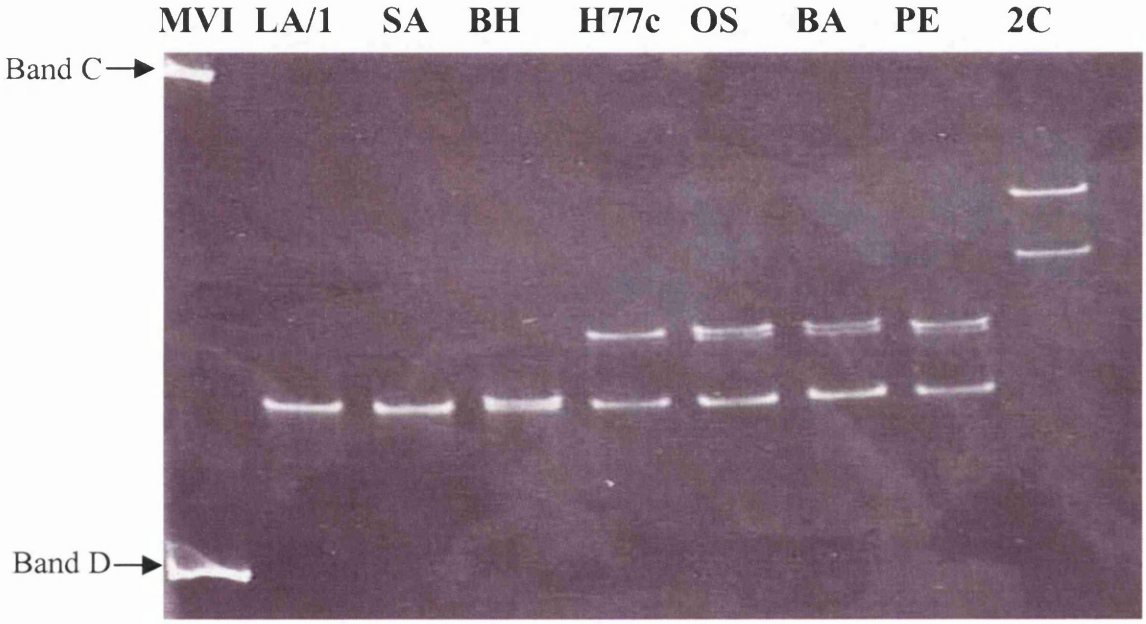
By comparing the SSCP pattern of DNAs derived from GT 1 clones, it was noticed that both sense and antisense strands of clones from 3 patients had the same relative mobility. As shown in Figure 7.3, clones LA, SA and BH have a difference at nt 107 (G to A) from the sequence in H77c.

The effect of nt change at position 107 on the separation of sense and antisense strands was investigated using three other GT 1 clones which were available. Clones OS, BA and PE with known nt A at position 107 were used. These clones differ from H77c by 1 to 3 nt but A was present in each at nt 107. PCR products amplified from these clones showed similar mobility to H77c (Figure 7.4). In other words, the two sense and antisense strands showed different mobilities. Unfortunately no GT 3 clone with guanosine at the equivalent position was available to see if both positive and negative strands would migrate at the same speed under denaturing condition.

It has already been suggested that the 3' end of minus strand of 5'UTR RNA does not fold into its mirror image (Schuster et al., 2002). In the secondary structure model of the 3' end of HCV GT 1a minus strand RNA proposed by this group, nt 107 (A) is the first nt of stem-loop SL-CI, a highly stable stem loop, which is located in a linker region between SL-BI and SL-CI (Figure 7.5). It should be noted that, with SSCP, we examine DNA not RNA secondary structure, but we were nevertheless interested in looking at this phenomenon further.

Figure 7.4. The effect of nucleotide differences at position 107 on SSCP.

To analyse the effect of nucleotide differences at position 107, seven clones with known sequences were subjected to SSCP analysis. Clones LA/1, SA and BH have nucleotide A at position 107 instead of G in H77c, OS, BA and PE. Clone BH has an additional change of A to C at position 204 which is not present in other two clones. Clone 2c was used as internal marker. Marker VI represents molecular weight marker and bands are assigned letters only and not size. Silver staining was not performed.



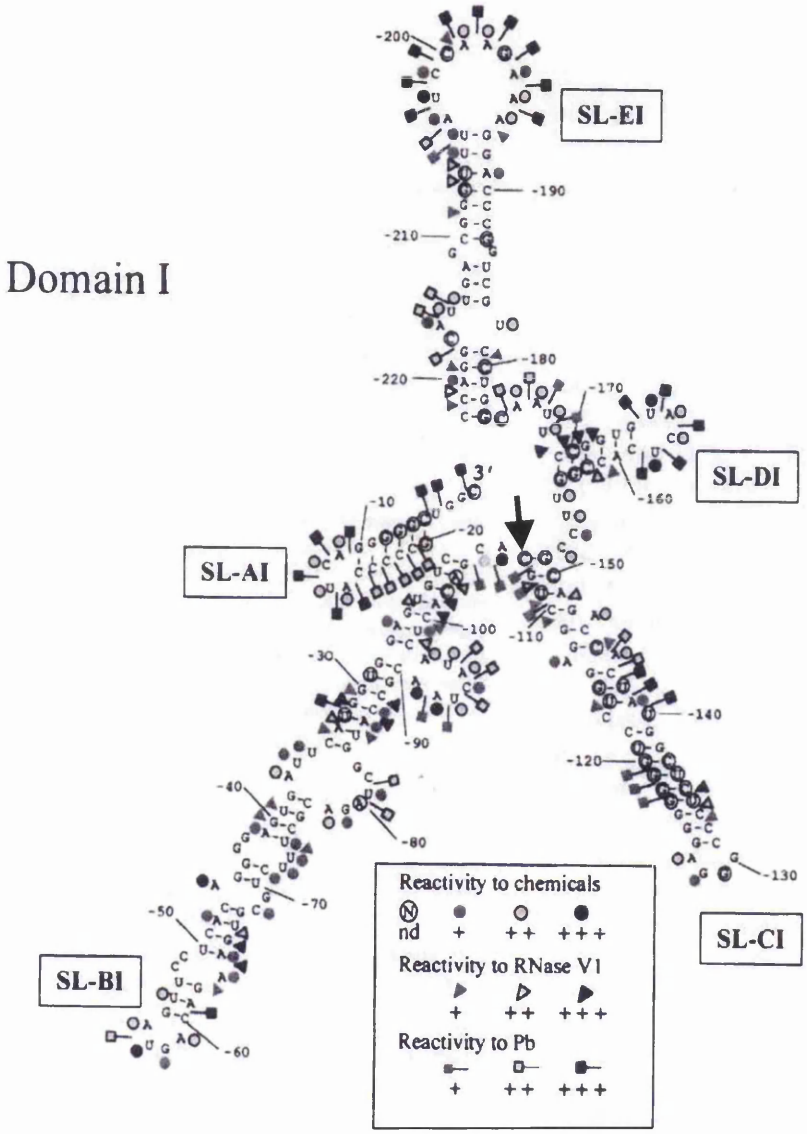


Figure 7.5. Secondary structure model of the 3' end of HCV GT 1a (-) strand RNA.

Domain I, spanning nt 230 to 1 of 5 UTR, is shown. The third stem-loop of domain I,SL-CI is highly stable. Nucleotide 107 is indicated by arrow . (Adapted from Schuster *et al* 2002).

I attempted to analyse the effect of the observed sequence difference at position 107 on the 5'UTR secondary structure of sense and antisense strands on using the Zuker MFOLD program version 3.1 (Zuker, 2003). This analysis was carried out on DNA and corresponding RNA sequences in the same nucleotide region that was examined by SSCP. In order to test minus strands, sense sequences were reversed and complemented. As can be seen in Figure 7.6. the substitution of guanosine by adenosine at position 107 in the H77c sequence resulted in a completely different predicted secondary structure from the sense strand of DNA. In contrast to the predicted secondary structure from H77c in which both sense and antisense strands were completely different, there was no difference in predicted structure in the sequence containing adenosine at nt 107. RNA analysis showed different folding pattern in antisense compared to corresponding sense strand (Figure 7.7). However, the substitution of G to A at nt 107 did not result in different secondary structures in the relevant RNA as it did with DNA. These data help explain the presence of different mobility patterns observed in sense and antisense strands of DNA samples sharing guanosine at nt 107 including H77c, OS, BA and PE but not when adenosine was present at the same location.

7.7. SSCP analysis of 5'UTR from GT 1 patients

To determine whether SSCP would reveal other variants within the quasispecies, fragments derived from the HCV 5'UTR were amplified from paired serum and liver samples from three patients (patients LA, SA and BH). SSCP analysis of PCR products amplified from clone, serum and liver revealed indistinguishable major band patterns, although the patients were known to have differences in the nt sequence of their majority variant (Figure 7.8). However, in patients LA and BH, SSCP analysis revealed the presence of an additional faint band present only in liver derived DNA and absent from clone or serum derived DNA. In all cases, SSCP bands were detected reproducibly in two separate SSCP runs using the same PCR products.

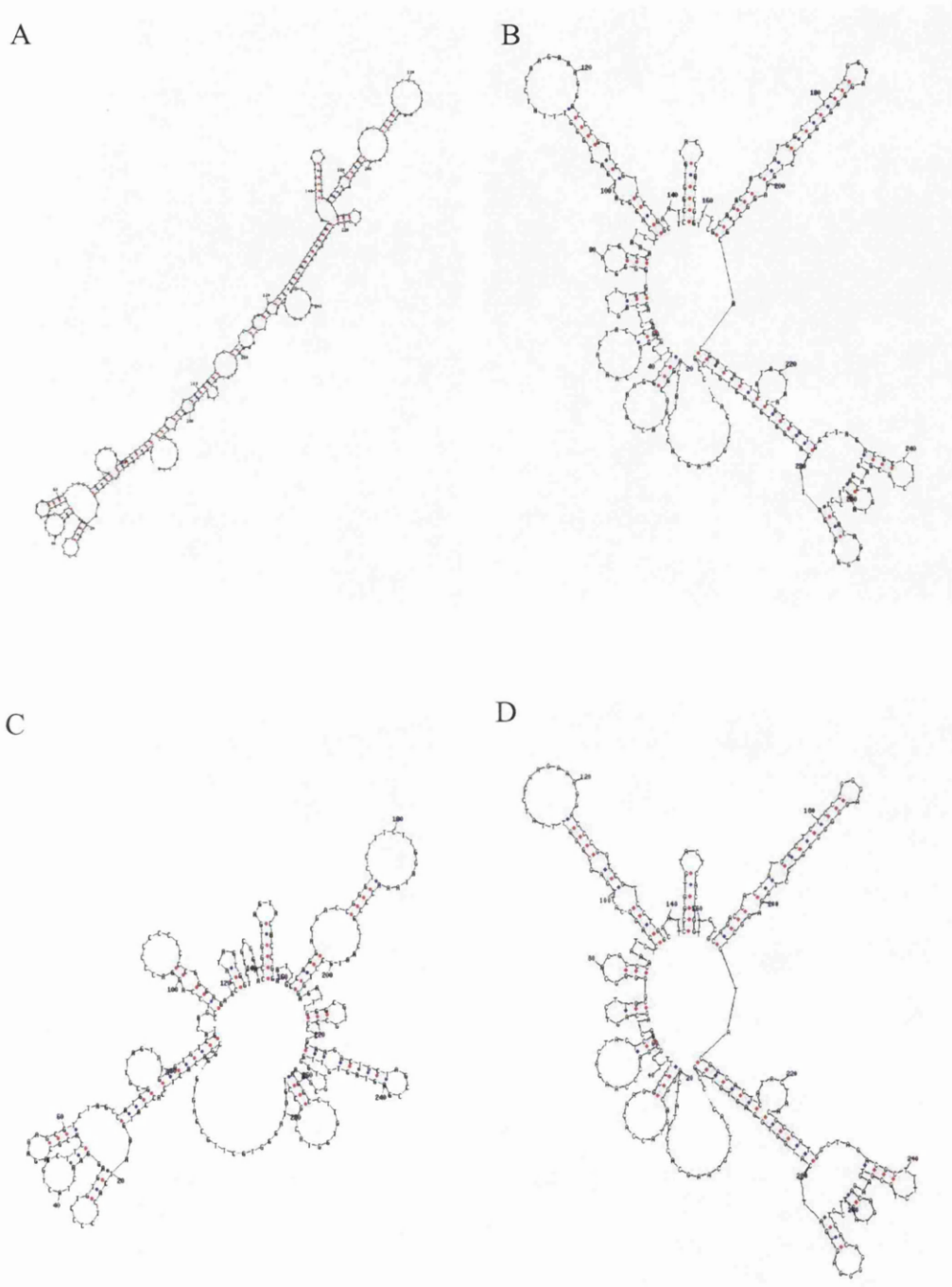


Figure 7.6 "Mfold" analysis of sense and antisense strands of the 5'UTR DNA.

The predicted secondary structure of the positive strand of H77c (A) and the relevant minus strand (B). Figures C and D show the secondary structure models of same sequence in which the nucleotide guanosine at position 107 has been substituted with adenosine.

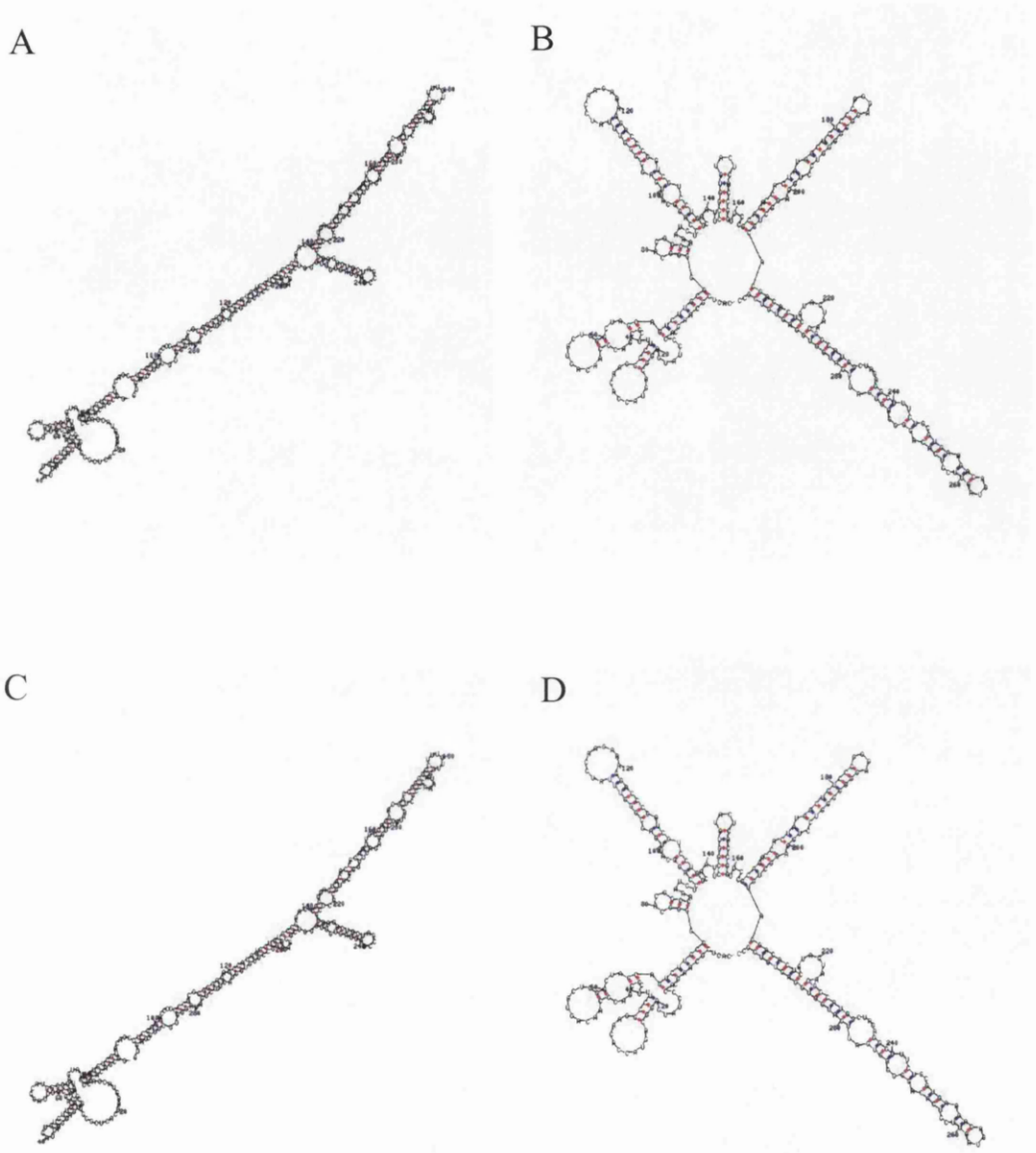


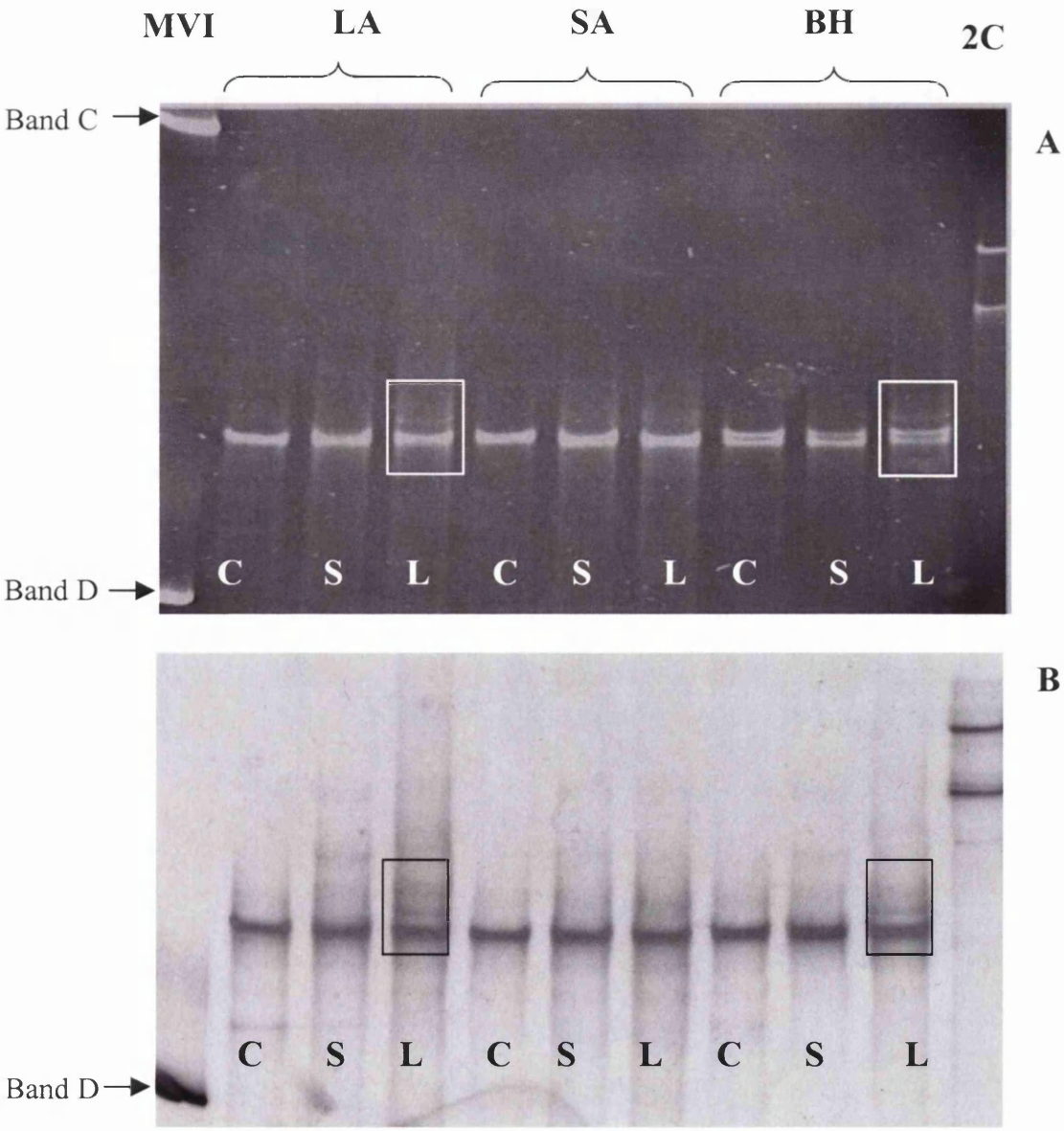
Figure 7.7 "Mfold" analysis of sense and antisense strands of the 5'UTR RNA.

The predicted secondary structure of the positive strand of H77c RNA (A) and the relevant minus strand (B). Figures C and D show the secondary structure models of the same sequence in which the nucleotide guanosine at position 107 has been substituted with adenosine.

