Evolution of the Predominant Sequence of the Hypervariable Region in the Putative Envelope Gene E2/NS1 of Hepatitis C Virus in Patients on Haemodialysis

by

Sharon Fitzpatrick

A thesis presented for the degree
of Doctor of Philosophy

in

The Faculty of Science
at the
University of Glasgow

Institute of Virology
Church Street
Glasgow
G11 5JR

October 1996

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

ACKNOWLEDGEMENTS

I am grateful to Professor J.H. Subak-Sharpe, and his successors, Professor Barklie Clements and Professor Duncan McGeoch for providing research facilities in the Institute of Virology.

I would like to thank my supervisor Dr Liz McCruden for her support throughout this project and her critical reading of this thesis.

I would also like to thank Dr Richard Elliott for his helpful advice and critical reading of parts of this thesis. I am also grateful to Dr Howard Marsden for his interesting discussions on aspects of this project, and to Dr June Munro and Dr Marion McElwee for their helpful advice throughout this project. Thanks also to the members of the Institute of Virology who answered my questions and gave me helpful advice on technical problems, especially Mr Aidan Dolan and Mrs Fiona Jamieson.

My thanks to the members of lab 311, past and present, in particular Michelle, Vicki and June for making the lab a very friendly place to work in and for providing the laughs which helped break the monotony of PCR. I would also like to thank all the friends that I have made throughout my time in the Institute.

I wish to thank the washroom and media staff who supplied the communal glassware and solutions used in this project.

I am grateful to my friends for putting up with my constant moaning throughout this project and during the preparation of this thesis.

Finally, I would like to thank my family, especially my parents, for their encouragement and support throughout the course of my studies.

The author was the recipient of an MRC studentship. Unless otherwise stated, all the work presented this thesis is by the author's own efforts.

SUMMARY

Follow-up studies of patients with acute HCV infection indicate that up to 50% of them develop a chronic infection. Hypervariability of the putative envelope protein E2/NS1 of HCV suggests that one mechanism by which the virus may persist in the host is by the generation of viral escape mutants. By this theory, a predominant population would evolve to which antibody is made which neutralises that viral population allowing the emergence of a new antigenically different predominant population. A characteristic clinical feature of HCV infection is the episodic fluctuating pattern of alanine transaminase (ALT) levels in which periods of elevated ALT levels are interspersed with periods of normal levels. The emergence of a new genetic variant may be expected to coincide with a period of liver dysfunction. Patients on long term renal dialysis are more likely to have elevated ALT levels than uninfected patients.

To characterise the emergence of genetic variants of the hypervariable region (HVR1) of the E2/NS1 protein and correlate their appearance with periods of liver dysfunction, a retrospective study of the HVR1 amplified from multiple serum samples taken over several years from three renal dialysis patients was performed. A total of 36 serum samples serially obtained from three patients with chronic hepatitis C were studied. HCV RNA was extracted from patient serum and the E2/NS1 HVR1 was amplified by reverse transcription and nested polymerase chain reaction. The amplified cDNA was analysed by direct chain termination nucleotide sequencing using the Sanger technique. The data indicate nucleotide sequence variation of HVR1 over time in each of the three renal dialysis patients but the emergence of a new sequence variant could not be linked with to a specific peak in ALT level. However, persistently raised ALT levels did appear to be associated with increased rate of change of the HVR1.

The second study presented in this thesis was undertaken to investigate the effect serum storage conditions had on the stability of HCV RNA and its detection by RT-PCR. Serum taken from a chronically infected patient was stored at 4°C for 14,

17, 19, 25 and 27 weeks and subjected to 1, 3, 5, 7 and 10 successive freeze-thaw cycles (-20°C to room temperature). To estimate the titre of HCV RNA, serial 10-fold dilutions of the serum were performed and HCV RNA extracted. PCR products were detected by agarose gel electrophoresis. Repeated freeze-thawing of serum or storage at 4°C did not result in a substantial reduction in HCV RNA levels detectable by RT-PCR compared to continuous storage at -20°C, indicating that HCV RNA present in serum is relatively resistant to degradation.

To allow quantitative PCR to be performed and to assess the relationship between serum HCV RNA levels and the emergence of new genetic variants, an attempt was made to synthesis control HCV RNA transcripts. PCR products amplified from the 5'NCR and core region of HCV were used as template to introduce a unique 28 nucleotide scramble sequence using PCR. Primers were designed with a unique sequence of bases towards the 5' end of each primer. Within the unique sequence an EcoR1 restriction enzyme site was incorporated. The first 10 bases of the unique sequence of the sense primer (Econ11) was complementary to the first 10 bases of the antisense primer (Econ12). This allowed the cDNA amplified with Econ11 and Econ12 to anneal together during amplification, resulting in a PCR product incorporating the unique sequence of bases. Unfortunately the final PCR product was contaminated with native HCV and was not cut when an EcoR1 digest was performed. The result of this was that control RNA transcripts were not synthesised and so quantitative PCR could not be performed. To overcome this the Econ11/Econ12 PCR product would need to be cloned into a plasmid containing a T7 promoter sequence and the resultant cDNA clones screened for those that contained the unique sequence of bases. Internal control HCV RNA could then be transcribed.

CONTENTS

Chapter 3 Methods

Chapter 4 Results: Section 1

Sequence evolution of the HVR1 of the putative envelope gene E2/NS1 of hepatitis C virus in patients on haemodialysis

Chapter 5 Results: Section 2

The effect of serum storage conditions on the detection of hepatitis C virus by the polymerase chain reaction

Chapter 6 Results: Section 3

Construction of a control RNA transcript for quantitative analysis of HCV infection

References

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

%	percentage
3'UTR	3' untranslated region
5'NCR	5' non-coding region
5'UTR	5' untranslated region
aa(s)	amino acid(s)
Ab	antibody
ALT	alanine aminotransferase level
anti-HCV	HCV antibody
ATP	adenosine triphosphate
B&W	buffer and wash solution
BCL	bootstrap confidence limit
bp	base pairs
BSA	bovine serum albumin
BVDV	bovine viral diarrhoea virus
C-	carboxy-
$CaCl_2$	calcium chloride
cDNA	complimentary deoxyribonucleic acid

CID	chimpanzee infectious doses
CTL	cytotoxic T lymphocyte
dATP	2'-deoxyadenosine- 5'-triphosphate
dCTP	2'-deoxycytidine- 5'-triphosphate
dGTP	2'-deoxyguanosine- 5'-triphosphate
dTTP	2'-deoxythymidine- 5'-triphosphate
ddATP	2', 3'-dideoxyadenosine-triphosphate
ddCTP	2', 3'-dideoxycytidine-triphosphate
ddGTP	2', 3'-dideoxyguanosine-triphosphate
ddTTP	2', 3'-dideoxythymidine-triphosphate
ddNTP	dideoxynucleotide-triphosphate
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide-triphosphate
DTT	dithiothreitol
E	envelope
EDTA	ethylenediamine tetraacetic acid
EIAV	equine infectious anaemia virus
ELISA	enzyme-linked immunosorbant assay
ER	endoplasmic reticulum
EtBr	ethidium bromide

FMDV	foot-and-mouth disease virus
g	gram
gp	glycoprotein
HAV	hepatitis A virus
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCV	hepatitis C virus
HDV	hepatitis D virus
HEV	hepatitis E virus
HFV	hepatitis F virus
HGV	hepatitis G virus
HDL	high density lipoprotein
HIV	human immunodeficiency virus
HIV-1	human immunodeficiency virus type-1
HLA	human leukocyte antigen
HoChV	hog cholera virus
HSV	herpes simplex virus
HVR	hypervariable region
IRES	internal ribosome entry site
LDL	low density lipoprotein

M	molar
M-MLV RT	Moloney Murine Leukaemia reverse transciptase
mAb	monoclonal antibody
mg	micrograms
$MgCl_2$	magnesium chloride
MHC	major histocompatability complex
mins	minutes
µl	microlitres
ml	millilitre
mM	millimolar
mmol	millimoles
N-	amino-
NaCl	sodium chloride
NANBH	non-A, non-B hepatitis
NaOH	sodium hydroxide
nm	nanometre
NS	non-structural
nt(s)	nucleotide(s)
O.D.	optical density

^{o}C	degrees centigrade
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pol	polymerase
PT	post-transfusion
RIBA	recombinant immunoblot assay
RNA	ribonucleic acid
RNAsin	ribonuclease inhibitor
RSV	respiratory syncytial virus
RT	reverse transcription
soln.	solution
SSCP	single stranded conformation polymorphism
TBE	tick-borne encephalitis
TCR	T-cell antigen receptor
TEMED	N, N, N'N - tetramethylenediamine
TFA	tubule forming agent
Tris-HCl	tris(hydroxymethy)-amino-methane
tRNA	transfer ribonucleic acid

U/L	units per litre
UPGMA	unweighted pair-group method with the arithmetic mean
UTR	untranslated region
UV	ultraviolet
V	variable
w/v	weight per volume
YFV	yellow fever virus