Figure 7.8. SSCP analysis of HCV 5'UTR sequences amplified from serum and liver from GT 1 infected patients.

Lanes "C" represents sequences from clones matched with majority sequences from patients. Lanes "S" and "L" represent viral sequences from serum and liver. Two patients, LA and BH, had additional bands in liver samples shown by boxes.

Marker VI (Boehringer Mannheim) represents a molecular weight marker. The band are assigned letters (C and D) only and not size as DNA does not migrate according to size on SSCP analysis. Clone 2C was used as marker. The bands revealed by A) ethidium bromide and B) silver staining.



Slight separation of both sense and antisense strands of sample BH was observed for the clone as well as serum and liver.

7.8. SSCP analysis of 5'UTR from GT 3 patients

Amplified PCR products from serum and liver from three patients (patients ME, MO and RJ) were analysed by SSCP. The fragment derived from the clone matched with majority sequence from the respective patient in each case. As shown in the Figure 7.9, the band pattern for each patient is quite distinctive due to differing single strand DNA (ssDNA) mobilities. However, an additional band present only in liver derived samples was not found in any of these patients. In contrast to GT 1 patients, SSCP profiles revealed that sense and antisense strands had differing mobilities in all patients.

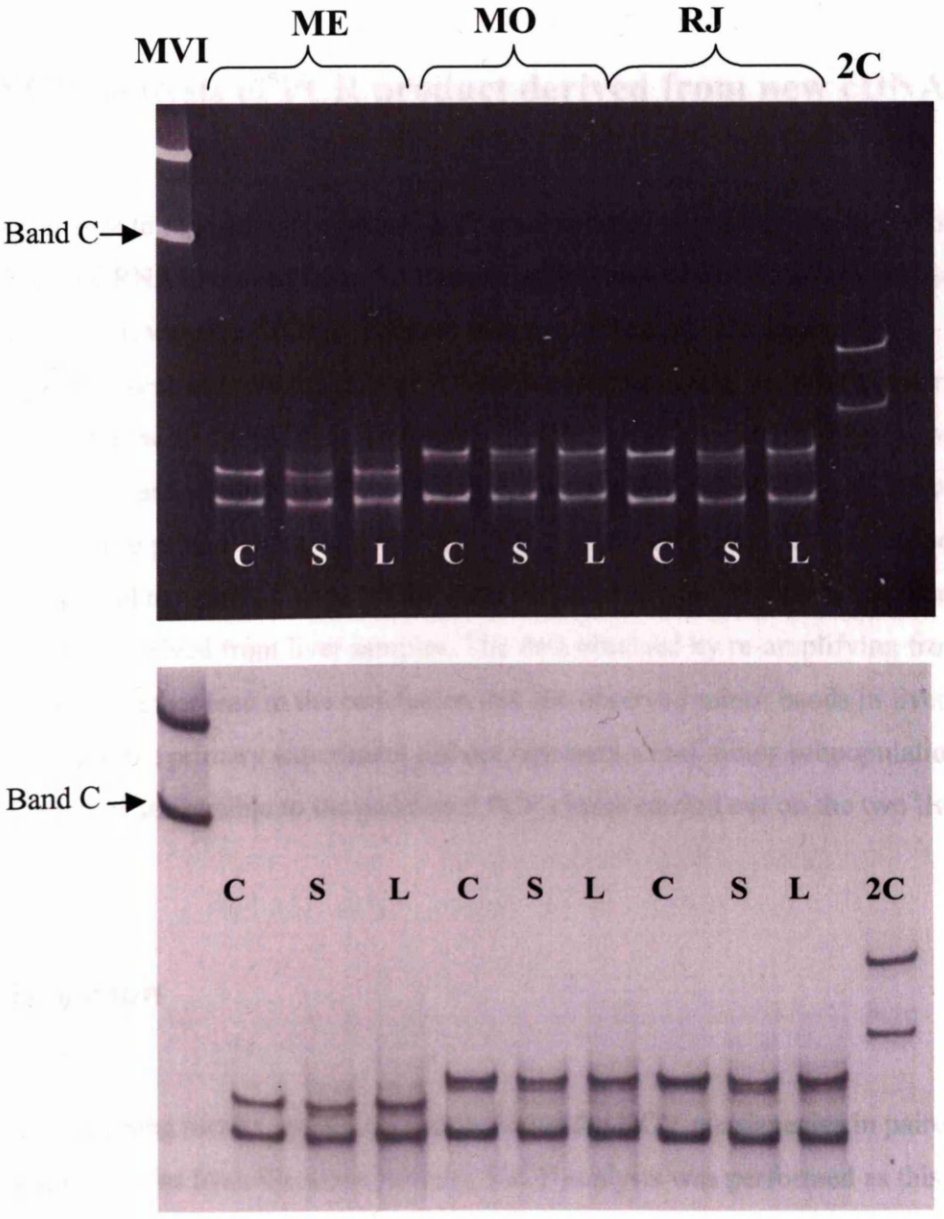
7.9. SSCP analysis of DNA extracted from silver stained polyacrylamide gel

The following steps were carried out by S. Erwa, a student in our laboratory, as part of her short term project under my supervision. To further clarify the issue of the additional band in SSCP from two liver samples (patients LA and BH), the minor band was excised from the gel shown in Figure 7.8. For comparison, major bands were excised from two corresponding serum samples (LA and BH). Extraction of DNA from the excised band was performed as described in section 2.2.11.5. Purified DNA was amplified using primers PP-AC5 and NCR4. Amplification of gel extracted DNA from the serum sample from LA was not successful. The other bands were successfully amplified by PCR and purified by gel extraction.

PCR products recovered from the gel were subjected to electrophoresis along with the relevant matched clone and the original PCR product from the liver for comparison. Silver

Figure 7.9. SSCP analysis of HCV 5'UTR sequences amplified from GT 3 infected patients.

Amplified DNA from serum and liver from three patients, ME, MO and RJ were analysed by SSCP. Lanes marked "C" represent sequences from clones matched with majority sequences from patients. Lanes "S" and "L" represent viral sequences from serum and liver. In all three patients, amplified DNA from clones, serum and liver showed identical mobility. Clone 2C were used as marker.
Marker VI represent molecular weight marker and bands are assigned letters only and not size.



staining of the gel revealed that all samples shared a major band with identical mobility with the relevant control clone DNA (Figure 7.10). None of the minor bands recovered from liver samples showed similar mobility to the additional band present in the original PCR product from liver samples. Interestingly, PCR products recovered from the major band from serum and the minor band from liver in patient BH showed migration patterns identical to each other. DNA recovered from the liver of patient LA was similar to patient BH with an extra band. No DNA recovered from the gel band from serum sample was available for comparison in patient LA.

7.10. SSCP analysis of PCR product derived from new cDNA

In a further attempt to reproduce the minor band, we attempted to amplify a new cDNA template from the RNA extracted from the liver using the original RT-PCR primers used by P. Preikschat. However, PCR amplification was not successful. Therefore, RNA extracted from two sera and two livers were reverse transcribed using Vtag-RT primer. Two rounds of PCR were carried out. Firstly, the 5'UTR region was amplified using sense primer PP-AC5 and antisense primer Vtag-RT. For the second round of PCR, NCR4 was used as an anti-sense primer along with PP-AC5. SSCP analysis (Figure 7.11) revealed a major band identical to control cloned cDNA from the same sample. No additional bands were seen in DNA derived from liver samples. The data obtained by re-amplifying from cDNA and RNA together lead to the conclusion that the observed minor bands in liver derived samples in the primary experiment did not represent a real minor subpopulation. They were possibly attributable to the additional PCR cycles carried out on the two liver samples.

7.11. Discussion

To evaluate the cloning results and assess the profile of the HCV quasispecies in paired liver and serum samples from the same patients, SSCP analysis was performed as this

Figure 7.10. SSCP analysis of DNA extracted from Silver stained gel.

The minor bands from the two liver samples (LA and BH) and major band (S*) from one serum sample (LA) were successfully extracted and amplified. Purified PCR products were loaded to SSCP and subjected to silver staining. Lanes C, L and S represent sequences obtained from matched clone, liver and serum respectively. L* and S* represent sequences recovered from silver stained gel. MVI, molecular weight marker, bands were assigned the letters C and D. Clone 2C was used as internal marker.

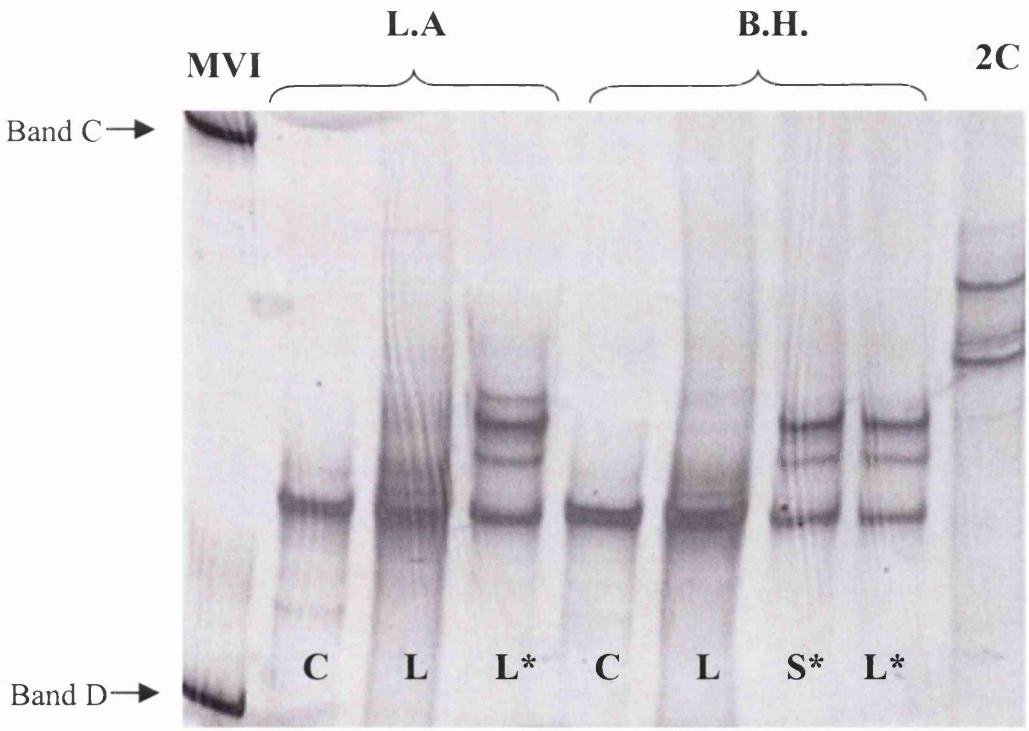
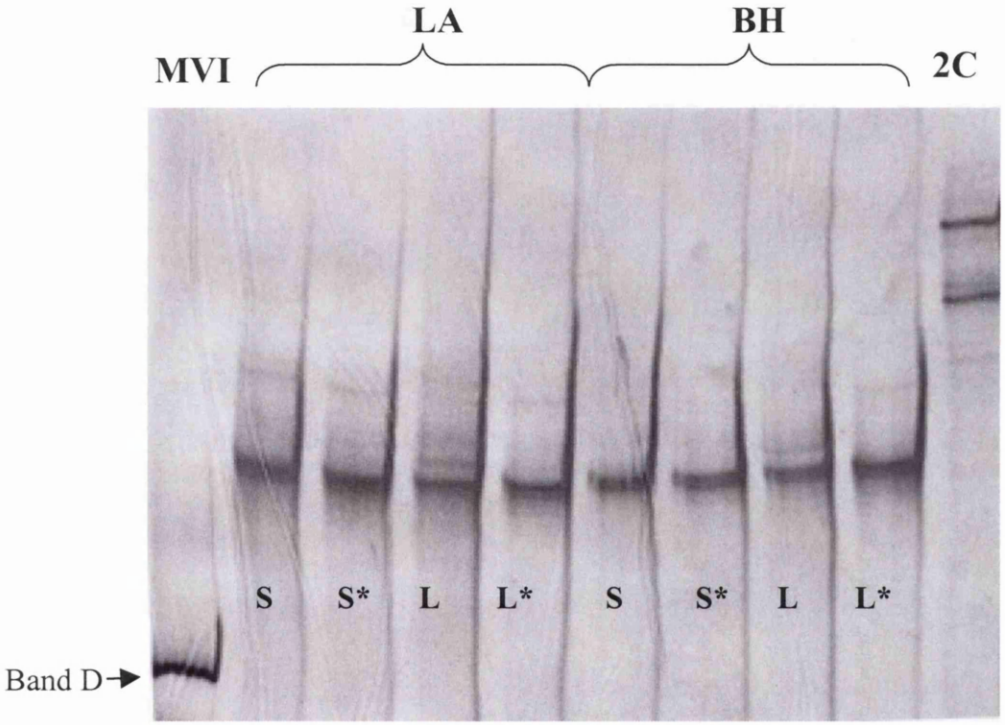


Figure 7.11. SSCP analysis of PCR product derived from new cDNA.

SSCP analysis was performed using PCR products generated from new cDNA from serum and liver samples from patients LA and BH. PCR product from new cDNA (S* and L*) were run along with previously made products (S and L). MVI, molecular weight marker, bands were assigned the letters C and D. Clone 2C was used as internal marker.



approach overcomes some of the problems associated with cloning of PCR products and the study of individual clones. For instance, it is more rapid and less prone to generate false heterogeneity as the result of misincorporation of bases by enzymes.

Using standard clones, it was shown that this technique could detect a subpopulation comprising only 2.5 % of the RNA within a simulated quasispecies of two variants. The data from the previous chapter showed the presence of single, distinct variants along with the majority of identical clones with the consensus sequence obtained from each patient. However, none of variants detected by cloning was detected by SSCP. This suggests that the real proportion of variants in the population may be overestimated using the cloning system due to the limited number of clones sequenced. In other words, none of single variants detected by sequencing could constitute 2.5% or more of the total population.

It is possible that some changes in nt sequence do not alter the secondary structure in a way that leads to a different mobility pattern in SSCP. This was seen in clones 1 and 6 from serum sample from patient LA which had only 1 difference at nt position 243. In contrast some changes appear to alter the secondary structure of DNA and resulted in a significant difference in mobility of single strand DNA (nt 107 in H77c). "Mfold" analysis showed that sense and antisense strands of H77c 5'UTR formed different secondary structures at DNA and RNA levels. However, substitution of guanosine with adenosine at position 107 resulted in a similar folding pattern for both sense and antisense strands, only in the case of DNA not RNA. These data explained the observed mobility difference between H77c and clones LA, SA and BH. It has been shown that that the 3' end of the minus strand of 5'UTR does not fold into its mirror image (Schuster et al., 2002). In the secondary structure model of the 3' end of HCV GT 1a minus strand RNA proposed by this group, nt 107 is located in a linker region between two stem loops (SL-BI and SL-CI) (Figure 7.6). Therefore, it is conceivable that any change in this nt could lead to a drastic effect on the secondary and tertiary structure of the 5'UTR with possible functional consequences. However, "Mfold" analysis on the equivalent RNA structures suggested that substitution of G107A did not change the secondary structure as it did with DNA.

At the first attempt, SSCP analysis of the serum and liver samples revealed the presence of a major band pattern, and a minor band in the liver samples from 2 of the 6 patients studied. In the remaining patients, the band pattern of serum and liver derived sequences were identical. Interestingly, the minor bands in liver samples were observed in those samples which involved three rounds of PCR. The association between the number of PCR cycles and presence of spurious bands in SSCP has been shown previously (McKechnie and McCruden, 2001). As mentioned earlier in this chapter, 3 rounds of PCR were carried out for 2 liver samples (Patients LA and BH) as well as serum sample from patient LA. In contrast to liver sample, no additional band was found in the matched serum from patient LA.

The presence of minor bands in the two liver samples was investigated further. Repeated SSCP analysis of PCR product resulting from amplification of extracted minor bands in two liver samples along with corresponding serum major bands, showed identical mobility pattern for major and minor bands extracted. To further investigate the reproducibility of minor variants observed in liver samples obtained from the two patients, SSCP analysis was repeated using new cDNA from the same extracted RNA samples. cDNA was generated using a different RT primer but still covering the entire region studied. No minor bands were observed in liver samples when reverse transcription using the different primer was performed. Therefore, it seems that the minor bands detected do not represent a real *in vivo* minor population of virus, but rather it was a part of the majority sequence with different migration due to an *in vitro* experimental factor probably *Taq* polymerase error during the additional rounds of PCR. We therefore decided not to proceed to the expense of sequencing the DNA from the minor band.

Only two studies were found which investigated the differences in quasispecies composition of 5'UTR between serum and liver using the SSCP technique (Jang et al., 1999; Laskus et al., 1998). Jang *et al* showed indistinguishable major bands compatible with the presence of identical master sequences in serum, liver and liver-derived negative strand from four out of six patients. In the remaining two patients, the results were not consistent. They also found an additional band specific to liver-derived positive strand

RNA, which were not present in either serum or liver-derived negative strand RNA in 4 of the 6 patients. In contrast, Laskus *et al.* (1998) compared the 5'UTR sequences obtained from serum and various organs including lymph nodes and liver in 5 HCV infected patients co-infected with HIV1 and found no differences in mobility pattern of bands from serum and liver samples in each patient. The results of our study revealed the presence of identical major bands in PCR products derived from paired serum and liver from each patient. However, no band specific to the liver samples was identified in line with the finding reported by Laskus *et al.*

One of the main problems with the use of SSCP in the analysis of quasispecies is that very few workers have rigorously standardised the method. Some groups did not report on the sensitivity of their assay and the amount of DNA loaded for SSCP analysis (Fujii et al., 1996; Toyoda et al., 1997; Vuillermoz et al., 2004). Loading a standard amount of DNA is important. If less DNA is loaded, it is possible that some of the minor variants in quasispecies will be missed. On the other hand, loading greater amounts of DNA would increase the thickness of bands and might mask some variants. Using MDE gel, SSCP profiles in this work were reproducible when run for 18 hours at room temperature as originally reported by V. McKechnie (2001). In addition, the presence of matched clones with the majority sequence from each patient allowed us to compare the band profile of each patient with its proper control.

In summary, comparison of the single strand conformation polymorphism patterns revealed that the number and the mobility of bands were the same between the serum and liver in the 6 patients studied. Therefore, there was no evidence to support the existence of different variants of the 5'UTR in liver from those in serum which occurred at a level of more than 2.5% of the total population. However, this conclusion must be tempered by the knowledge that some majority sequences had the same mobility as others known to differ by one nucleotide. This means that I cannot exclude the presence of minor populations differing in nucleotide sequence but having identical mobility pattern to major population.

Chapter 8

General Discussion

The first aim of this study was to compare the translation efficiencies of the IRES of HCV isolates from GT1 and GT3 infected patients. While several previous studies have examined the translation efficiency of a single sequence as a representative of one genotype, this is the first report, to our knowledge, in which the translation efficiencies of two genotypes were compared using a number of sequences from each genotype. There is some variation in the 5'UTR even between different isolates of the same genotype. One problem with using a single sequence is that it is not clear whether the observed translational activity is specific for particular sequence used or reflects real genotype specific activity. I ensured that all 5'UTR constructs used for this study were completely matched with the majority sequence obtained from each patient. This means that the observed translation activities were a true reflection of the translational activity of the majority of virions existing in each patient.

I demonstrated in BHKsinT7 cells, which are of non hepatic origin, that the mean translational activity of GT1 samples was significantly higher than that of GT3 isolates in 3 out of 5 experiments. However, in HuH7 cells, which are of hepatic origin, no significant difference was observed. It might be expected that liver derived cells would be more relevant for the study of the HCV life cycle including protein translation. HuH7 cells are permissive for replication of HCV using the subgenomic replicon system. The discrepancy between results obtained from two cell lines may be because of the differences in the way in which T7 polymerase was generated. Renilla luciferase light output levels were much higher in HuH7cells in which T7 polymerase was generated from a vaccinia construct which may have meant that translation was maximally

stimulated and became limited by factors such as the supply of amino acids for incorporation into protein.

It was shown that nucleotide differences between isolates led to different translation activities in the two cell lines studied. Some of the substitutions led to lower translation efficiency such as the substitution of cytosine to uracil at positions 121 in patient RJ.

A bicistronic construct containing renilla luciferase as the upstream reporter gene and firefly luciferase as the downstream reporter gene was used for this study. It was based on that previously constructed by Collier (1998). Importantly, both reporter enzymes are assayed in the same cell lysate preparation. The upstream translation product acts to control internally for differences in transfection efficiency. IRES activities were studied *in vivo* using two different cell lines. Many authors have examined the IRES activity in cell free systems (*in vitro*) such as in rabbit reticulocyte lysates. One complication of the cell free system is that the optimal potassium ion concentration differs for cap-dependent and IRES dependent translation (Borman et al., 1995). Translations are carried out under conditions which favour either one or the other activity, making comparisons difficult.

I would contend that studying IRES activity in cell lines is more appropriate since it is likely that the conditions are closer to those in which the HCV 5'UTR initiates the translation of the polyprotein in the infected liver. It is important, however, to bear in mind that this system is only an experimental model for studying IRES activity which may be different from that which occurs in actual infection. It is also possible that other regions in the HCV genome influence translation initiation in addition to the sequences included in the present study. For example, the effect of core protein on IRES has been reported by several groups (Kim et al., 2003; Wang et al., 2000). There are controversial reports regarding the influence of the HCV 3'UTR on translation efficiency. One study reported the enhancement of IRES activity by the 3'UTR (Ito et al., 1998), whereas, others reported downregulation (Murakami et al., 2001) or absence of regulation of IRES activity by the 3'UTR (Fang and Moyer, 2000).

Chapter 4 provides preliminary data regarding translational activity of the 5'UTR in the presence of the core encoding RNA and translated protein in GT1 and GT3 infected patients. Attempts to make the 5'UTR plus core constructs match with the majority sequence of patients were unsuccessful. Therefore, as an alternative approach, it was decided to test the translation efficiency of 3 unmatched clones from each patient blindly. In total, 6 patients were studied. It was reasoned that if there was a consistent and significant difference in translation between genotypes with such constructs, it would be evident. The mean translational activity of all GT1 constructs was compared with that obtained from three GT3 infected patients and showed no significant difference. It was shown that the inclusion of core sequence resulted in a marked reduction in firefly activity in BHKsinT7 cells which, in turn, caused a sharp decrease in firefly to renilla ratio, as indicative of IRES activity. The results from western blot analysis for expressed core and firefly proteins suggested that there was no reduction of the expressed firefly protein in comparison with constructs containing only 5'UTR. Instead it was likely that the presence of core increased the relative production of firefly protein. It seemed that the results obtained from luciferase assay and western blot were contradictory. Translation of the 5'UTR (without core) constructs would be expected to occur in cytoplasm. The presence of signal sequences, however, in the nascent core polypeptide will result in the translocation of the translational machinery into the ER. This would result in other modifications such as the glycosylation and the formation of disulfide bonds in the firefly protein. Amino acid sequence analysis of the firefly protein showed the presence of three predicted glycosylation sites. An increased apparent size of the firefly protein detected by western blot was consistent with glycosylation occurring at the predicted sites. Therefore, the observed reduction in firefly activity in the luciferase assay could be explained by reduction of firefly protein enzymatic activity rather than downregulation of the IRES activity. Unfortunately, there was no time to carry out electrophoresis following deglycosylation prior to SDS-PAGE to confirm this speculation. Taken together, our results did not support the concept that the presence of core coding sequence suppresses IRES activity. The data presented here are in contrast to previous reports suggesting that the presence of core downregulates the IRES activity (Shimoike et al., 1999; Wang et al., 2000).

Chapter 5 provides data in which postulated associations between translation efficiency and viral load or liver histology in our patient cohort were investigated. Real time quantitation of serum viral loads in patients showed no significant differences between the two genotypes. The data did not support the presence of a correlation between the IRES activity and serum viral load. There was also no significant correlation between circulating HCV RNA titers and the degree of inflammation or fibrosis in the liver. Examination of liver biopsies showed the presence of mild disease in most patients. As reported elsewhere (Adinolfi et al., 2001a; Rubbia-Brandt et al., 2004), our data demonstrated that patients infected with GT 3a had a higher prevalence of steatosis than those infected with GT1a.

The work presented in chapter 6, investigated the quasispecies makeup of 5'UTR in paired serum and liver samples. One hypothesis to explain the development of chronic infection was that the selection of variants of virus which replicated at low level occurred which would allow survival of infected cells in the presence of an immune response. Selection of mutations in a number of genes including the RNA polymerase could result in the generation of low-replication variants, but selection of variants with a less efficient IRES would result in a reduced level of protein translation which would have a direct effect on the amount of antigen produced. The resulting lower level of protein expression would reduce the possibility of the antigens being recognised by the immune system. Direct sequencing analysis of matched serum and liver samples from 26 chronically HCV infected patients carried out by P. Preikschat in our laboratory showed the presence of identical majority sequences in both tissues suggesting that the majority of hepatocytes were infected with the same variants that were present in the circulation. However, majority sequence analysis is not able to detect minor variants in the quasispecies composition of liver samples. In order to investigate this hypothesis further, the quasispecies composition of the 5'UTR in serum and relevant liver samples were examined in 6 patients using two different approaches. Firstly, a cloning and sequencing strategy was used. The majority of clones obtained from serum and matched liver samples had the identical sequence to the consensus sequence obtained from direct sequencing in each patient. It was demonstrated that a proportion of the observed substitutions in the

5'UTR derived by cloning and sequencing must represent real variants present in the patient. The data from this study provided evidence for a quasispecies distribution in the 5'UTR on which selection can occur for the fittest variants for the environment in which infection is occurring. Immune selection directly by antibody would not be an issue as this is not a protein coding region but the RNA sequence could be subjected to other selection processes such as change in secondary and tertiary structure altering the interactions with cellular proteins. No significant difference was observed in the complexity of sequences in serum and liver. The presence of quasispecies in the 5'UTR of HCV has been supported by some other studies (Laporte et al., 2000; Soler et al., 2002). Sequencing of cloned PCR products is the most commonly used method for quasispecies analysis. However, it has some limitations. One problem is that only limited number of clones can be sequenced due to the high cost of sequencing. This technique is also prone to artefactual polymorphism introduced during RT-PCR. Therefore, it was decided to use a different approach as an alternative method to examine the quasispecies makeup of the 5'UTR region.

The work presented in chapter 7 was undertaken to evaluate the quasispecies composition of the 5'UTR in serum and liver using SSCP analysis. Although SSCP has its own limitations, some of the problems of cloning and sequencing individual clones are overcome. The SSCP technique performed in this work could detect minor variants present as only 2.5% of the total population. I demonstrated that the band profiles obtained from matched serum and liver samples were identical in the 6 patients studied. This means that none of observed unique variants detected by cloning and sequencing were present in more than 2.5% of total population. Taken together, the results of sequencing and SSCP analysis of the 5'UTR, did not support the presence of different minor variants in the liver. I conclude therefore, that the hypothesis that variants in the liver with low translational activity exist which escape from immune surveillance is not consistent with the data I obtained.

8.1. Future studies

It would be necessary to repeat the experiments using the pRLN constructs in BHKsinT7 cells to remove doubts that transfection levels were adequate. To establish whether any differences observed in translation levels were really related to the cell line or to the different methods for supplying T7 RNA polymerase in BHKsinT7 and HuH7 cells, pRLN constructs could be transfected into BHK cells preinfected with vaccinia virus expressing T7 polymerase.

In order to discover whether a cryptic promoter exists in the 5'UTR DNA, it would be necessary to perform Northern blot and/or RT-PCR analysis in all cell lines used. The presence of shorter transcripts would suggest the presence of promoter in the 5'UTR DNA.

My preliminary data suggested that the presence of full-length core encoding RNA and protein did not suppress the IRES activity as reported by other studies. Instead, it was suggested that the presence of core reduced firefly enzymatic activity by translocation of the protein into the ER which resulted in post translation modifications. If time had permitted, I would have repeated the whole experiment and carried out electrophoresis following deglycosylation prior to SDS-PAGE to confirm this speculation.

The most established biological difference between GT1 and 3 is the better response of patients infected with HCV GT 3 to treatment with interferon-α (IFN-α), either alone or in combination with ribavirin. The causes and mechanisms of HCV resistance to IFN-α treatment are not understood, but several viral genomic regions or viral proteins may affect the antiviral process induced by IFN (Pawlotsky, 2003b). No work has been done in order to address the differences of the effect of IFN-α on IRES activity in GT 1 and GT 3. As discussed in chapter 3, the bicistronic constructs made for the translation study included 8 different 5'UTR isolates for each genotype. If time had permitted, a number of these constructs from each GT would have been transfected into different cell lines

followed by addition of IFN-α in serial dilutions. However, our constructs contain the T7 promoter which drives the transcription of mRNA. Using vaccinia virus for production of T7 RNA polymerase in HuH7 cells is problematic as vaccinia virus is very sensitive to IFN-α. In order to avoid this, I suggest that plasmid DNAs could be transcribed *in vitro* to RNA before transfection into cells. Another suggestion would be the replacement of the T7 promoter with one which can be processed in mammalian cells such as the cytomegalovirus (CMV) promoter. The problem with this system is that the transcribed mRNA might be subjected to post transcription modification such as splicing in the nucleus. From this study, a number of questions could be answered:
Does IFN-α have an inhibitory effect on IRES activity?
If it does, is there a difference between the inhibition of GT 1 and 3 constructs?

I would also be interested to study the IRES activity of HCV GT2 using a number of isolates obtained from patients. Previous reports suggested that GT2b contained the most efficient IRES (Collier et al., 1998; Kamoshita et al., 1997). However, these groups used only a single 5'UTR sequence as representative of the genotype activity. Using the same methods applied in this study, the IRES activity of number of GT2 isolates could be tested in different cell lines.

Appendix

Figure A.1. Alignment of the majority nucleotide seqnence of the 5'UTR and 5' end of core region (nt 1-360) obtained from GT 1a infected patients.

Sequences derived from serum and liver from each patient are indicated by S and L respectively. H77c (accession no: AF 011751) was used as reference sequence for comparison. "." represents those bases which are identical to the consensus sequence. "r" indicates a mixture of guanosine or adenosine at the relevant position. Small case letters indicate nucleotide differences from H77c sequence.

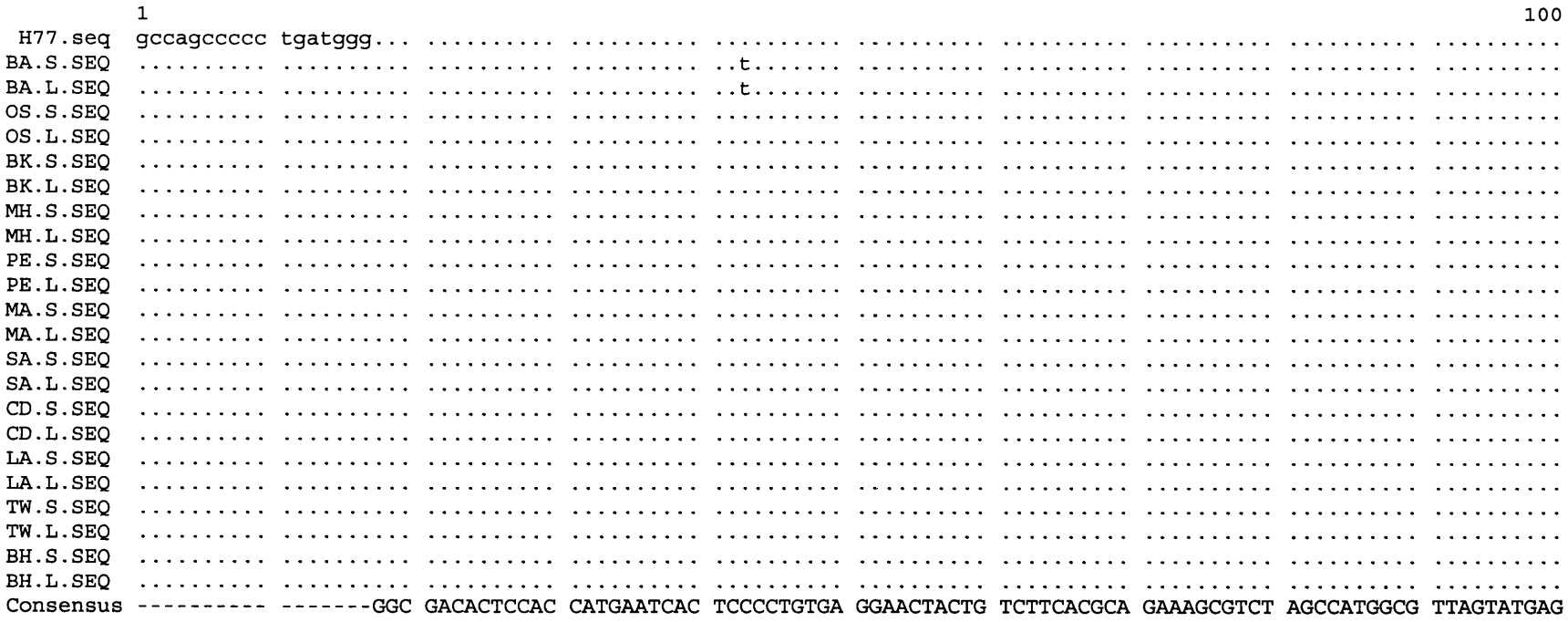


Figure A.1 Part 1

Appendix

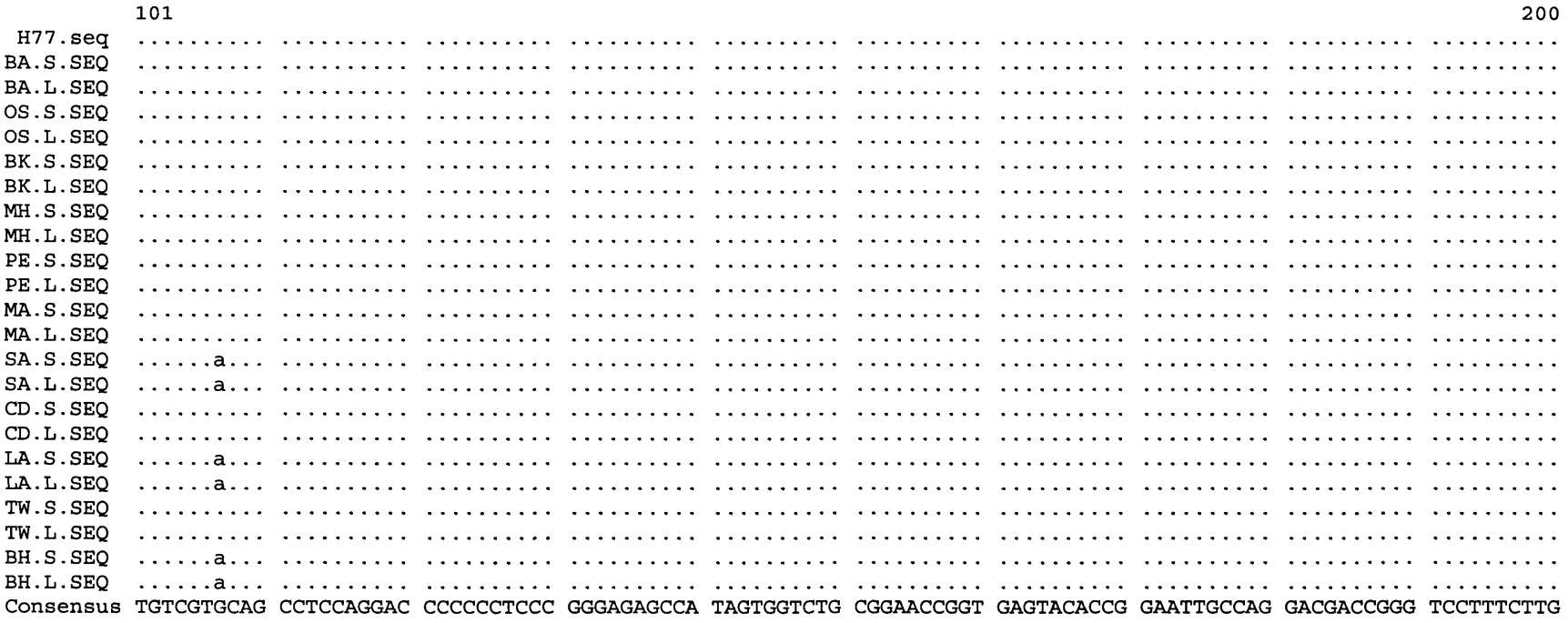


Figure A.1 Part 2

Appendix

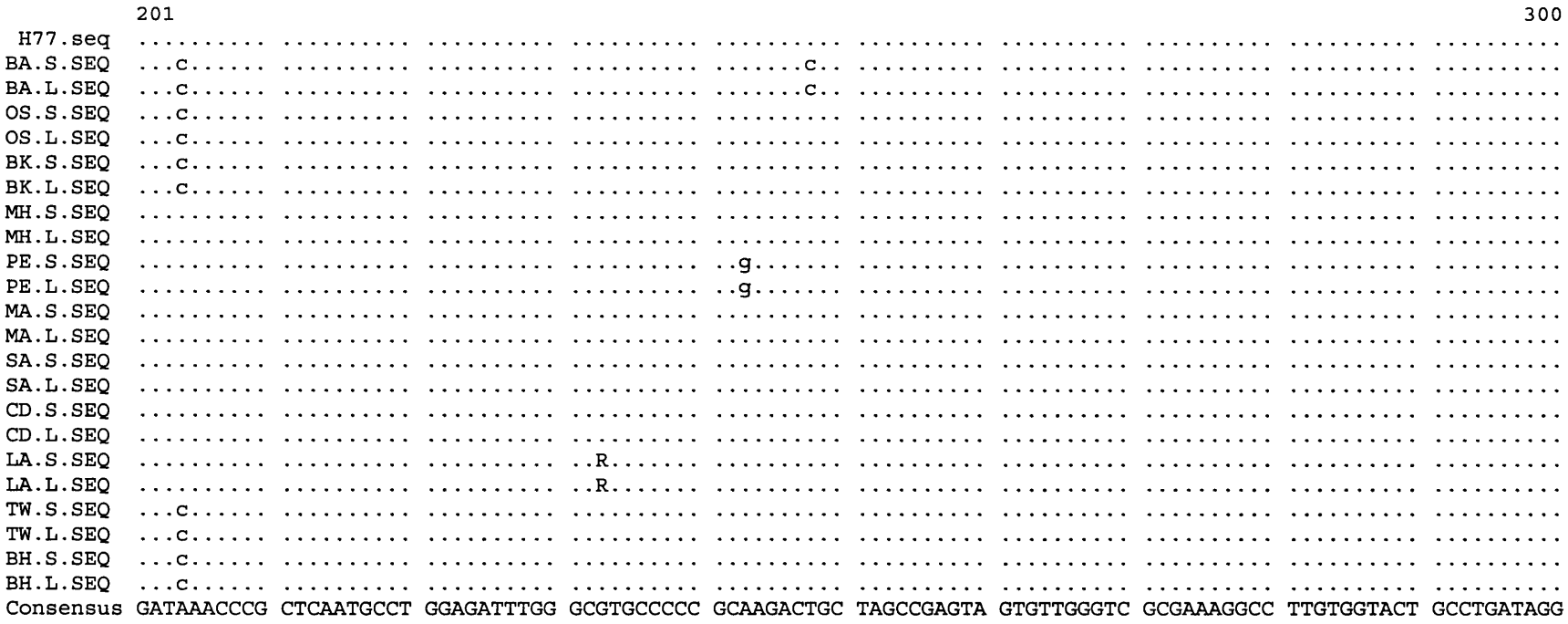


Figure A.1 Part 3

Appendix

```
          301                                                                360
 H77.seq  .......... .......... .......... .......... .......... ..........
BA.S.SEQ  .......... .......... .......... .......... .......... ..........
BA.L.SEQ  .......... .......... .......... .......... .......... ..........
OS.S.SEQ  .......... .......... .......... .......... .......... ..........
OS.L.SEQ  .......... .......... .......... .......... .......... ..........
BK.S.SEQ  .......... .......... .......... .......... .......... ..........
BK.L.SEQ  .......... .......... .......... .......... .......... ..........
MH.S.SEQ  .......... .......... .......... .......... .......... ..........
MH.L.SEQ  .......... .......... .......... .......... .......... ..........
PE.S.SEQ  .......... .......... .......... .......... .......... ..........
PE.L.SEQ  .......... .......... .......... .......... .......... ..........
MA.S.SEQ  .......... .......... .......... .......... .......... ..........
MA.L.SEQ  .......... .......... .......... .......... .......... ..........
SA.S.SEQ  .......... .......... .......... .......... .......... ..........
SA.L.SEQ  .......... .......... .......... .......... .......... ..........
CD.S.SEQ  .......... .......... .......... .........t .........a ..........
CD.L.SEQ  .......... .......... .......... .........t .........a ..........
LA.S.SEQ  .......... .......... .......... .......... .......... ..........
LA.L.SEQ  .......... .......... .......... .......... .......... ..........
TW.S.SEQ  .......... .......... .......... .......... .......... ..........
TW.L.SEQ  .......... .......... .......... .......... .......... ..........
BH.S.SEQ  .......... .......... .......... .......... .......... ..........
BH.L.SEQ  .......... .......... .......... .......... .......... ..........
Consensus GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGCAC CATGAGCACG AATCCTAAAC
```

Figure A.1 Part 4

Appendix

Figure A.2. Alignment of majority nucleotide sequence of 5'UTR and 5' end of core region (nt 1-360) obtained from GT 3a infected patients.