GCG PROGRAM NUCLEOTIDE SYMBOLS

Symbol	Meaning
A	A
C	C
G	G
T/U	T
M	A or C
R	A or G
W	A or T
S	C or G
Y	C or T
K	G or T
V	A or C or G
H	A or C or T
D	A or G or T
B	C or G or T
X/N	G or A or T or C

ONE AND THREE LETTER ABBREVIATIONS FOR AMINO ACIDS

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

Chapter 1 Introduction

1.1 Hepatitis Viruses: Description and classification

Viral hepatitis is caused by at least five distinct viruses. Hepatitis A virus (HAV) and hepatitis E virus (HEV) transmission is by the faecal-oral route and both viruses cause acute self-limiting hepatitis. The main route of hepatitis B virus (HBV) transmission is by sexual contact or from mother to child. The principal route of transmission for hepatitis C virus (HCV) and hepatitis D virus (HDV) is through exposure to contaminated blood or blood products. Unlike HAV or HEV, both HBV and HCV are associated with chronic liver disease. HCV in particular has a high rate of chronicity. Infection with several other viruses, for example cytomegalovirus and Epstein-Barr virus can also cause hepatitis (Haldor & Margolis, 1991).

Hepatitis A virus: HAV is the only hepatitis causing virus to have been isolated and propagated in cell culture (Purcell, 1994). The hepatitis A virion is a spherical particle, consisting of a single stranded positive sense RNA genome approximately 7.5kb long which is surrounded by a protein capsid comprising three proteins, VP1, VP2 and VP3. The HAV capsid is thought to have icosahedral symmetry. HAV is an enveloped virus. The genomic organisation of HAV is similar to viruses belonging to the *Picornaviridae* family of naked positive sense RNA viruses. Initially HAV was classified in the genus *enterovirus*, which includes the polio- and Coxsackieviruses, viruses within this genus preferentially inhabit the intestinal tract of the host. HAV has now been classified into a separate genus within the *Picornaviridae* family termed *hepatovirus*. HAV is very distinct in sequence from other entero viruses and has unique features, including hepatocyte cell tropism, possibly absence of VP4 protein and thermostability (reviewed by Lemon, 1994) .

Hepatitis B virus: HBV is an enveloped virus containing a circular partially double stranded DNA genome, approximately 3.2kb long. Unlike many other envelope viruses, no knobs or spikes have been observed in the HBV envelope (reviewed by Gerlich, 1993). HBV belongs to the family *Hepadnaviridae*, genus *orthohepadnavirus.* Other viruses included in this family are woodchuck hepatitis virus and ground squirrel hepatitis virus (Purcell, 1994). All viruses within this family exhibit narrow host ranges, hepatocyte tropism and an ability to cause persistent infection within the host.

Hepatitis C virus: The genome of HCV like that of HAV is a single stranded positive sense RNA virus, approximately 9.5kb long (Choo *et al.*, 1991) but unlike HAV, HCV is an enveloped virus (Bradley *et al.*, 1985). HCV displays sequence homology with pesti- and flaviviruses, small enveloped, single stranded positive sense RNA viruses, known to cause disease among animals (Miller & Purcell, 1990; Choo *et al.*, 1991). Miller and Purcell also demonstrated that HCV displays sequence homology with carmo- and potyviruses, a group of plant viruses. Positive stranded RNA viruses have been classified into 3 supergroups based upon analysis of their RNA dependent RNA polymerases (Koonin, 1991). HCV was classified into supergroup II with pestiviruses and carmoviruses but not with flaviviruses. Therefore, HCV has been classified into a separate as yet un-named genus within the family *Flaviviridae.* Evidence to support the classification of HCV as a member of the family *Flaviviridae* is discussed in detail later in this thesis.

Hepatitis D virus: HDV is only found in individuals with ongoing HBV infection. HDV is dependent on the presence of HBV for transmission and infection because it uses the hepatitis B surface antigen as its virion coat. HDV contains a small circular single stranded RNA genome, approximately 1700 nucleotides (nts). No decision has yet been reached to the taxonomic status of HDV. Initially HDV was classified as a member of the family *Hepadnaviridae* since it was only found when HBV was present. However HDV resembles viroids and plant satellite RNAs, in that the RNA genome is

circular, displays a high degree of intramolecular base pairing and has a similar mode of replication.

Hepatitis E virus: HEV is a spherical, non-enveloped virus containing a single-stranded, positive sense RNA genome, approximately 7.5kb. Like HDV, HEV is at present unclassified. It's genomic organisation resembles that of caliciviruses (Purcell, 1994), naked RNA viruses associated with respiratory and gastrointestinal infections in animals.