Sequences derived from serum and paired liver from each patient are indicated by S and L respectively. NZL1 (EMBL accession no: D17763) was used as reference sequence for comparison. . "." represents those bases which are identical to the consensus sequence. Small case letters represent bases in direct sequences obtained from patients which were different from the NZL1 sequence. "y" indicates a mixture of cytosine or uridine at the relevant position.

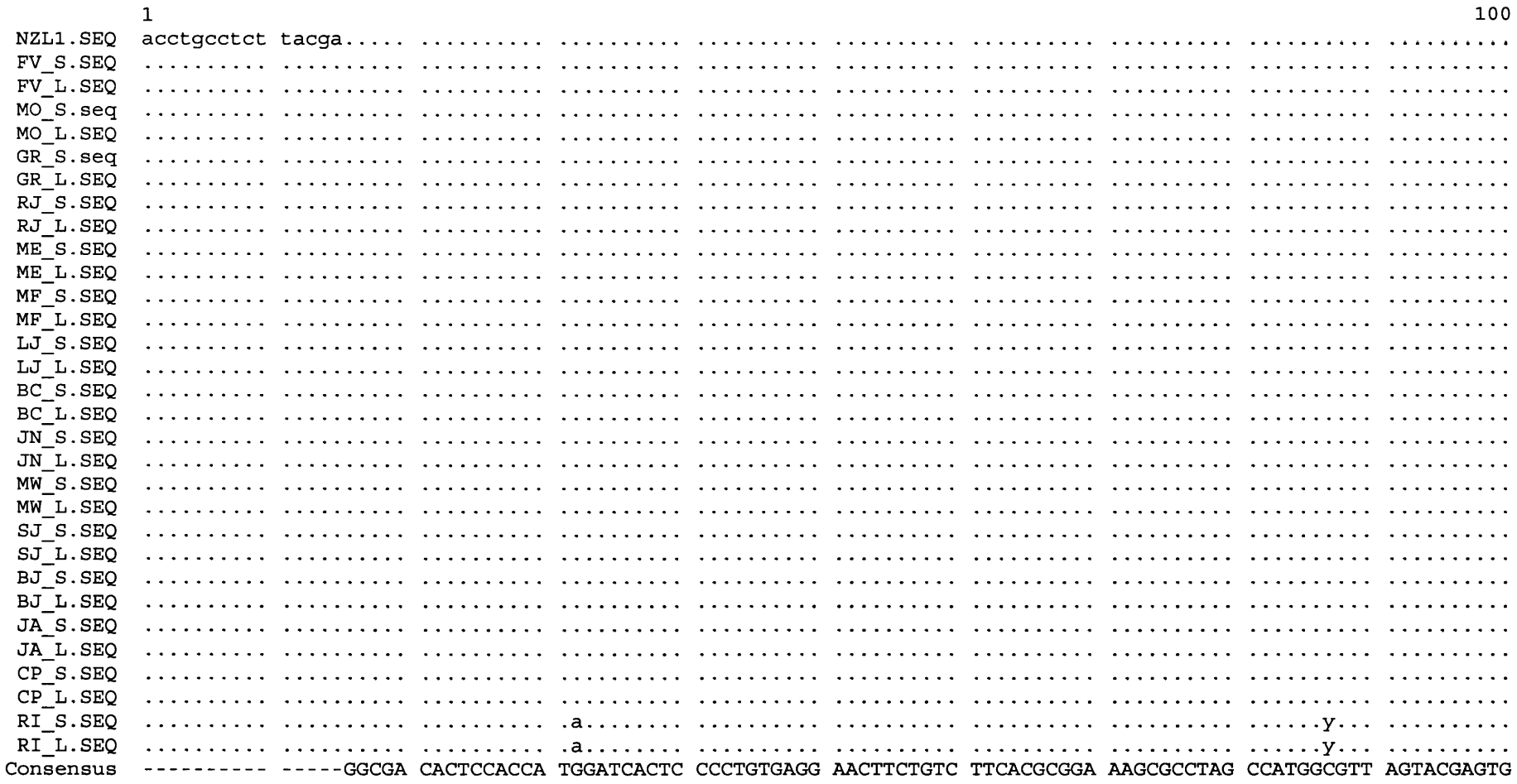


Figure A.2 Part 1

Appendix

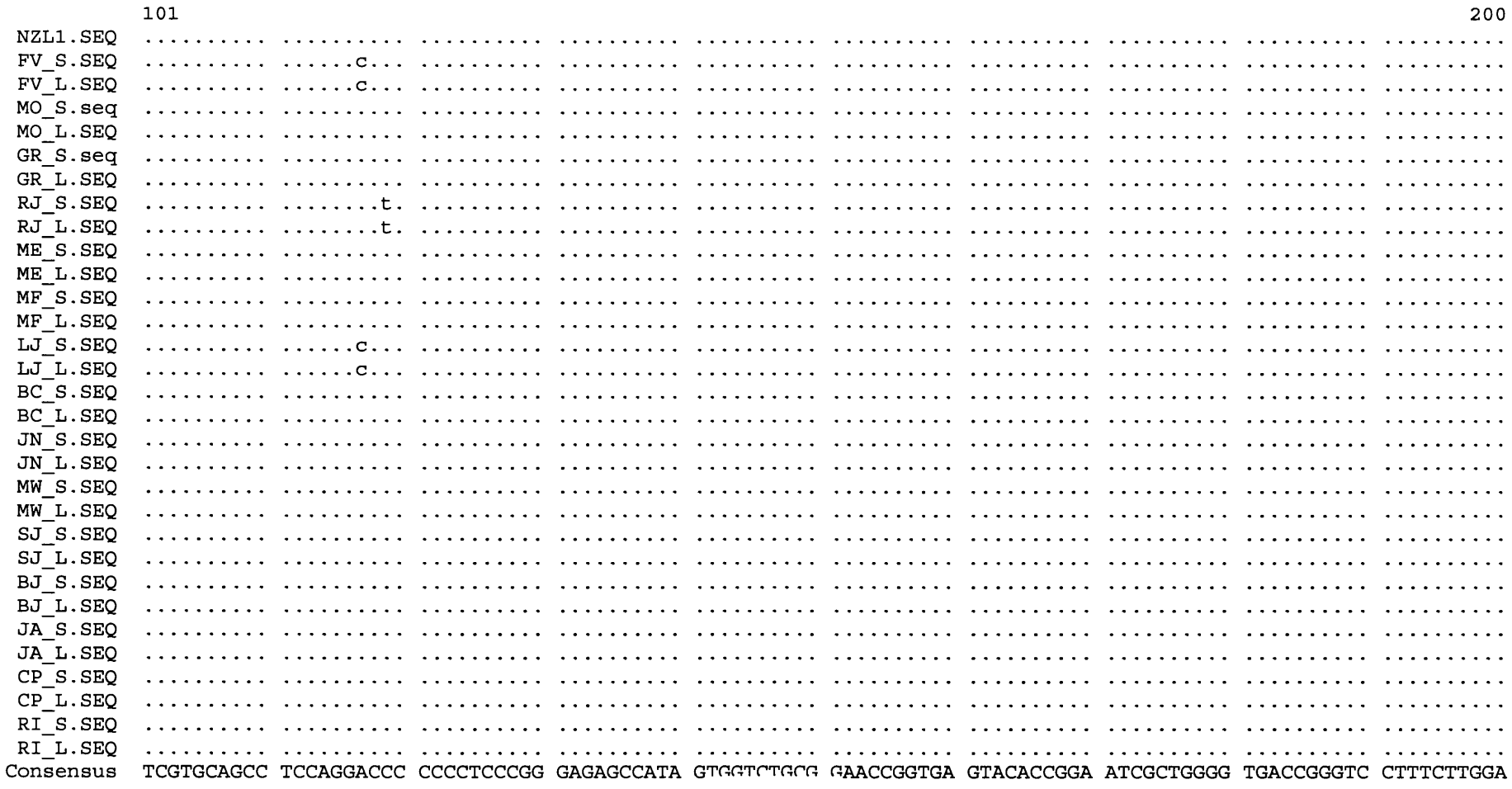


Figure A.2 Part 2

Appendix

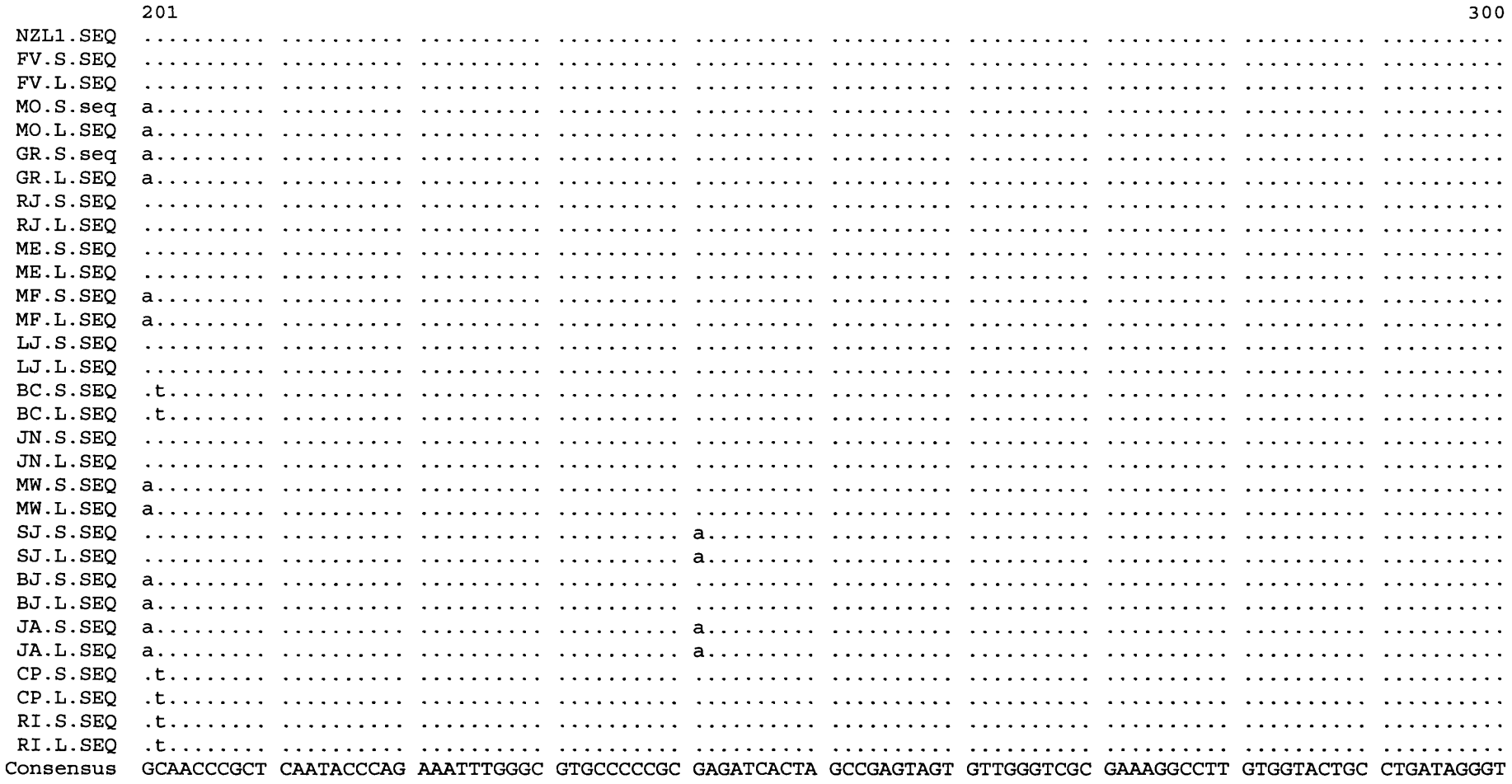


Figure A.2 Part 3

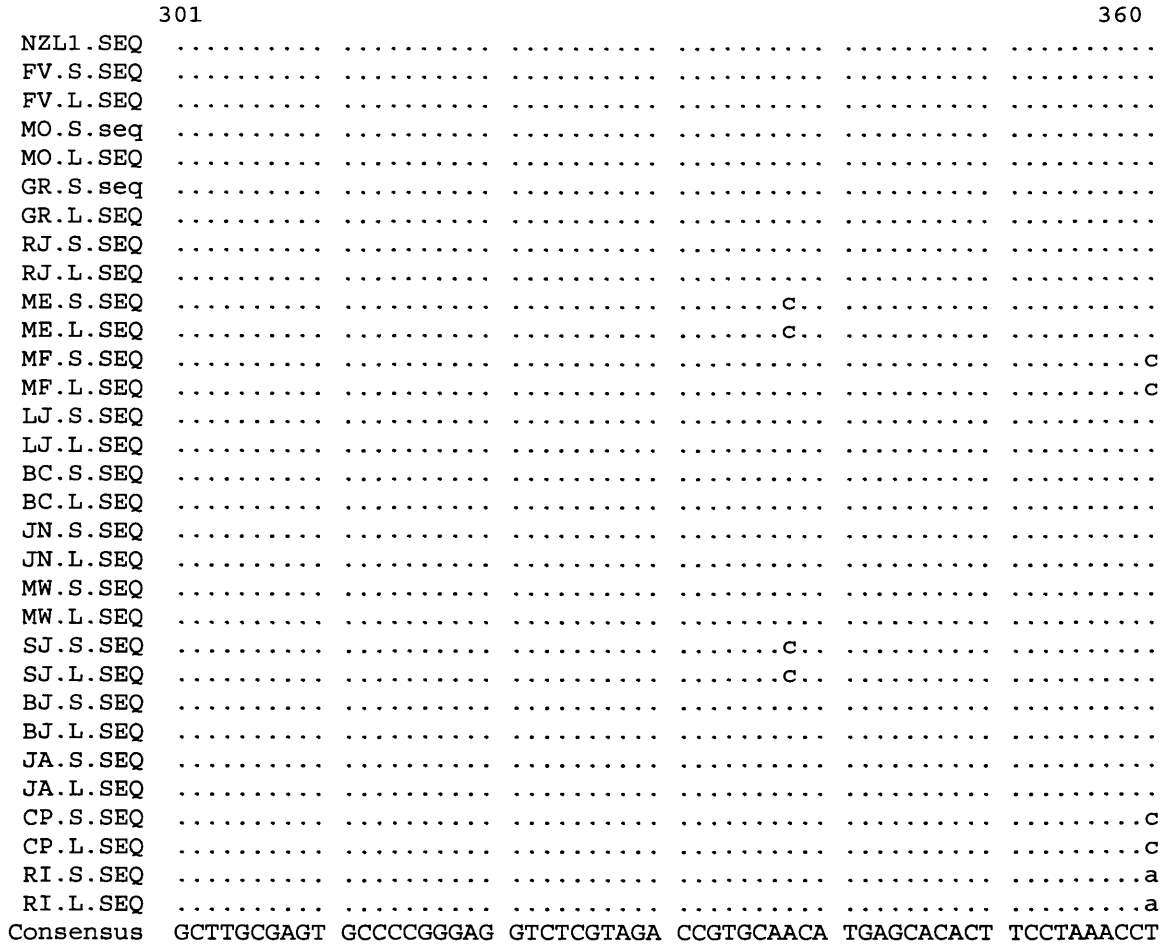


Figure A.2 Part 4

Appendix

References

Adinolfi, L.E., Gambardella, M., Andreana, A., Tripodi, M.F., Utili, R. and Ruggiero, G. (2001a) Steatosis accelerates the progression of liver damage of chronic hepatitis C patients and correlates with specific HCV genotype and visceral obesity. *Hepatology*, **33**, 1358-1364.

Adinolfi, L.E., Utili, R., Andreana, A., Tripodi, M.F., Marracino, M., Gambardella, M., Giordano, M. and Ruggiero, G. (2001b) Serum HCV RNA levels correlate with histological liver damage and concur with steatosis in progression of chronic hepatitis C. *Dig Dis Sci*, **46**, 1677-1683.

Agapov, E.V., Frolov, I., Lindenbach, B.D., Pragai, B.M., Schlesinger, S. and Rice, C.M. (1998) Noncytopathic Sindbis virus RNA vectors for heterologous gene expression. *Proc Natl Acad Sci U S A*, **95**, 12989-12994.

Agnello, V., Abel, G., Elfahal, M., Knight, G.B. and Zhang, Q.X. (1999) Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proc Natl Acad Sci U S A*, **96**, 12766-12771.

Aizaki, H., Saito, S., Ogino, T., Miyajima, N., Harada, T., Matsuura, Y., Miyamura, T. and Kohase, M. (2000) Suppression of interferon-induced antiviral activity in cells expressing hepatitis C virus proteins. *J Interferon Cytokine Res*, **20**, 1111-1120.

Ali, N., Pruijn, G.J., Kenan, D.J., Keene, J.D. and Siddiqui, A. (2000) Human La antigen is required for the hepatitis C virus internal ribosome entry site-mediated translation. *J Biol Chem*, **275**, 27531-27540.

Ali, N. and Siddiqui, A. (1995) Interaction of polypyrimidine tract-binding protein with the 5' noncoding region of the hepatitis C virus RNA genome and its functional requirement in internal initiation of translation. *J Virol*, **69**, 6367-6375.

Ali, N. and Siddiqui, A. (1997) The La antigen binds 5' noncoding region of the hepatitis C virus RNA in the context of the initiator AUG codon and stimulates internal ribosome entry site-mediated translation. *Proc Natl Acad Sci U S A*, **94**, 2249-2254.

Alter, H.J., Holland, P.V., Morrow, A.G., Purcell, R.H., Feinstone, S.M. and Moritsugu, Y. (1975) Clinical and serological analysis of transfusion-associated hepatitis. *Lancet*, **2**, 838-841.

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

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

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. and et al. (1992) The natural history of community-acquired hepatitis C in the United States. The Sentinel Counties Chronic non-A, non-B Hepatitis Study Team. *N Engl J Med*, **327**, 1899-1905.

Anand, B.S. and Velez, M. (2004) Assessment of correlation between serum titers of hepatitis c virus and severity of liver disease. *World J Gastroenterol*, **10**, 2409-2411.

Arthur, R.R., Hassan, N.F., Abdallah, M.Y., el-Sharkawy, M.S., Saad, M.D., Hackbart, B.G. and Imam, I.Z. (1997) Hepatitis C antibody prevalence in blood donors in different governorates in Egypt. *Trans R Soc Trop Med Hyg*, **91**, 271-274.

Assy, N. and Minuk, G.Y. (2000) Serum aspartate but not alanine aminotransferase levels help to predict the histological features of chronic hepatitis C viral infections in adults. *Am J Gastroenterol*, **95**, 1545-1550.

Bain, C., Fatmi, A., Zoulim, F., Zarski, J.P., Trepo, C. and Inchauspe, G. (2001) Impaired allostimulatory function of dendritic cells in chronic hepatitis C infection. *Gastroenterology*, **120**, 512-524.

Banner, B.F., Barton, A.L., Cable, E.E., Smith, L. and Bonkovsky, H.L. (1995) A detailed analysis of the Knodell score and other histologic parameters as predictors of response to interferon therapy in chronic hepatitis C. *Mod Pathol*, **8**, 232-238.

Barba, G., Harper, F., Harada, T., Kohara, M., Goulinet, S., Matsuura, Y., Eder, G., Schaff, Z., Chapman, M.J., Miyamura, T. and Brechot, C. (1997) Hepatitis C virus core protein shows a cytoplasmic localization and associates to cellular lipid storage droplets. *Proc Natl Acad Sci USA*, **94**, 1200-1205.

Bartenschlager, R., Ahlborn-Laake, L., Mous, J. and Jacobsen, H. (1994) Kinetic and structural analyses of hepatitis C virus polyprotein processing. *J Virol*, **68**, 5045-5055.

Bartosch, B., Dubuisson, J. and Cosset, F.-L. (2003) Infectious Hepatitis C Virus Pseudo-particles Containing Functional E1-E2 Envelope Protein Complexes. *J. Exp. Med.*, **197**, 633-642.

Baumert, T.F., Ito, S., Wong, D.T. and Liang, T.J. (1998) Hepatitis C virus structural proteins assemble into viruslike particles in insect cells. *J Virol*, **72**, 3827-3836.

Bedard, K.M. and Semler, B.L. (2004) Regulation of picornavirus gene expression. *Microbes and Infection*, **6**, 702-713.

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

Belsham, G.J. and Sonenberg, N. (2000) Picornavirus RNA translation: roles for cellular proteins. *Trends in Microbiology*, **8**, 330-335.

Belsham, G.J., Sonenberg, N. and Svitkin, Y.V. (1995) The role of the La autoantigen in internal initiation. *Curr Top Microbiol Immunol*, **203**, 85-98.

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

Bergqvist, A., Sundstrom, S., Dimberg, L.Y., Gylfe, E. and Masucci, M.G. (2003) The Hepatitis C Virus Core Protein Modulates T Cell Responses by Inducing Spontaneous and Altering T-cell Receptor-triggered Ca2+ Oscillations. *J. Biol. Chem.*, **278**, 18877-18883.

Bertoletti, A. and Ferrari, C. (2003) Kinetics of the immune response during HBV and HCV infection. *Hepatology*, **38**, 4-13.

Blight, K., Lesniewski, R., LaBrooy, J., Trowbridge, R. and Gowans, E. (1993) Localisation of hepatitis C virus proteins in infected liver tissue by immunofluorescence. *Gastroenterol Jpn*, **28 Suppl 5**, 55-58.

Blight, K. and Rice, C. (1997) Secondary structure determination of the conserved 98-base sequence at the 3' terminus of hepatitis C virus genome RNA. *J. Virol.*, **71**, 7345-7352.

Blight, K.J., Kolykhalov, A.A. and Rice, C.M. (2000) Efficient initiation of HCV RNA replication in cell culture. *Science*, **290**, 1972-1974.

Blight, K.J., McKeating, J.A., Marcotrigiano, J. and Rice, C.M. (2003) Efficient replication of hepatitis C virus genotype 1a RNAs in cell culture. *J Virol*, **77**, 3181-3190.

Booth, J.C., O'Grady, J. and Neuberger, J. (2001) Clinical guidelines on the management of hepatitis C. *Gut*, **49 Suppl 1**, I1-21.

Boulikas, T. and Hancock, R. (1981) A highly sensitive technique for staining DNA and RNA in polyacrylamide gels using silver. *J Biochem Biophys Methods*, **5**, 219-228.

Bradley, D.W., Krawczynski, K., Ebert, J.W., McCaustland, K.A., Choo, Q.L., Houghton, M.A. and Kuo, G. (1990) Parenterally transmitted non-A, non-B hepatitis: virus-specific antibody response patterns in hepatitis C virus-infected chimpanzees. *Gastroenterology*, **99**, 1054-1060.

Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R.L., Mathieu, M., De Francesco, R. and Rey, F.A. (1999) Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proc Natl Acad Sci USA*, **96**, 13034-13039.

Bressler, B.L., Guindi, M., Tomlinson, G. and Heathcote, J. (2003) High body mass index is an independent risk factor for nonresponse to antiviral treatment in chronic hepatitis C. *Hepatology*, **38**, 639-644.

Brewer, J.W. and Diehl, J.A. (2000) PERK mediates cell-cycle exit during the mammalian unfolded protein response. *Proc Natl Acad Sci USA*, **97**, 12625-12630.

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

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

Bukh, J., Miller, R.H. and Purcell, R.H. (1995) Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin Liver Dis*, **15**, 41-63.

Bukh, J., Purcell, R. and Miller, R. (1992) Sequence Analysis of the 5' Noncoding Region of Hepatitis C Virus. *Proc Natl Acad Sci USA*, **89**, 4942-4946.

Bukh, J., Purcell, R.H. and Miller, R.H. (1994) Sequence analysis of the core gene of 14 hepatitis C virus genotypes. *Proc Natl Acad Sci U S A*, **91**, 8239-8243.

Buratti, E., Gerotto, M., Pontisso, P., Alberti, A., Tisminetzky, S.G. and Baralle, F.E. (1997) In vivo translational efficiency of different hepatitis C virus 5'-UTRs. *FEBS Lett*, **411**, 275-280.

Cabot, B., Esteban, J.I., Martell, M., Genesca, J., Vargas, V., Esteban, R., Guardia, J. and Gomez, J. (1997) Structure of replicating hepatitis C virus (HCV) quasispecies in the liver may not be reflected by analysis of circulating HCV virions. *J Virol*, **71**, 1732-1734.

Cabot, B., Martell, M., Esteban, J.I., Piron, M., Otero, T., Esteban, R., Guardia, J. and Gomez, J. (2001) Longitudinal evaluation of the structure of replicating and circulating hepatitis C virus quasispecies in nonprogressive chronic hepatitis C patients. *J Virol*, **75**, 12005-12013.

Cabot, B., Martell, M., Esteban, J.I., Sauleda, S., Otero, T., Esteban, R., Guardia, J. and Gomez, J. (2000) Nucleotide and amino acid complexity of hepatitis C virus quasispecies in serum and liver. *J Virol*, **74**, 805-811.

Carrere-Kremer, S., Montpellier-Pala, C., Cocquerel, L., Wychowski, C., Penin, F. and Dubuisson, J. (2002) Subcellular Localization and Topology of the p7 Polypeptide of Hepatitis C Virus. *J. Virol.*, **76**, 3720-3730.

Cerino, A., Meola, A., Segagni, L., Furione, M., Marciano, S., Triyatni, M., Liang, T.J., Nicosia, A. and Mondelli, M.U. (2001) Monoclonal antibodies with broad specificity for hepatitis C virus hypervariable region 1 variants can recognize viral particles. *Journal Of Immunology (Baltimore, Md.: 1950)*, **167**, 3878-3886.

Chamberlain, R.W., Adams, N., Saeed, A.A., Simmonds, P. and Elliott, R.M. (1997a) Complete nucleotide sequence of a type 4 hepatitis C virus variant, the predominant genotype in the Middle East. *J Gen Virol*, **78**, 1341-1347.

Chamberlain, R.W., Adams, N.J., Taylor, L.A., Simmonds, P. and Elliott, R.M. (1997b) The complete coding sequence of hepatitis C virus genotype 5a, the predominant genotype in South Africa. *Biochem Biophys Res Commun*, **236**, 44-49.

Chen, C.M., You, L.R., Hwang, L.H. and Lee, Y.H. (1997) Direct interaction of hepatitis C virus core protein with the cellular lymphotoxin-beta receptor

modulates the signal pathway of the lymphotoxin-beta receptor. *J Virol*, **71**, 9417-9426.

Cheng, J. C., Chang, M. F. and Chang, S.C. (1999) Specific Interaction between the Hepatitis C Virus NS5B RNA Polymerase and the 3' End of the Viral RNA. *J. Virol.*, **73**, 7044-7049.

Cho, H.-S., Ha, N. C., Kang, L. W., Chung, K.M., Back, S.H., Jang, S.K. and Oh, B. H. (1998) Crystal Structure of RNA Helicase from Genotype 1b Hepatitis C Virus. A feasible mechanism of unwinding duplex RNA *J. Biol. Chem.*, **273**, 15045-15052.

Choo, Q. L., Kuo, G., Weiner, A.J., Overby, L.R., Bradley, D.W. and 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. and Houghton, M. (1991) Genetic organization and diversity of the hepatitis C virus. *Proc Natl Acad Sci USA*, **88**, 2451-2455.

Choo, S.H., So, H.S., Cho, J.M. and Ryu, W.S. (1995) Association of hepatitis C virus particles with immunoglobulin: a mechanism for persistent infection. *J Gen Virol*, **76 (Pt 9)**, 2337-2341.

Cione, A.L. and Torno, A.E. (1998) Assignment of the bony fish Propygidium-Primaevus (a supposed siluriform from the tertiary of Patagonia) to the order Perciformes. *Journal Of Paleontology*, **62**, 656.

Cocquerel, L., Meunier, J.C., Pillez, A., Wychowski, C. and Dubuisson, J. (1998) A retention signal necessary and sufficient for endoplasmic reticulum localization maps to the transmembrane domain of hepatitis C virus glycoprotein E2. *J Virol*, **72**, 2183-2191.

Cocquerel, L., Wychowski, C., Minner, F., Penin, F. and 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. *J Virol*, **74**, 3623-3633.

Collier, A.J., Tang, S. and Elliott, R.M. (1998) Translation efficiencies of the 5' untranslated region from representatives of the six major genotypes of hepatitis C virus using a novel bicistronic reporter assay system. *J Gen Virol*, **79**, 2359-2366.

Cooper, S., Erickson, A.L., Adams, E.J., Kansopon, J., Weiner, A.J., Chien, D.Y., Houghton, M., Parham, P. and Walker, C.M. (1999) Analysis of a successful immune response against hepatitis C virus. *Immunity*, **10**, 439-449.

Cormier, E.G., Tsamis, F., Kajumo, F., Durso, R.J., Gardner, J.P. and Dragic, T. (2004) CD81 is an entry coreceptor for hepatitis C virus. *Proc Natl Acad Sci U S A*, **101**, 7270-7274.

Cribier, B., Schmitt, C., Bingen, A., Kirn, A. and Keller, F. (1995) In vitro infection of peripheral blood mononuclear cells by hepatitis C virus. *J Gen Virol*, **76**, 2485-2491.

Davidson, F., Simmonds, P., Ferguson, J.C., Jarvis, L.M., Dow, B.C., Follett, E.A.C., Seed, C.R.G., Krusius, T., Lin, C., Medgyesi, G.A., Kiyokawa, H., Olim, G., Duraisamy, G., Cuypers, T., Saeed, A.A., Teo, D., Conradie, J., Kew, M.C., Yap, P.L. and s. (1995) Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding region. *J Gen Virol*, **76**, 1197-1204.

Di Bisceglie, A.M. (1998) Hepatitis C. *Lancet*, **351**, 351-355.

Di Bisceglie, A.M., Conjeevaram, H.S., Fried, M.W., Sallie, R., Park, Y., Yurdaydin, C., Swain, M., Kleiner, D.E., Mahaney, K. and Hoofnagle, J.H. (1995) Ribavirin as Therapy for Chronic Hepatitis C: A Randomized, Double-Blind, Placebo-Controlled Trial. *Ann Intern Med*, **123**, 897-903.

Dubuisson, J., Hsu, H.H., Cheung, R.C., Greenberg, H.B., Russell, D.G. and Rice, C.M. (1994) Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and Sindbis viruses. *J Virol*, **68**, 6147-6160.

Dumas, E., Staedel, C., Colombat, M., Reigadas, S., Chabas, S., Astier-Gin, T., Cahour, A., Litvak, S. and Ventura, M. (2003) A promoter activity is present in the DNA sequence corresponding to the hepatitis C virus 5' UTR. *Nucl. Acids. Res.*, **31**, 1275-1281.

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

Egger, D., Wolk, B., Gosert, R., Bianchi, L., Blum, H.E., Moradpour, D. and Bienz, K. (2002) Expression of Hepatitis C Virus Proteins Induces Distinct Membrane

Alterations Including a Candidate Viral Replication Complex. *J. Virol.*, **76**, 5974-5984.

Elroy-Stein, O., Fuerst, T.R. and Moss, B. (1989) Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. *Proc Natl Acad Sci U S A*, **86**, 6126-6130.

Enomoto, M., Nishiguchi, S., Shiomi, S., Tanaka, M., Fukuda, K., Ueda, T., Tamori, A., Habu, D., Takeda, T., Yano, Y. and Otani, S. (2001) Comparison of real-time quantitative polymerase chain reaction with three other assays for quantitation of hepatitis C virus. *J Gastroenterol Hepatol*, **16**, 904-909.

Enomoto, N., Kurosaki, M., Tanaka, Y., Marumo, F. and Sato, C. (1994) Fluctuation of hepatitis C virus quasispecies in persistent infection and interferon treatment revealed by single-strand conformation polymorphism analysis. *J Gen Virol*, **75 (Pt 6)**, 1361-1369.

Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Izumi, N., Marumo, F. and 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. *J Clin Invest*, **96**, 224-230.

Erickson, A.L., Kimura, Y., Igarashi, S., Eichelberger, J., Houghton, M., Sidney, J., McKinney, D., Sette, A., Hughes, A.L. and Walker, C.M. (2001) The Outcome of Hepatitis C Virus Infection Is Predicted by Escape Mutations in Epitopes Targeted by Cytotoxic T Lymphocytes. *Immunity*, **15**, 883-895.

Fang, J.W.S. and Moyer, R.W. (2000) The effects of the conserved extreme 3' end sequence of hepatitis C virus (HCV) RNA on the in vitro stabilization and translation of the HCV RNA genome. *J hepatol*, **33**, 632-639.

Fanning, L., Kenny, E., Sheehan, M., Cannon, B., Whelton, M., O'Connell, J., Collins, J.K. and Shanahan, F. (1999) Viral load and clinicopathological features of chronic hepatitis C (1b) in a homogeneous patient population. *Hepatology*, **29**, 904-907.

Farci, P., Alter, H.J., Wong, D., Miller, R.H., Shih, J.W., Jett, B. and Purcell, R.H. (1991) A long-term study of hepatitis C virus replication in non-A, non-B hepatitis. *N Engl J Med*, **325**, 98-104.

Farci, P., Alter, H.J., Wong, D.C., Miller, R.H., Govindarajan, S., Engle, R., Shapiro, M. and Purcell, R.H. (1994) Prevention of hepatitis C virus infection in chimpanzees after antibody- mediated in vitro neutralization. *Proc Natl Acad Sci USA*, **91**, 7792-7796.

Farci, P., Shimoda, A., Wong, D., Cabezon, T., De Gioannis, D., Strazzera, A., Shimizu, Y., Shapiro, M., Alter, H.J. and Purcell, R.H. (1996) Prevention of hepatitis C virus infection in chimpanzees by hyperimmune serum against the hypervariable region 1 of the envelope 2 protein. *Proc Natl Acad Sci USA*, **93**, 15394-15399.

Fargion, S. and Fracanzani, A.L. (2003) Prevalence of hepatitis C virus infection in porphyria cutanea tarda. *J Hepatol*, **39**, 635-638.

Feinstone, S.M., Kapikian, A.Z., Purcell, R.H., Alter, H.J. and Holland, P.V. (1975) Transfusion-associated hepatitis not due to viral hepatitis type A or B. *N Engl J Med*, **292**, 767-770.

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

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

Foy, E., Li, K., Wang, C., Sumpter, R., Jr, Ikeda, M., Lemon, S.M. and Gale, M., Jr. (2003) Regulation of interferon regulatory factor-3 by the hepatitis C virus serine protease. *Science*, **300**, 1145-1148.

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

Friebe, P. and Bartenschlager, R. (2002) Genetic Analysis of Sequences in the 3' Nontranslated Region of Hepatitis C Virus That Are Important for RNA Replication. *J. Virol.*, **76**, 5326-5338.

Friebe, P., Lohmann, V., Krieger, N. and Bartenschlager, R. (2001) Complete 5' noncoding region is necessary for the efficient internal initiation of hepatitis C virus RNA. *J Virol*, **75**, 12047-12057.

Fried, M.W., Shiffman, M.L., Reddy, K.R., Smith, C., Marinos, G., Goncales, F.L., Jr., Haussinger, D., Diago, M., Carosi, G., Dhumeaux, D., Craxi, A., Lin, A., Hoffman, J. and Yu, J. (2002) Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med*, **347**, 975-982.

Fuerst, T.R., Niles, E.G., Studier, F.W. and Moss, B. (1986) Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci U S A*, **83**, 8122-8126.

Fujie, H., Yotsuyanagi, H., Moriya, K., Shintani, Y., Tsutsumi, T., Takayama, T., Makuuchi, M., Matsuura, Y., Miyamura, T., Kimura, S. and Koike, K. (1999) Steatosis and intrahepatic hepatitis C virus in chronic hepatitis. *J Med Virol*, **59**, 141-145.

Fujii, K., Hino, K., Okazaki, M., Okuda, M., Kondoh, S. and Okita, K. (1996) Differences in Hypervariable Region 1 Quasispecies of Hepatitis C Virus between Human Serum and Peripheral Blood Mononuclear Cells. *Biochemical and Biophysical Research Communications*, **225**, 771-776.

Fukushi, S., Katayama, K., Kurihara, C., Ishiyama, N., Hoshino, F.B., Ando, T. and Oya, A. (1994) Complete 5' noncoding region is necessary for the efficient internal initiation of hepatitis C virus RNA. *Biochem Biophys Res Commun*, **199**, 425-432.

Fukushi, S., Kurihara, C., Ishiyama, N., Hoshino, F., Oya, A. and Katayama, K. (1997) The sequence element of the internal ribosome entry site and a 25- kilodalton cellular protein contribute to efficient internal initiation of translation of hepatitis C virus RNA. *J. Virol.*, **71**, 1662-1666.

Fukushi, S., Okada, M., Stahl, J., Kageyama, T., Hoshino, F.B. and Katayama, K. (2001) Ribosomal Protein S5 Interacts with the Internal Ribosomal Entry Site of Hepatitis C Virus. *J. Biol. Chem.*, **276**, 20824-20826.

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

Gale, M.J., Jr., Korth, M.J., Tang, N.M., Tan, S.L., Hopkins, D.A., Dever, T.E., Polyak, S.J., Gretch, D.R. and 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.

Gao, B., Hong, F. and Radaeva, S. (2004) Host factors and failure of interferon-alpha treatment in hepatitis C virus. *Hepatology*, **39**, 880-890.

Gardner, J.P., Durso, R.J., Arrigale, R.R., Donovan, G.P., Maddon, P.J., Dragic, T. and Olson, W.C. (2003) L-SIGN (CD 209L) is a liver-specific capture receptor for hepatitis C virus. *Proc Natl Acad Sci U S A*, **100**, 4498-4503.

Gebauer, F. and Hentze, M.W. (2004) Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol*, **5**, 827-835.

Giannini, C., Caini, P., Giannelli, F., Fontana, F., Kremsdorf, D., Brechot, C. and Zignego, A.L. (2002) Hepatitis C virus core protein expression in human B-cell lines does not significantly modify main proliferative and apoptosis pathways. *J Gen Virol*, **83**, 1665-1671.

Goffard, A. and Dubuisson, J. (2003) Glycosylation of hepatitis C virus envelope proteins. *Biochimie*, **85**, 295-301.

Gosert, R., Chang, K.H., Rijnbrand, R., Yi, M., Sangar, D.V. and Lemon, S.M. (2000) Transient Expression of Cellular Polypyrimidine-Tract Binding Protein Stimulates Cap-Independent Translation Directed by Both Picornaviral and Flaviviral Internal Ribosome Entry Sites In Vivo. *Mol. Cell. Biol.*, **20**, 1583-1595.

Gosert, R., Egger, D., Lohmann, V., Bartenschlager, R., Blum, H.E., Bienz, K. and Moradpour, D. (2003) Identification of the Hepatitis C Virus RNA Replication Complex in Huh-7 Cells Harboring Subgenomic Replicons. *J. Virol.*, **77**, 5487-5492.

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

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

Gretch, D., Corey, L., Wilson, J., dela Rosa, C., Willson, R., Carithers, R., Jr., Busch, M., Hart, J., Sayers, M. and Han, J. (1994) Assessment of hepatitis C virus RNA levels by quantitative competitive RNA polymerase chain reaction: high-titer viremia correlates with advanced stage of disease. *J Infect Dis*, **169**, 1219-1225.

Gretch, D.R., Polyak, S.J., Wilson, J.J., Carithers, R.L., Jr., Perkins, J.D. and Corey, L. (1996) Tracking hepatitis C virus quasispecies major and minor variants in symptomatic and asymptomatic liver transplant recipients. *J Virol*, **70**, 7622-7631.

Griffin, S.D., Beales, L.P., Clarke, D.S., Worsfold, O., Evans, S.D., Jaeger, J., Harris, M.P. and Rowlands, D.J. (2003) The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett*, **535**, 34-38.

Hadziyannis, S.J., Sette, H., Jr., Morgan, T.R., Balan, V., Diago, M., Marcellin, P., Ramadori, G., Bodenheimer, H., Jr., Bernstein, D., Rizzetto, M., Zeuzem, S., Pockros, P.J., Lin, A. and Ackrill, A.M. (2004) Peginterferon-{alpha}2a and Ribavirin Combination Therapy in Chronic Hepatitis C: A Randomized Study of Treatment Duration and Ribavirin Dose. *Ann Intern Med*, **140**, 346-355.

Hagan, H., Jarlais, D.C., Friedman, S.R., Purchase, D. and Alter, M.J. (1995) Reduced risk of hepatitis B and hepatitis C among injection drug users in the Tacoma syringe exchange program. *Am J Public Health*, **85**, 1531-1537.

Hahm, B., Kim, Y.K., Kim, J.H., Kim, T.Y. and Jang, S.K. (1998) Heterogeneous Nuclear Ribonucleoprotein L Interacts with the 3' Border of the Internal Ribosomal Entry Site of Hepatitis C Virus. *J. Virol.*, **72**, 8782-8788.

Hawkins, A., Davidson, F. and Simmonds, P. (1997) Comparison of plasma virus loads among individuals infected with hepatitis C virus (HCV) genotypes 1, 2, and 3 by quantiplex HCV RNA assay versions 1 and 2, Roche Monitor assay, and an in-house limiting dilution method. *J Clin Microbiol*, **35**, 187-192.

Hayashi, J., Kawakami, Y., Nabeshima, A., Kishihara, Y., Furusyo, N., Sawayama, Y., Kinukawa, N. and Kashiwagi, S. (1998a) Comparison of HCV RNA levels by branched DNA probe assay and by competitive polymerase chain reaction to predict effectiveness of interferon treatment for patients with chronic hepatitis C virus. *Dig Dis Sci*, **43**, 384-391.

Hayashi, J., Kishihara, Y., Ueno, K., Yamaji, K., Kawakami, Y., Furusyo, N., Sawayama, Y. and Kashiwagi, S. (1998b) Age-related response to interferon alfa treatment in women vs men with chronic hepatitis C virus infection. *Arch Intern Med*, **158**, 177-181.

He, L.F., Alling, D., Popkin, T., Shapiro, M., Alter, H.J. and Purcell, R.H. (1987) Determining the size of non-A, non-B hepatitis virus by filtration. *J Infect Dis*, **156**, 636-640.

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

Hellen, C.U., Pestova, T.V., Litterst, M. and Wimmer, E. (1994) The cellular polypeptide p57 (pyrimidine tract-binding protein) binds to multiple sites in the poliovirus 5' nontranslated region. *J Virol*, **68**, 941-950.

Hellen, C.U.T. and Sarnow, P. (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.*, **15**, 1593-1612.

Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M. and Shimotohno, K. (1991) Gene Mapping of the Putative Structural Region of the Hepatitis C Virus Genome by In Vitro Processing Analysis. *Proc Natl Acad Sci USA*, **88**, 5547-5551.

Hijikata, M., Mizushima, H., Akagi, T., Mori, S., Kakiuchi, N., Kato, N., Tanaka, T., Kimura, K. and Shimotohno, K. (1993a) Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus. *J Virol*, **67**, 4665-4675.

Hijikata, M., Mizushima, H., Tanji, Y., Komoda, Y., Hirowatari, Y., Akagi, T., Kato, N., Kimura, K. and Shimotohno, K. (1993b) Proteolytic Processing and Membrane Association of Putative Nonstructural Proteins of Hepatitis C Virus. *Proc Natl Acad Sci USA*, **90**, 10773-10777.

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

Hofgartner, W.T., Kant, J.A. and Weck, K.E. (2000) Hepatitis C virus quantitation: optimization of strategies for detecting low-level viremia. *J Clin Microbiol*, **38**, 888-891.

Holland, P., Abramson, R., Watson, R. and Gelfand, D. (1991) Detection of Specific Polymerase Chain Reaction Product by Utilizing the 5' {rightarrow} 3' Exonuclease Activity of Thermus aquaticus DNA Polymerase. *Proc Natl Acad Sci USA*, **88**, 7276-7280.

Honda, M., Kaneko, S., Matsushita, E., Kobayashi, K., Abell, G.A. and Lemon, S.M. (2000) Cell cycle regulation of hepatitis C virus internal ribosomal entry site-directed translation. *Gastroenterology*, **118**, 152-162.

Honda, M., Beard, M.R., Ping, L.H. and 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. *J Virol*, **73**, 1165-1174.

Honda, M., Brown, E.A. and Lemon, S.M. (1996a) Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation on hepatitis C virus RNA. *RNA*, **2**, 955-968.

Honda, M., Ping, L.H., Rijnbrand, R.C., Amphlett, E., Clarke, B., Rowlands, D. and Lemon, S.M. (1996b) Structural requirements for initiation of translation by internal ribosome entry within genome-length hepatitis C virus RNA. *Virology*, **222**, 31-42.

Honda, M., Rijnbrand, R., Abell, G., Kim, D. and Lemon, S.M. (1999c) Natural variation in translational activities of the 5' nontranslated RNAs of hepatitis C virus genotypes 1a and 1b: evidence for a long- range RNA-RNA interaction outside of the internal ribosomal entry site. *J Virol*, **73**, 4941-4951.

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

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

Hope, R.G., Murphy, D.J. and McLauchlan, J. (2002) The domains required to direct core proteins of hepatitis C virus and GB virus-B to lipid droplets share common features with plant oleosin proteins. *J Biol Chem*, **277**, 4261-4270.

Hsieh, T.-Y., Matsumoto, M., Chou, H.-C., Schneider, R., Hwang, S.B., Lee, A.S. and Lai, M.M.C. (1998) Hepatitis C Virus Core Protein Interacts with Heterogeneous Nuclear Ribonucleoprotein K. *J. Biol. Chem.*, **273**, 17651-17659.

Hussy, P., Langen, H., Mous, J. and 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. and Lai, M.M.C. (1997) Hepatitis C Virus NS5B Protein Is a Membrane-Associated Phosphoprotein with a Predominantly Perinuclear Localization. *Virology*, **227**, 439-446.

Imbert, I., Dimitrova, M., Kien, F., Kieny, M.P. and Schuster, C. (2003) Hepatitis C virus IRES efficiency is unaffected by the genomic RNA 3'NTR even in the presence of viral structural or non-structural proteins. *J Gen Virol*, **84**, 1549-1557.

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. and Thaler, H. (1995) Histological grading and staging of chronic hepatitis. *J hepatol*, **22**, 696-699.

Ishido, S., Fujita, T. and Hotta, H. (1998) Complex formation of NS5B with NS3 and NS4A proteins of hepatitis C virus. *Biochem Biophys Res Commun*, **244**, 35-40.

Ito, T. and Lai, M.M. (1997) Determination of the secondary structure of and cellular protein binding to the 3'-untranslated region of the hepatitis C virus RNA genome. *J Virol*, **71**, 8698-8706.

Ito, T. and Lai, M.M. (1999) An internal polypyrimidine-tract-binding protein-binding site in the hepatitis C virus RNA attenuates translation, which is relieved by the 3'-untranslated sequence. *Virology*, **254**, 288-296.

Ito, T., Mukaigawa, J., Zuo, J., Hirabayashi, Y., Mitamura, K. and 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. *J Gen Virol*, **77 (Pt 5)**, 1043-1054.

Ito, T., Tahara, S. and Lai, M. (1998) The 3'-untranslated region of hepatitis C virus RNA enhances translation from an internal ribosomal entry site. *J Virol*, **72**, 8789-8796.

Jang, S.J., Wang, L.F., Radkowski, M., Rakela, J. and Laskus, T. (1999) Differences between hepatitis C virus 5' untranslated region quasispecies in serum and liver. *J Gen Virol*, **80**, 711-716.

Jang, S.K., Krausslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C. and Wimmer, E. (1988) A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol*, **62**, 2636-2643.

Jin, L. and Peterson, D.L. (1995) Expression, isolation, and characterization of the hepatitis C virus ATPase/RNA helicase. *Arch Biochem Biophys*, **323**, 47-53.

Jinushi, M., Takehara, T., Kanto, T., Tatsumi, T., Groh, V., Spies, T., Miyagi, T., Suzuki, T., Sasaki, Y. and Hayashi et, a. (2003) Critical role of MHC class I-related chain A and B expression on IFN-alpha-stimulated dendritic cells in NK cell activation: impairment in chronic hepatitis C virus infection. *Journal Of Immunology (Baltimore, Md.: 1950)*, **170**, 1249-1256.

Jouet, P., Roudot-Thoraval, F., Dhumeaux, D. and Metreau, J.M. (1994) Comparative efficacy of interferon alfa in cirrhotic and noncirrhotic patients with non-A, non-B, C hepatitis. Le Groupe Francais pour l'Etude du Traitement des Hepatites Chroniques NANB/C. *Gastroenterology*, **106**, 686-690.

Jubin, R., Vantuno, N.E., Kieft, J.S., Murray, M.G., Doudna, J.A., Lau, J.Y. and Baroudy, B.M. (2000) Hepatitis C virus internal ribosome entry site (IRES) stem loop IIId contains a phylogenetically conserved GGG triplet essential for translation and IRES folding. *J Virol*, **74**, 10430-10437.

Kaito, M., Watanabe, S., Tsukiyama-Kohara, K., Yamaguchi, K., Kobayashi, Y., Konishi, M., Yokoi, M., Ishida, S., Suzuki, S. and Kohara, M. (1994) Hepatitis C virus particle detected by immunoelectron microscopic study. *J Gen Virol*, **75 (Pt 7)**, 1755-1760.

Kalliampakou, K.I., Psaridi-Linardaki, L. and Mavromara, P. (2002) Mutational analysis of the apical region of domain II of the HCV IRES. *FEBS Letters*, **511**, 79-84.

Kaminski, A., Hunt, S.L., Patton, J.G. and Jackson, R.J. (1995) Direct evidence that polypyrimidine tract binding protein (PTB) is essential for internal initiation of translation of encephalomyocarditis virus RNA. *RNA*, **1**, 924-938.

Kamoshita, N., Tsukiyama-Kohara, K., Kohara, M. and Nomoto, A. (1997) Genetic analysis of internal ribosomal entry site on hepatitis C virus RNA: implication for involvement of the highly ordered structure and cell type-specific transacting factors. *Virology*, **233**, 9-18.

Kanai, A., Tanabe, K. and Kohara, M. (1995) Poly(U) binding activity of hepatitis C virus NS3 protein, a putative RNA helicase. *FEBS Letters*, **376**, 221-224.

Kaneko, T., Tanji, Y., Satoh, S., Hijikata, M., Asabe, S., Kimura, K. and Shimotohno, K. (1994) Production of two phosphoproteins from the NS5A region of the hepatitis C viral genome. *Biochem Biophys Res Commun*, **205**, 320-326.

Kanistanon, D., Neelamek, M., Dharakul, T. and Songsivilai, S. (1997) Genotypic distribution of hepatitis C virus in different regions of Thailand. *J. Clin. Microbiol.*, **35**, 1772-1776.

Kato, J., Kato, N., Moriyama, M., Goto, T., Taniguchi, H., Shiratori, Y. and Omata, M. (2002) Interferons specifically suppress the translation from the internal ribosome entry site of hepatitis C virus through a double-stranded RNA-activated protein kinase-independent pathway. *J Infect Dis*, **186**, 155-163.

Kato, N. (2001) Molecular virology of hepatitis C virus. *Acta Med Okayama*, **55**, 133-159.

Kato, N., Ootsuyama, Y., Ohkoshi, S., Nakazawa, T., Sekiya, H., Hijikata, M. and Shimotohno, K. (1992) Characterization of hypervariable regions in the putative envelope protein of hepatitis C virus. *Biochem Biophys Res Commun*, **189**, 119-127.

Keck, Z.-Y., Sung, V.M.H., Perkins, S., Rowe, J., Paul, S., Liang, T.J., Lai, M.M.C. and Foung, S.K.H. (2004) Human Monoclonal Antibody to Hepatitis C Virus E1 Glycoprotein That Blocks Virus Attachment and Viral Infectivity. *J. Virol.*, **78**, 7257-7263.

Kieft, J.S., Zhou, K., Jubin, R. and Doudna, J.A. (2001) Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA*, **7**, 194-206.

Kieft, J.S., Zhou, K., Jubin, R., Murray, M.G., Lau, J.Y.N. and Doudna, J.A. (1999) The Hepatitis C Virus Internal Ribosome Entry Site Adopts an Ion-dependent Tertiary Fold,. *Journal of Molecular Biology*, **292**, 513-529.

Kim, D.W., Gwack, Y., Han, J.H. and Choe, J. (1995) C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. *Biochem Biophys Res Commun*, **215**, 160-166.

Kim, D.W., Kim, J., Gwack, Y., Han, J.H. and Choe, J. (1997) Mutational analysis of the hepatitis C virus RNA helicase. *J Virol*, **71**, 9400-9409.

Kim, Y.K., Lee, S.H., Kim, C.S., Seol, S.K. and Jang, S.K. (2003) Long-range RNA-RNA interaction between the 5' nontranslated region and the core-coding sequences of hepatitis C virus modulates the IRES- dependent translation. *RNA*, **9**, 599-606.

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

Koev, G., Duncan, R.F. and Lai, M.M.C. (2002) Hepatitis C Virus IRES-Dependent Translation Is Insensitive to an eIF2[alpha]-Independent Mechanism of Inhibition by Interferon in Hepatocyte Cell Lines. *Virology*, **297**, 195-202.

Kolupaeva, V.G., Pestova, T.V. and Hellen, C.U.T. (2000) An Enzymatic Footprinting Analysis of the Interaction of 40S Ribosomal Subunits with the Internal Ribosomal Entry Site of Hepatitis C Virus. *J. Virol.*, **74**, 6242-6250.

Kolykhalov, A.A., Feinstone, S.M. and Rice, C.M. (1996) Identification of a highly conserved sequence element at the 3' terminus of hepatitis C virus genome RNA. *J Virol*, **70**, 3363-3371.

Kolykhalov, A.A., Mihalik, K., Feinstone, S.M. and 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. *J. Virol.*, **74**, 2046-2051.

Konan, K.V., Giddings, T.H., Jr., Ikeda, M., Li, K., Lemon, S.M. and Kirkegaard, K. (2003) Nonstructural protein precursor NS4A/B from hepatitis C virus alters function and ultrastructure of host secretory apparatus. *J Virol*, **77**, 7843-7855.

Kong, L.K. and Sarnow, P. (2002) Cytoplasmic Expression of mRNAs Containing the Internal Ribosome Entry Site and 3' Noncoding Region of Hepatitis C Virus: Effects of the 3' Leader on mRNA Translation and mRNA Stability. *J. Virol.*, **76**, 12457-12462.

Konnick, E.Q., Erali, M., Ashwood, E.R. and Hillyard, D.R. (2002) Performance Characteristics of the COBAS Amplicor Hepatitis C Virus (HCV) Monitor, Version 2.0, International Unit Assay and the National Genetics Institute HCV Superquant Assay. *J. Clin. Microbiol.*, **40**, 768-773.

Kozak, M. (1989) The scanning model for translation: an update. *J. Cell Biol.*, **108**, 229-241.

Kozak, M. (2003) Alternative ways to think about mRNA sequences and proteins that appear to promote internal initiation of translation. *Gene*, **318**, 1-23.

Kumar, U., Thomas, H.C. and Monjardino, J. (1994) Serum HCV RNA levels in chronic HCV hepatitis measured by quantitative PCR assay; correlation with serum AST. *J Virol Methods*, **47**, 95-102.

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. and Houghton, M. (1989) An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science*, **244**, 362-364.

Lafuente, E., Ramos, R. and Martinez-Salas, E. (2002) Long-range RNA-RNA interactions between distant regions of the hepatitis C virus internal ribosome entry site element. *J Gen Virol*, **83**, 1113-1121.

Lanford, R.E., Sureau, C., Jacob, J.R., White, R. and 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.

Laporte, J., Bain, C., Maurel, P., Inchauspe, G., Agut, H. and Cahour, A. (2003) Differential distribution and internal translation efficiency of hepatitis C virus quasispecies present in dendritic and liver cells. *Blood*, **101**, 52-57.

Laporte, J., Malet, I., Andrieu, T., Thibault, V., Toulme, J.J., Wychowski, C., Pawlotsky, J.M., Huraux, J.M., Agut, H. and Cahour, A. (2000) Comparative analysis of translation efficiencies of hepatitis C virus 5' untranslated regions among intraindividual quasispecies present in chronic infection: opposite behaviors depending on cell type. *J Virol*, **74**, 10827-10833.

Laskus, T., Radkowski, M., Wang, L.F., Jang, S.J., Vargas, H. and Rakela, J. (1998) Hepatitis C virus quasispecies in patients infected with HIV-1: correlation with extrahepatic viral replication. *Virology*, **248**, 164-171.

Lau, J.Y.N., Davis, G.L., Prescott, L.E., Maertens, G., Lindsay, K.L., Qian, K., Mizokami, M., Simmonds, P., Perrillo, R.P., Schiff, E.R., Bodenheimer, H.C., Balart, L.A., Regenstein, F., Dienstag, J.L., Katkov, W.N., Tamburro, C.H., Goff, J.S., Everson, G.T., Goodman, Z. and et al. (1996) Distribution of hepatitis C virus genotypes determined by line probe assay in patients with chronic hepatitis C seen at tertiary referral centers in the United States. *Annals of Internal Medicine*, **124**, 868-876.

Layden-Almer, J.E., Ribeiro, R.M., Wiley, T., Perelson, A.S. and Layden, T.J. (2003) Viral dynamics and response differences in HCV-infected African American and white patients treated with IFN and ribavirin. *Hepatology*, **37**, 1343-1350.

Le, S.Y. and Maizel, J.V., Jr. (1998) Evolution of a common structural core in the internal ribosome entry sites of picornavirus. *Virus Genes*, **16**, 25-38.

Lee, Y.S., Yoon, S.K., Chung, E.S., Bae, S.H., Choi, J.Y., Han, J.Y., Chung, K.W., Sun, H.S., Kim, B.S. and Kim, B.K. (2001) The relationship of histologic

activity to serum ALT, HCV genotype and HCV RNA titers in chronic hepatitis C. *J Korean Med Sci*, **16**, 585-591.

Lehmann, M., Meyer, M.F., Monazahian, M., Tillmann, H.L., Manns, M.P. and Wedemeyer, H. (2004) High rate of spontaneous clearance of acute hepatitis C virus genotype 3 infection. *J Med Virol*, **73**, 387-391.

Lerat, H., Shimizu, Y.K. and Lemon, S.M. (2000) Cell Type-Specific Enhancement of Hepatitis C Virus Internal Ribosome Entry Site-Directed Translation due to 5' Nontranslated Region Substitutions Selected during Passage of Virus in Lymphoblastoid Cells. *J. Virol.*, **74**, 7024-7031.

Lesburg, C.A., Cable, M.B., Ferrari, E., Hong, Z., Mannarino, A.F. and Weber, P.C. (1999) Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat Struct Biol*, **6**, 937-943.

Levis, J., Kenny-Walsh, E., O'Sullivan, K., Horgan, M., Whelton, M., Shanahan, F. and Fanning, L. (2001) Strategy for the maximization of clinically relevant information from hepatitis C virus, RT-PCR quantification. *J Clin Virol*, **20**, 163-171.

Li, D., Takyar, S.T., Lott, W.B. and Gowans, E.J. (2003) Amino acids 1-20 of the hepatitis C virus (HCV) core protein specifically inhibit HCV IRES-dependent translation in HepG2 cells, and inhibit both HCV IRES- and cap-dependent translation in HuH7 and CV-1 cells. *J Gen Virol*, **84**, 815-825.

Lin, C., Lindenbach, B., Pragai, B., McCourt, D. and Rice, C. (1994) Processing in the hepatitis C virus E2-NS2 region: identification of p7 and two distinct E2-specific products with different C termini. *J. Virol.*, **68**, 5063-5073.

Lin, C., Wu, J.W., Hsiao, K. and Su, M.S. (1997) The hepatitis C virus NS4A protein: interactions with the NS4B and NS5A proteins. *J Virol*, **71**, 6465-6471.

Lissen, E., Alter, H.J., Abad, M.A., Torres, Y., Perez-Romero, M., Leal, M., Pineda, J.A., Torronteras, R. and Sanchez-Quijano, A. (1993) Hepatitis C virus infection among sexually promiscuous groups and the heterosexual partners of hepatitis C virus infected index cases. *Eur J Clin Microbiol Infect Dis*, **12**, 827-831.

Liu, Q., Bhat, R.A., Prince, A.M. and Zhang, P. (1999) The Hepatitis C Virus NS2 Protein Generated by NS2-3 Autocleavage Is Required for NS5A

Phosphorylation*1. *Biochemical and Biophysical Research Communications*, **254**, 572-577.

Livak, K.J., Flood, S.J., Marmaro, J., Giusti, W. and Deetz, K. (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl*, **4**, 357-362.

Lo, S., Selby, M. and Ou, J. (1996) Interaction between hepatitis C virus core protein and E1 envelope protein. *J. Virol.*, **70**, 5177-5182.

Lo, S.-Y., Selby, M., Tong, M. and Ou, J.-H. (1994) Comparative studies of the core gene products of two different hepatitis C virus isolates: two alternative forms determined by a single amino acid substitution. *Virology*, **199**, 124-131.

Logvinoff, C., Major, M.E., Oldach, D., Heyward, S., Talal, A., Balfe, P., Feinstone, S.M., Alter, H., Rice, C.M. and McKeating, J.A. (2004) Neutralizing antibody response during acute and chronic hepatitis C virus infection. *Proc Natl Acad Sci USA*, **101**, 10149-10154.

Lohmann, V., Körner, F., Koch, J.-O., Herian, U., Theilmann, L. and Bartenschlager, R. (1999) Replication of Subgenomic Hepatitis C Virus RNAs in a Hepatoma Cell Line. *Science*, **285**, 110-113.

Lu, H.-h. and Wimmer, E. (1996) Poliovirus chimeras replicating under the translational control of genetic elements of hepatitis C virus reveal unusual properties of the internal ribosomal entry site of hepatitis C virus. *Proc Natl Acad Sci USA* , **93**, 1412-1417.

Lu, M., Kruppenbacher, J. and Roggendorf, M. (2000) The importance of the quasispecies nature of hepatitis C virus (HCV) for the evolution of HCV populations in patients: study on a single source outbreak of HCV infection. *Arch Virol*, **145**, 2201-2210.

Lukavsky, P.J., Otto, G.A., Lancaster, A.M., Sarnow, P. and Puglisi, J.D. (2000) Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function. *Nat Struct Biol*, **7**, 1105-1110.

Lundin, M., Monne, M., Widell, A., Von Heijne, G. and Persson, M.A. (2003) Topology of the membrane-associated hepatitis C virus protein NS4B. *J Virol*, **77**, 5428-5438.

Luo, G., Xin, S. and Cai, Z. (2003) Role of the 5'-proximal stem-loop structure of the 5' untranslated region in replication and translation of hepatitis C virus RNA. *J Virol*, **77**, 3312-3318.

Luo, J.C., Hwang, S.J., Lai, C.R., Lu, C.L., Li, C.P., Tsay, S.H., Wu, J.C., Chang, F.Y. and Lee, S.D. (1998) Relationships between serum aminotransferase levels, liver histologies and virological status in patients with chronic hepatitis C in Taiwan. *J Gastroenterol Hepatol*, **13**, 685-690.

Lytle, J.R., Wu, L. and Robertson, H.D. (2002) Domains on the hepatitis C virus internal ribosome entry site for 40s subunit binding. *RNA*, **8**, 1045-1055.

Ma, H.C., Ke, C.H., Hsieh, T.Y. and Lo, S.Y. (2002) The first hydrophobic domain of the hepatitis C virus E1 protein is important for interaction with the capsid protein. *J Gen Virol*, **83**, 3085-3092.

Macdonald, A. and Harris, M. (2004) Hepatitis C virus NS5A: tales of a promiscuous protein. *J Gen Virol*, **85**, 2485-2502.

Mahaney, K., Tedeschi, V., Maertens, G., Di Bisceglie, A.M., Vergalla, J., Hoofnagle, J.H. and Sallie, R. (1994) Genotypic analysis of hepatitis C virus in American patients. *Hepatology*, **20**, 1405-1411.

Malet, I., Belnard, M., Agut, H. and Cahour, A. (2003) From RNA to quasispecies: a DNA polymerase with proofreading activity is highly recommended for accurate assessment of viral diversity. *J Virol Methods*, **109**, 161-170.

Manns, M.P. and Rambusch, E.G. (1999) Autoimmunity and extrahepatic manifestations in hepatitis C virus infection. *J hepatol Supplement*, **31**, 39-42.

Marcellin, P. (1999a) EASL International Consensus Conference on Hepatitis C. *J hepatol*, **31**, 1-2.

Marcellin, P. (1999b) Hepatitis C: the clinical spectrum of the disease. *J Hepatol*, **31 Suppl 1**, 9-16.

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

Martinot-Peignoux, M., Boyer, N., Le Breton, V., Le Guludec, G., Castelnau, C., Akremi, R. and Marcellin, P. (2000) A new step toward standardization of serum hepatitis C virus-RNA quantification in patients with chronic hepatitis C. *Hepatology*, **31**, 726-729.

Martinot-Peignoux, M., Marcellin, P., Pouteau, M., Castelnau, C., Boyer, N., Poliquin, M., Degott, C., Descombes, I., Le Breton, V., Milotova, V. and et al. (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 C. *Hepatology*, **22**, 1050-1056.

Mason, A.L., Lau, J.Y., Hoang, N., Qian, K., Alexander, G.J., Xu, L., Guo, L., Jacob, S., Regenstein, F.G., Zimmerman, R., Everhart, J.E., Wasserfall, C., Maclaren, N.K. and Perrillo, R.P. (1999) Association of diabetes mellitus and chronic hepatitis C virus infection. *Hepatology*, **29**, 328-333.

McCormick, S.E., Goodman, Z.D., Maydonovitch, C.L. and Sjogren, M.H. (1996) Evaluation of liver histology, ALT elevation, and HCV RNA titer in patients with chronic hepatitis C. *Am J Gastroenterol*, **91**, 1516-1522.

McCullough, A.J. (2003) Obesity and its nurturing effect on hepatitis C. *Hepatology*, **38**, 557-559.

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. and Albrecht, J.K. (1998) Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. *N Engl J Med*, **339**, 1485-1492.

McKechnie, V.M. and McCruden, E.A. (2001) Optimisation and sensitivity of single-stranded conformation polymorphism for the detection of hepatitis C virus quasi-species. *J Virol Methods*, **92**, 131-139.

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

McLauchlan, J., Lemberg, M.K., Hope, G. and Martoglio, B. (2002) Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *Embo J*, **21**, 3980-3988.

McLaughlin, K., Cameron, S., Good, T., McCruden, E., Ferguson, J., Davidson, F., Simmonds, P., Mactier, R. and McMillan, M. (1997) Nosocomial transmission of hepatitis C virus within a British dialysis centre. *Nephrol. Dial. Transplant.*, **12**, 304-309.

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. and Simmonds, P. (1994) Geographical distribution of hepatitis C virus genotypes in blood donors:and internation collaborative survey. *Journal of Clinical Microbiology*, **32**, 884-892.

Meerovitch, K., Svitkin, Y.V., Lee, H.S., Lejbkowicz, F., Kenan, D.J., Chan, E.K., Agol, V.I., Keene, J.D. and Sonenberg, N. (1993) La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *J Virol*, **67**, 3798-3807.

Mehta, S.H., Brancati, F.L., Strathdee, S.A., Pankow, J.S., Netski, D., Coresh, J., Szklo, M. and Thomas, D.L. (2003) Hepatitis C virus infection and incident type 2 diabetes. *Hepatology*, **38**, 50-56.

Memon, M.I. and Memon, M.A. (2002) Hepatitis C: an epidemiological review. *J Viral Hepat*, **9**, 84-100.

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., Tyrrell, D.L. and Kneteman, N.M. (2001) Hepatitis C virus replication in mice with chimeric human livers. *Nat Med*, **7**, 927-933.

Michalak, J.P., Wychowski, C., Choukhi, A., Meunier, J.C., Ung, S., Rice, C.M. and Dubuisson, J. (1997) Characterization of truncated forms of hepatitis C virus glycoproteins. *J Gen Virol*, **78**, 2299-2306.

Miller, R.H. and 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. *Proc Natl Acad Sci U S A*, **87**, 2057-2061.

Mizutani, T., Kato, N., Ikeda, M., Sugiyama, K. and Shimotohno, K. (1996) Long-term human T-cell culture system supporting hepatitis C virus replication. *Biochem Biophys Res Commun*, **227**, 822-826.

Monazahian, M., Bohme, I., Bonk, S., Koch, A., Scholz, C., Grethe, S. and Thomssen, R. (1999) Low density lipoprotein receptor as a candidate receptor for hepatitis C virus. *J Med Virol*, **57**, 223-229.

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

Moribe, T., Hayashi, N., Kanazawa, Y., Mita, E., Fusamoto, H., Negi, M., Kaneshige, T., Igimi, H., Kamada, T. and Uchida, K. (1995) Hepatitis C viral complexity

detected by single-strand conformation polymorphism and response to interferon therapy. *Gastroenterology*, **108**, 789-795.

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

Murakami, K., Abe, M., Kageyama, T., Kamoshita, N. and Nomoto, A. (2001) Down-regulation of translation driven by hepatitis C virus internal ribosomal entry site by the 3' untranslated region of RNA. *Arch Virol*, **146**, 729-741.

Murphy, E.L., Bryzman, S., Williams, A.E., Co-Chien, H., Schreiber, G.B., Ownby, H.E., Gilcher, R.O., Kleinman, S.H., Matijas, L., Thomson, R.A. and Nemo, G.J. (1996) Demographic determinants of hepatitis C virus seroprevalence among blood donors. *Jama*, **275**, 995-1000.

Mutimer, D.J., Harrison, R.F., O'Donnell, K.B., Shaw, J., Martin, B.A., Atrah, H., Ala, F.A., Skidmore, S., Hubscher, S.G., Neuberger, J.M. and et al. (1995) Hepatitis C virus infection in the asymptomatic British blood donor. *J Viral Hepat*, **2**, 47-53.

Nakata, S., Song, P., Duc, D.D., Nguyen, X.Q., Murata, K., Tsuda, F. and Okamoto, H. (1994) Hepatitis C and B virus infections in populations at low or high risk in Ho Chi Minh and Hanoi, Vietnam. *J Gastroenterol Hepatol*, **9**, 416-419.

Nakazawa, T., Kato, N., Ohkoshi, S., Shibuya, A. and Shimotohno, K. (1994) Characterization of the 5' noncoding and structural region of the hepatitis C virus genome from patients with non-A, non-B hepatitis responding differently to interferon treatment. *J Hepatol*, **20**, 623-629.

Navas, S., Martin, J., Quiroga, J.A., Castillo, I. and Carreno, V. (1998) Genetic diversity and tissue compartmentalization of the hepatitis C virus genome in blood mononuclear cells, liver, and serum from chronic hepatitis C patients. *J Virol*, **72**, 1640-1646.

Neddermann, P., Clementi, A. and De Francesco, R. (1999) Hyperphosphorylation of the hepatitis C virus NS5A protein requires an active NS3 protease, NS4A, NS4B, and NS5A encoded on the same polyprotein. *J Virol*, **73**, 9984-9991.

Nguyen, V.A., Chen, J., Hong, F., Ishac, E.J. and Gao, B. (2000) Interferons activate the p42/44 mitogen-activated protein kinase and JAK-STAT (Janus kinase-signal transducer and activator transcription factor) signalling pathways in

hepatocytes: differential regulation by acute ethanol via a protein kinase C-dependent mechanism. *Biochem J*, **349**, 427-434.

Nousbaum, J.-B., Polyak, S.J., Ray, S.C., Sullivan, D.G., Larson, A.M., Carithers, R.L., Jr. and Gretch, D.R. (2000) Prospective Characterization of Full-Length Hepatitis C Virus NS5A Quasispecies during Induction and Combination Antiviral Therapy. *J. Virol.*, **74**, 9028-9038.

Odreman-Macchioli, F., Baralle, F.E. and Buratti, E. (2001) Mutational analysis of the different bulge regions of hepatitis C virus domain II and their influence on internal ribosome entry site translational ability. *The Journal Of Biological Chemistry*, **276**, 41648-41655.

O'Farrell, D., Trowbridge, R., Rowlands, D. and Jager, J. (2003) Substrate Complexes of Hepatitis C Virus RNA Polymerase (HC-J4): Structural Evidence for Nucleotide Import and De-novo Initiation. *Journal of Molecular Biology*, **326**, 1025-1035.

Ogata, N., Alter, H.J., Miller, R.H. and Purcell, R.H. (1991) Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proc Natl Acad Sci U S A*, **88**, 3392-3396.

Oh, J.-W., Ito, T. and Lai, M.M.C. (1999) A Recombinant Hepatitis C Virus RNA-Dependent RNA Polymerase Capable of Copying the Full-Length Viral RNA. *J. Virol.*, **73**, 7694-7702.

Oh, J.-W., Sheu, G.-T. and Lai, M.M.C. (2000) Template Requirement and Initiation Site Selection by Hepatitis C Virus Polymerase on a Minimal Viral RNA Template. *J. Biol. Chem.*, **275**, 17710-17717.

Ohto, H., Terazawa, S., Sasaki, N., Hino, K., Ishiwata, C., Kako, M., Ujiie, N., Endo, C., Matsui, A. and et al. (1994) Transmission of hepatitis C virus from mothers to infants. The Vertical Transmission of Hepatitis C Virus Collaborative Study Group. *N Engl J Med*, **330**, 744-750.

Okamoto, H., Kurai, K., Okada, S., Yamamoto, K., Lizuka, H., Tanaka, T., Fukuda, S., Tsuda, F. and Mishiro, S. (1992) Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. *Virology*, **188**, 331-341.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci U S A*, **86**, 2766-2770.

Orito, E., Mizokami, M., Nakano, T., Terashima, H., Nojiri, O., Sakakibara, K., Mizuno, M., Ogino, M., Nakamura, M., Matsumoto, Y. and et al. (1994) Serum hepatitis C virus RNA level as a predictor of subsequent response to interferon-alpha therapy in Japanese patients with chronic hepatitis C. *J Med Virol*, **44**, 410-414.

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

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

Pawlotsky, J.M. (2003a) Hepatitis C virus genetic variability: pathogenic and clinical implications. *Clin Liver Dis*, **7**, 45-66.

Pawlotsky, J.-M. (2003b) Mechanisms of antiviral treatment efficacy and failure in chronic hepatitis C. *Antiviral Research*, **59**, 1-11.

Pawlotsky, J.-M. (2004) Pathophysiology of hepatitis C virus infection and related liver disease. *Trends in Microbiology*, **12**, 96-102.

Pawlotsky, J.M., Pellerin, M., Bouvier, M., Roudot-Thoraval, F., Germanidis, G., Bastie, A., Darthuy, F., Remire, J., Soussy, C.J. and Dhumeaux, D. (1998) Genetic complexity of the hypervariable region 1 (HVR1) of hepatitis C virus (HCV): influence on the characteristics of the infection and responses to interferon alfa therapy in patients with chronic hepatitis C. *J Med Virol*, **54**, 256-264.

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

Pelletier, J. and Sonenberg, N. (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*, **334**, 320-325.

Penin, F., Combet, C., Germanidis, G., Frainais, P.-O., Deleage, G. and Pawlotsky, J.-M. (2001) Conservation of the Conformation and Positive Charges of Hepatitis C Virus E2 Envelope Glycoprotein Hypervariable Region 1 Points to a Role in Cell Attachment. *J. Virol.*, **75**, 5703-5710.

Penin, F., Dubuisson, J., Rey, F.A., Moradpour, D. and Pawlotsky, J.M. (2004) Structural biology of hepatitis C virus. *Hepatology*, **39**, 5-19.

Perlemuter, G., Sabile, A., Letteron, P., Vona, G., Topilco, A., Chretien, Y., Koike, K., Pessayre, D., Chapman, J., Barba, G. and Brechot, C. (2002) Hepatitis C virus core protein inhibits microsomal triglyceride transfer protein activity and very low density lipoprotein secretion: A model of viral-related steatosis. *FASEB Journal*, **16**, 185-194.

Pestova, T., Hellen, C. and Shatsky, I. (1996a) Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol. Cell. Biol.*, **16**, 6859-6869.

Pestova, T., Shatsky, I. and Hellen, C. (1996b) Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol. Cell. Biol.*, **16**, 6870-6878.

Pestova, T.V. and Hellen, C.U.T. (1999) Internal Initiation of Translation of Bovine Viral Diarrhea Virus RNA. *Virology*, **258**, 249-256.

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

Petrik, J., Parker, H. and 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. *J Gen Virol*, **80**, 3109-3113.

Pietschmann, T., Lohmann, V., Kaul, A., Krieger, N., Rinck, G., Rutter, G., Strand, D. and Bartenschlager, R. (2002) Persistent and Transient Replication of Full-Length Hepatitis C Virus Genomes in Cell Culture. *J. Virol.*, **76**, 4008-4021.

Pileri, P., Uematsu, Y., Campagnoli, S., Galli, G., Falugi, F., Petracca, R., Weiner, A.J., Houghton, M., Rosa, D., Grandi, G. and Abrignani, S. (1998) Binding of hepatitis C virus to CD81. *Science*, **282**, 938-941.

Pilipenko, E., Blinov, V., Chernov, B., Dmitrieva, T. and Agol, V. (1989) Conservation of the secondary structure elements of the 5'-untranslated region of cardio- and aphthovirus RNAs. *Nucl. Acids. Res.*, **17**, 5701-5711.

Polyak, S.J., Khabar, K.S., Paschal, D.M., Ezelle, H.J., Duverlie, G., Barber, G.N., Levy, D.E., Mukaida, N. and Gretch, D.R. (2001) Hepatitis C virus

nonstructural 5A protein induces interleukin-8, leading to partial inhibition of the interferon-induced antiviral response. *J Virol*, **75**, 6095-6106.

Polywka, S., Schroter, M., Feucht, H.H., Zollner, B. and Laufs, R. (1999) Low risk of vertical transmission of hepatitis C virus by breast milk. *Clin Infect Dis*, **29**, 1327-1329.

Poynard, T., Leroy, V., Cohard, M., Thevenot, T., Mathurin, P., Opolon, P. and Zarski, J.P. (1996) Meta-analysis of interferon randomized trials in the treatment of viral hepatitis C: effects of dose and duration. *Hepatology*, **24**, 778-789.

Poynard, T., Marcellin, P., Lee, S.S., Niederau, C., Minuk, G.S., Ideo, G., Bain, V., Heathcote, J., Zeuzem, S., Trepo, C. and 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 for treatment of chronic infection with hepatitis C virus. *The Lancet*, **352**, 1426-1432.

Poynard, T., Ratziu, V., McHutchison, J., Manns, M., Goodman, Z., Zeuzem, S., Younossi, Z. and Albrecht, J. (2003) Effect of treatment with peginterferon or interferon alfa-2b and ribavirin on steatosis in patients infected with hepatitis C. *Hepatology*, **38**, 75-85.

Ramalho, F. (2003) Hepatitis C virus infection and liver steatosis. *Antiviral Res*, **60**, 125-127.

Racanelli, V. and Rehermann, B. (2003) Hepatitis C virus infection: when silence is deception. *Trends in Immunology*, **24**, 456-464.

Ray, R.B., Meyer, K. and Ray, R. (1996) Suppression of apoptotic cell death by hepatitis C virus core protein. *Virology*, **176**, 182-185.

Reddy, K.R., Hoofnagle, J.H., Tong, M.J., Lee, W.M., Pockros, P., Heathcote, E.J., Albert, D. and Joh, T. (1999) Racial differences in responses to therapy with interferon in chronic hepatitis C. Consensus Interferon Study Group. *Hepatology*, **30**, 787-793.

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

Rehermann, B. (2000) Interaction between the hepatitis C virus and the immune system. *Semin Liver Dis*, **20**, 127-141.

Reynolds, J., Kaminski, A., Kettinen, H., Grace, K., Clarke, B., Carroll, A., Rowlands, D. and Jackson, R. (1995) Unique features of internal initiation of hepatitis C virus RNA translation. *EMBO J.*, **14**, 6010-6020.

Reynolds, J.E., Kaminski, A., Carroll, A.R., Clarke, B.E., Rowlands, D.J. and Jackson, R.J. (1996) Internal initiation of translation of hepatitis C virus RNA: the ribosome entry site is at the authentic initiation codon. *RNA*, **2**, 867-878.

Rijnbrand, R., Bredenbeek, P., van der Straaten, T., Whetter, L., Inchauspe, G., Lemon, S. and Spaan, W. (1995) Almost the entire 5' non-translated region of hepatitis C virus is required for cap-independent translation. *FEBS Lett*, **365**, 115-119.

Rijnbrand, R., Bredenbeek, P.J., Haasnoot, P.C., Kieft, J.S., Spaan, W.J. and Lemon, S.M. (2001) The influence of downstream protein-coding sequence on internal ribosome entry on hepatitis C virus and other flavivirus RNAs. *RNA*, **7**, 585-597.

Rijnbrand, R.C. and Lemon, S.M. (2000) Internal ribosome entry site-mediated translation in hepatitis C virus replication. *Curr Top Microbiol Immunol*, **242**, 85-116.

Robertson, B., Myers, G., Howard, C., Brettin, T., Bukh, J., Gaschen, B., Gojobori, T., Maertens, G., Mizokami, M., Nainan, O., Netesov, S., Nishioka, K., Shin i, T., Simmonds, P., Smith, D., Stuyver, L. and Weiner, A. (1998) Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. International Committee on Virus Taxonomy. *Arch Virol*, **143**, 2493-2503.

Rockey, D.C. (2000) Hepatic fibrogenesis and hepatitis C. *Seminars in Gastrointestinal Disease*, **11**, 69-83.

Rose, J.K., Buonocore, L. and Whitt, M.A. (1991) A new cationic liposome reagent mediating nearly quantitative transfection of animal cells. *Biotechniques*, **10**, 520-525.

Roussel, J., Pillez, A., Montpellier, C., Duverlie, G., Cahour, A., Dubuisson, J. and Wychowski, C. (2003) Characterization of the expression of the hepatitis C virus F protein. *J Gen Virol*, **84**, 1751-1759.

Rubbia-Brandt, L., Fabris, P., Paganin, S., Leandro, G., Male, P.-J., Giostra, E., Carlotto, A., Bozzola, L., Smedile, A. and Negro, F. (2004) Steatosis affects chronic hepatitis C progression in a genotype specific way. *Gut*, **53**, 406-412.

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

Ruggieri, A., Harada, T., Matsuura, Y. and 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., Bizollon, T., Gouillat, C., Gripon, P., Guguen-Guillouzo, C. and Inchauspe et, a. (1999) Dynamic analysis of hepatitis C virus replication and quasispecies selection in long-term cultures of adult human hepatocytes infected in vitro. *The J Gen Virol*, **80 (Pt 11)**, 3007-3018.

Saiz, J.C., Lopez de Quinto, S., Ibarrola, N., Lopez-Labrador, F.X., Sanchez-Tapias, J.M., Rodes, J. and Martinez-Salas, E. (1999) Internal initiation of translation efficiency in different hepatitis C genotypes isolated from interferon treated patients. *Arch Virol*, **144**, 215-229.

Saldanha, J. (1999) Standardization: a Progress Report. *Biologicals*, **27**, 285-289.

Sanger, F., Nicklen, S. and Coulson, A.R. (1992) DNA sequencing with chain-terminating inhibitors. 1977. *Biotechnology*, **24**, 104-108.

Santolini, E., Migliaccio, G. and La Monica, N. (1994) Biosynthesis and biochemical properties of the hepatitis C virus core protein. *J Virol*, **68**, 3631-3641.

Saunier, B., Triyatni, M., Ulianich, L., Maruvada, P., Yen, P. and Kohn, L.D. (2003) Role of the asialoglycoprotein receptor in binding and entry of hepatitis C virus structural proteins in cultured human hepatocytes. *J Virol*, **77**, 546-559.

Scarselli, E., Ansuini, H., Cerino, R., Roccasecca, R.M., Acali, S., Filocamo, G., Traboni, C., Nicosia, A., Cortese, R. and Vitelli, A. (2002) The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *Embo J*, **21**, 5017-5025.

Schuster, C., Isel, C., Imbert, I., Ehresmann, C., Marquet, R. and Kieny, M.P. (2002) Secondary structure of the 3' terminus of hepatitis C virus minus- strand RNA. *J Virol*, **76**, 8058-8068.

Sentinelli, F., Lovari, S., Vitale, M., Giorgi, G., Di Mario, U. and Baroni, M.G. (2000) A simple method for non-radioactive PCR-SSCP using MDE (TM) gel solution and a midi gel format: Application for the detection of variants in the GLUT1 and CTLA-4 genes. *Journal of Biotechnology*, **78**, 201-204.

Serfaty, L., Andreani, T., Giral, P., Carbonell, N., Chazouilleres, O. and Poupon, R. (2001) Hepatitis C virus induced hypobetalipoproteinemia: a possible mechanism for steatosis in chronic hepatitis C. *J hepatol*, **34**, 428-434.

Sherf, B.A. (1996) Dual-luciferase reporter assay: an advanced co-repoter technology integrating firefly and renilla luciferase assays. *Promega Notes*, **57**, 2-9.

Shi, S.T., Polyak, S.J., Tu, H., Taylor, D.R., Gretch, D.R. and Lai, M.M. (2002) Hepatitis C virus NS5A colocalizes with the core protein on lipid droplets and interacts with apolipoproteins. *Virology*, **292**, 198-210.

Shimazaki, T., Honda, M., Kaneko, S. and Kobayashi, K. (2002) Inhibition of internal ribosomal entry site-directed translation of HCV by recombinant IFN-alpha correlates with a reduced La protein. *Hepatology*, **35**, 199-208.

Shimizu, Y.K., Feinstone, S.M., Kohara, M., Purcell, R.H. and Yoshikura, H. (1996) Hepatitis C virus: detection of intracellular virus particles by electron microscopy. *Hepatology*, **23**, 205-209.

Shimizu, Y.K., Igarashi, H., Kanematu, T., Fujiwara, K., Wong, D.C., Purcell, R.H. and Yoshikura, H. (1997) Sequence analysis of the hepatitis C virus genome recovered from serum, liver, and peripheral blood mononuclear cells of infected chimpanzees. *J Virol*, **71**, 5769-5773.

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

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

Simmonds, P. (2004) Genetic diversity and evolution of hepatitis C virus - 15 years on. *J Gen Virol*, **85**, 3173-3188.

Simmonds, P., Alberti, A., Alter, H.J., Bonino, F., Bradley, D.W., Brechot, C., Brouwer, J.T., Chan, S.-W., Chayama, K., Chen, D.-S., Choo, Q.-L., Colombo, M., Cuypers, H.T.M., Date, T., Dusheiko, G.M., Esteban, J.I., Fay, O., Hadziyannis, S.J., Han, J., Hatzakis, A., Holmes, E.C., Hotta, H., Houghton, M., Irvine, B., Kohara, M., Kolberg, J.A., Kuo, G., Lau, J.Y.N., Lelie, P.N., Maertens, G., McOmish, F., Miyamura, T., Mizokami, M., Nomoto, A., Prince, A.M., Reesink, H.W., Rice, C., Roggendorf, M., Schalm, S.W., Shikata, T., Shimotohno, K., Stuyver, L., Trepo, C., Weiner, A., Yap, P.L. and Urdea, M.S. (1994) A proposed system for the nomenclature of hepatitis C viral genotypes. *Hepatology*, **19**, 1321-1324.

Simmonds, P., Mellor, J., Craxi, A., Sanchez-Tapias, J.M., Alberti, A., Prieto, J., Colombo, M., Rumi, M.G., Lo Iacano, O., Ampurdanes-Mingall, S., Forns-Bernhardt, X., Chemello, L., Civeira, M.P., Frost, C. and Dusheiko, G. (1996) Epidemiological, clinical and therapeutic associations of hepatitis C types in western European patients. *J Hepatol*, **24**, 517-524.

Sizova, D.V., Kolupaeva, V.G., Pestova, T.V., Shatsky, I.N. and Hellen, C.U.T. (1998) Specific Interaction of Eukaryotic Translation Initiation Factor 3 with the 5' Nontranslated Regions of Hepatitis C Virus and Classical Swine Fever Virus RNAs. *J. Virol.*, **72**, 4775-4782.

Smith, D.B., Cuceanu, N., Davidson, F., Jarvis, L.M., Mokili, J.L., Hamid, S., Ludlam, C.A. and Simmonds, P. (1997a) Discrimination of hepatitis G virus/GBV-C geographical variants by analysis of the 5' non-coding region. *J Gen Virol*, **78 (Pt 7)**, 1533-1542.

Smith, D.B., Davidson, F. and Simmonds, P. (1995a) Hepatitis C virus variants and the role of genotyping. *J hepatol Supplement*, **23**, 26-31.

Smith, D.B., Davidson, F., Yap, P., Brown, H., Kolberg, J., Detmer, J., Urdea, M., Simmonds, P. and International, H.C.V.C. (1996) Levels of hepatitis C virus in blood donors infected with different viral genotypes. International HCV Collaborative Study Group. *Journal of Infectious Diseases*, **173**, 727-730.

Smith, D.B., McAllister, J., Casino, C. and Simmonds, P. (1997b) Virus 'quasispecies' making a mountain out of a molehill? *J Gen Virol*, **78**, 1511-1519.

Smith, D.B., Mellor, J., Jarvis, L.M., Davidson, F., Kolberg, J., Urdea, M., Yap, P.L. and Simmonds, P. (1995b) Variation of the hepatitis C virus 5' non-coding region: implications for secondary structure, virus detection and typing. The International HCV Collaborative Study Group. *J Gen Virol*, **76**, 1749-1761.

Smith, D.B. and Simmonds, P. (1997) Review: molecular epidemiology of hepatitis C virus. *J Gastroenterol Hepatol*, **12**, 522-527.

Soler, M., Pellerin, M., Malnou, C.E., Dhumeaux, D., Kean, K.M. and Pawlotsky, J.M. (2002) Quasispecies heterogeneity and constraints on the evolution of the 5' noncoding region of hepatitis C virus (HCV): relationship with HCV resistance to interferon-alpha therapy. *Virology*, **298**, 160-173.

Spahn, C.M.T., Kieft, J.S., Grassucci, R.A., Penczek, P.A., Zhou, K., Doudna, J.A. and Frank, J. (2001) Hepatitis C Virus IRES RNA-Induced Changes in the Conformation of the 40S Ribosomal Subunit. *Science*, **291**, 1959-1962.

Spangberg, K., Goobar-Larsson, L., Wahren-Herlenius, M. and 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. *J Hum Virol*, **2**, 296-307.

Spangberg, K., Wiklund, L. and Schwartz, S. (2000) HuR, a Protein Implicated in Oncogene and Growth Factor mRNA Decay, Binds to the 3' Ends of Hepatitis C Virus RNA of Both Polarities. *Virology*, **274**, 378-390.

Spangberg, K., Wiklund, L. and Schwartz, S. (2001) Binding of the La autoantigen to the hepatitis C virus 3' untranslated region protects the RNA from rapid degradation in vitro. *J Gen Virol*, **82**, 113-120.

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

Tabor, E., Gerety, R.J., Drucker, J.A., Seeff, L.B., Hoofnagle, J.H., Jackson, D.R., April, M., Barker, L.F. and Pineda-Tamondong, G. (1978) Transmission of non-A, non-B hepatitis from man to chimpanzee. *Lancet*, **1**, 463-466.

Tai, C., Chi, W., Chen, D. and Hwang, L. (1996) The helicase activity associated with hepatitis C virus nonstructural protein 3 (NS3). *J. Virol.*, **70**, 8477-8484.

Takada, N., Takase, S., Takada, A. and Date, T. (1993) Differences in the hepatitis C virus genotypes in different countries. *J Hepatol*, **17**, 277-283.

Takahashi, M., Yamada, G., Miyamoto, R., Doi, T., Endo, H. and Tsuji, T. (1993) Natural course of chronic hepatitis C. *Am J Gastroenterol*, **88**, 240-243.

Takeuchi, T., Katsume, A., Tanaka, T., Abe, A., Inoue, K., Tsukiyama-Kohara, K., Kawaguchi, R., Tanaka, S. and Kohara, M. (1999) Real-time detection system for quantification of hepatitis C virus genome. *Gastroenterology*, **116**, 636-642.

Tan, S.L., Pause, A., Shi, Y. and Sonenberg, N. (2002) Hepatitis C therapeutics: current status and emerging strategies. *Nat Rev Drug Discov*, **1**, 867-881.

Tanaka, T., Kato, N., Cho, M.-J., Sugiyama, K. and Shimotohno, K. (1996) Structure of the 3' terminus of the hepatitis C genome. *J Virol*, **70**, 3307-3312.

Tanaka, T., Kato, N., Hijikata, M. and Shimotohno, K. (1993) Base transitions and base transversions seen in mutations among various types of the hepatitis C viral genome. *FEBS Letters*, **315**, 201-203.

Tanji, Y., Hijikata, M., Hirowatari, Y. and Shimotohno, K. (1994) Hepatitis C virus polyprotein processing: kinetics and mutagenic analysis of serine proteinase-dependent cleavage. *J Virol*, **68**, 8418-8422.

Tanji, Y., Hijikata, M., Satoh, S., Kaneko, T. and Shimotohno, K. (1995a) Hepatitis C virus-encoded nonstructural protein NS4A has versatile functions in viral protein processing. *J. Virol.*, **69**, 1575-1581.

Tanji, Y., Kaneko, T., Satoh, S. and Shimotohno, K. (1995b) Phosphorylation of hepatitis C virus-encoded nonstructural protein NS5A. *J Virol*, **69**, 3980-3986.

Taylor, A., Goldberg, D., Hutchinson, S., Cameron, S., Gore, S.M., McMenamin, J., Green, S., Pithie, A. and Fox, R. (2000) Prevalence of Hepatitis C Virus Infection Among Injecting Drug Users in Glasgow 1990-1996: are Current Harm Reduction Strategies Working? *Journal of Infection*, **40**, 176-183.

Teoh, N.C. and Farrell, G.C. (2004) Management of chronic hepatitis C virus infection: a new era of disease control. *Intern Med J*, **34**, 324-337.

Thelu, M.A., Drouet, E., Hilleret, M.N. and Zarski, J.P. (2004) Lack of clinical significance of variability in the internal ribosome entry site of hepatitis C virus. *J Med Virol*, **72**, 396-405.

Thimme, R., Oldach, D., Chang, K.-M., Steiger, C., Ray, S.C. and Chisari, F.V. (2001) Determinants of Viral Clearance and Persistence during Acute Hepatitis C Virus Infection. *J. Exp. Med.*, **194**, 1395-1406.

Tokita, H., Shrestha, S.M., Okamoto, H., Sakamoto, M., Horikita, M., Iizuka, H., Shrestha, S., Miyakawa, Y. and Mayumi, M. (1994) Hepatitis C virus variants from Nepal with novel genotypes and their classification into the third major group. *J Gen Virol*, **75 (Pt 4)**, 931-936.

Toyoda, H., Kumada, T., Nakano, S., Takeda, I., Sugiyama, K., Osada, T., Kiriyama, S., Sone, Y., Kinoshita, M. and Hadama, T. (1997) Quasispecies nature of hepatitis C virus and response to alpha interferon: significance as a predictor of direct response to interferon. *J hepatol*, **26**, 6-13.

Tseng, C.T.K. and Klimpel, G.R. (2002) Binding of the hepatitis C virus envelope protein E2 to CD81 inhibits natural killer cell functions. *The Journal Of Experimental Medicine*, **195**, 43-49.

Tsuchihara, K., Tanaka, T., Hijikata, M., Kuge, S., Toyoda, H., Nomoto, A., Yamamoto, N. and 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. *J. Virol.*, **71**, 6720-6726.

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. and Nomoto, A. (1992) Internal ribosome entry site within hepatitis C virus RNA. *J Virol*, **66**, 1476-1483.

Tsung, K., Yim, J., Marti, W., Buller, R. and Norton, J. (1996) Gene expression and cytopathic effect of vaccinia virus inactivated by psoralen and long-wave UV light. *J. Virol.*, **70**, 165-171.

Urdea, M.S., Horn, T., Fultz, T.J., Anderson, M., Running, J.A., Hamren, S., Ahle, D. and Chang, C.A. (1991) Branched DNA amplification multimers for the sensitive, direct detection of human hepatitis viruses. *Nucleic Acids Symp Ser*, **24**, 197-200.

Van Damme, P. and Vellinga, A. (1998) Epidemiology of hepatitis B and C in Europe. *Acta Gastroenterol Belg*, **61**, 175-182.

van Leeuwen, H.C., Reusken, C.B.E.M., Roeten, M., Dalebout, T.J., Riezu-Boj, J.I., Ruiz, J. and Spaan, W.J.M. (2004) Evolution of naturally occurring 5' non-translated region variants of hepatitis C virus genotype 1b in selectable replicons. *J Gen Virol*, **85**, 1859-1866.

Varaklioti, A., Georgopoulou, U., Kakkanas, A., Psaridi, L., Serwe, M., Caselmann, W.H. and Mavromara, P. (1998) Mutational Analysis of Two Unstructured Domains of the 5' Untranslated Region of HCV RNA. *Biochemical and Biophysical Research Communications*, **253**, 678-685.

Varaklioti, A., Vassilaki, N., Georgopoulou, U. and Mavromara, P. (2002) Alternate translation occurs within the core coding region of the hepatitis C viral genome. *J Biol Chem*, **277**, 17713-17721.

Vizmanos, J.L., Gonzalez-Navarro, C.J., Novo, F.J., Civeira, M.P., Prieto, J., Gullon, A. and Garcia-Delgado, M. (1998) Degree and distribution of variability in the 5' untranslated, E1, E2/NS1 and NS5 regions of the hepatitis C virus (HCV). *J Viral Hepat*, **5**, 227-240.

Vuillermoz, I., Khattab, E., Sablon, E., Ottevaere, I., Durantel, D., Vieux, C., Trepo, C. and Zoulim, F. (2004) Genetic variability of hepatitis C virus in chronically infected patients with viral breakthrough during interferon-ribavirin therapy. *J Med Virol*, **74**, 41-53.

Wang, C., Le, S.Y., Ali, N. and Siddiqui, A. (1995) An RNA pseudoknot is an essential structural element of the internal ribosome entry site located within the hepatitis C virus 5' noncoding region. *RNA*, **1**, 526-537.

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

Wang, T.H., Rijnbrand, R.C. and 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. *J Virol*, **74**, 11347-11358.

Wedemeyer, H., He, X.-S., Nascimbeni, M., Davis, A.R., Greenberg, H.B., Hoofnagle, J.H., Liang, T.J., Alter, H. and Rehermann, B. (2002) Impaired effector function of hepatitis C virus-specific CD8+ T cells in chronic hepatitis C virus infection. *Journal Of Immunology (Baltimore, Md.: 1950)*, **169**, 3447-3458.

Weiner, A.J., Brauer, M.J., Rosenblatt, J., Richman, K.H., Tung, J., Crawford, K., Bonino, F., Saracco, G., Choo, Q.L., Houghton, M. and Han, J.H. (1991a) 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., Christopherson, C., Hall, J.E., Bonino, F., Saracco, G., Brunetto, M.R., Crawford, K., Marion, C.D., Crawford, K.A., Venkatakrishna, S. and et al. (1991b) Sequence variation in hepatitis C viral isolates. *J Hepatol*, **13 Suppl 4**, S6-14.

Westh, H., Worm, A.M., Jensen, B.L., Kroon, S., Kvinesdal, B., Nielsen, C.M. and Wantzin, P. (1993) Hepatitis C virus antibodies in homosexual men and intravenous drug users in Denmark. *Infection*, **21**, 115-117.

WHO. (1997) Hepatitis C: global prevalence. *Wkly Epidemiol Rec*, **72**, 341-344.

WHO. (1999) Global surveillance and control of hepatitis C. Report of a WHO Consultation organized in collaboration with the Viral Hepatitis Prevention Board, Antwerp, Belgium. *J Viral Hepat*, Vol. 6, pp. 35-47.

Wiese, M., Berr, F., Lafrenz, M., Porst, H. and Oesen, U. (2000) Low frequency of cirrhosis in a hepatitis C (genotype 1b) single-source outbreak in germany: a 20-year multicenter study. *Hepatology*, **32**, 91-96.

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

Wolk, B., Sansonno, D., Krausslich, H.-G., Dammacco, F., Rice, C.M., Blum, H.E. and Moradpour, D. (2000b) Subcellular Localization, Stability, and trans-

Cleavage Competence of the Hepatitis C Virus NS3-NS4A Complex Expressed in Tetracycline-Regulated Cell Lines. *J. Virol.*, **74**, 2293-2304.

Xu, Z., Choi, J., Lu, W. and Ou, J. h. (2003) Hepatitis C Virus F Protein Is a Short-Lived Protein Associated with the Endoplasmic Reticulum. *J. Virol.*, **77**, 1578-1583.

Xu, Z., Choi, J., Yen, T.S., Lu, W., Strohecker, A., Govindarajan, S., Chien, D., Selby, M.J. and Ou, J. (2001) Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. *Embo J*, **20**, 3840-3848.

Yamaga, A.K. and Ou, J. h. (2002) Membrane Topology of the Hepatitis C Virus NS2 Protein. *J. Biol. Chem.*, **277**, 33228-33234.

Yamamoto, C., Enomoto, N., Kurosaki, M., Yu, S.H., Tazawa, J., Izumi, N., Marumo, F. and Sato, C. (1997) Nucleotide sequence variations in the internal ribosome entry site of hepatitis C virus-1b: no association with efficacy of interferon therapy or serum HCV-RNA levels. *Hepatology*, **26**, 1616-1620.

Yan, Y., Li, Y., Munshi, S., Sardana, V., Cole, J.L., Sardana, M., Steinkuehler, C., Tomei, L., De-francesco, R., Kuo, L.C. and Chen, Z. (1998) Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: A 2.2 A resolution structure in a hexagonal crystal form. *Protein Sci*, **7**, 837-847.

Yanagi, M., Purcell, R.H., Emerson, S.U. and 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. *Proc Natl Acad Sci U S A*, **94**, 8738-8743.

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. and Simmonds, P. (1994) Hepatitis C virus transmission by intravenous immunoglobulin. *J hepatol*, **21**, 455-460.

Young, K., Archer, J., Yokosuka, O., Omata, M. and Resnick, R. (1995) Detection of hepatitis C virus RNA by a combined reverse transcription PCR assay: comparison with nested amplification and antibody testing. *J. Clin. Microbiol.*, **33**, 654-657.

Yuki, N., Hayashi, N., Moribe, T., Matsushita, Y., Tabata, T., Inoue, T., Kanazawa, Y., Ohkawa, K., Kasahara, A., Fusamoto, H. and Kamada, T. (1997) Relation of disease activity during chronic hepatitis C infection to complexity of hypervariable region 1 quasispecies. *Hepatology*, **25**, 439-444.

Zanetti, A.R., Tanzi, E., Romano, L., Zuin, G., Minola, E., Vecchi, L. and Principi, N. (1998) A prospective study on mother-to-infant transmission of hepatitis C virus. *Intervirology*, **41**, 208-212.

Zeuzem, S., Franke, A., Lee, J. H., Herrmann, G., Ruster, B. and Roth, W.K. (1996) Phylogenetic analysis of hepatitis C virus isolates and their correlation to viremia, liver function tests and histology. *Hepatology*, **24**, 1003-1009.

Zhang, J., Yamada, O., Yoshida, H., Iwai, T. and Araki, H. (2002) Autogenous translational inhibition of core protein: implication for switch from translation to RNA replication in hepatitis C virus. *Virology*, **293**, 141-150.

Zhu, N., Khoshnan, A., Schneider, R., Matsumoto, M., Dennert, G., Ware, C. and 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. *J Virol*, **72**, 3691-3697.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Acids Res.*, **31**, 3406-3415.