Despite diagnostic tests for the detection of known hepatitis viruses being available, the etiology of a proportion of post-transfusion (PT) and community-acquired hepatitis cases remain unexplained. Deka *et al.* (1994) reported the isolation of virus particles from human stool samples from patients suffering from sporadic NANBH. This virus has been provisionally designated hepatitis F virus (HFV). The genome of HFV, whose entire sequence is still to be determined, was found to be 20kb in size. The clinical importance of this virus is not yet known. In addition to the discovery of HFV, hepatitis G virus (HGV) a transfusion-transmissible agent isolated from the plasma of a patient with chronic PT hepatitis has also been reported. HGV is an RNA virus approximately 9.4kb long (Linnen *et al.*, 1996). The frequency with which HGV-RNA has been found in the population is similar to that of HCV. HGV is frequently found in coinfection with HBV and/or HCV.

Recent work by Simon *et al.* (1995) has shown the existence of two flavivirus-like genomes in plasma from a tamarin infected with the GB hepatitis agent (obtained from the serum of a 34 year old surgeon with acute hepatitis). The GB inoculum consists of two viruses, GB virus A (GBV-A) and GB virus B (GBV-B). The GBV-B agent is thought to have originated from tamarins since it replicates in the liver of these animals. GBV-A does not replicate in liver, but is found in lymphoid tissue (Muerhoff *et al.*, 1995). A third virus, GBV-C, has been isolated from a human specimen. GBV-C has been shown to have sequence similarity to GBV-A. Analysis of the genomes of

GBV-A, B, and C with that of HCV showed the three GB viruses were not genotypes of HCV (Schlauder *et al.*, 1995). To assess the clinical relevance of the GB-viruses and HGV it will be necessarily to undertake comprehensive studies to investigate their prevalence in blood donors and blood recipients, their worldwide distribution, their associtation with liver diseases.

1.2 Non-A, non-B hepatitis: Identification of hepatitis C virus

1.2.1 Early studies

The discovery in the mid-1960's of the hepatitis B surface antigen (HBsAg) led to the development of diagnostic tests with which to identify infection with HBV. HBV was the major cause of post-transfusion (PT) hepatitis prior to the introduction of screening for HBsAg. Despite the screening of blood donations for HBsAg, transfusion-associated hepatitis persisted. The development of diagnostic tests for HAV in the early 1970's indicated that the majority of PT hepatitis cases cases (a major world-wide health problem) were not the result of infection by HAV, HBV or by any other known hepatotropic viruses. Prince *et al.* (1974) therefore postulated the existence of another agent(s) responsible for PT-hepatitis. They studied 204 patients who had undergone cardiovascular surgery, in thirty-six of the fifty-one (71%) patients who developed PT-hepatitis there was no evidence of the HBV antigen (HBsAg) or anti-HBsAg. The possibility of HAV being responsible was excluded on the basis of the long incubation period. CMV was excluded because of equal exposure in those who did not contract hepatitis. Similar findings in studies conducted by Alter *et al.* (1975) and Feinstone *et al.* (1975) supported the existence of a new agent(s) as the cause of PT-hepatitis, this agent was termed non-A, non-B hepatitis (NANBH) virus.

1.2.2 Transmission studies in chimpanzees

In 1978 two separate studies reported the transmission of the NANBH agent to chimpanzees (Alter *et al.*, 1978; Tabor *et al.*, 1978). In that by Alter *et al.*, five chimpanzees were inoculated with plasma or serum obtained from patients with acute or chronic NANBH or from a blood donor implicated in two cases of NANBH. All five chimpanzees developed hepatitis based upon biochemical and histological evidence. No evidence of HAV or HBV infection was found in any of them, suggesting hepatitis was transmitted by the plasma or serum derived from both acute and chronic NANBH patients. In a similar study by Tabor *et al.* (1978), four chimpanzees were inoculated with serum obtained from a chronic NANBH patient or from two blood donors whose blood appeared to transmit hepatitis, all four chimpanzees developed hepatitis. Both studies demonstrated that the agent responsible for PT-NANBH was transmissible and that a chronic carrier state existed. Further evidence that the agent responsible for PT-NANBH was transmissible came from Bradley *et al.* (1979), who demonstrated that chimpanzees inoculated with a Factor VIII concentrate, implicated in the transmission of NANBH to human recipients, developed hepatitis.

The chimpanzee model for NANBH proved invaluable in the absence of specific serological assays for the agent responsible. Early cross-challenge and electron microscopic studies described the existence of more than one agent responsible for NANBH (Bradley *et al.*, 1980; Hollinger *et al.*, 1980). Chimpanzees were inoculated with either serum, plasma or Factor XIII concentrates implicated in cases of NANBH, after recovery from an episode of hepatitis the chimpanzees were re-challenged with a different inoculum and another episode of hepatitis developed. It is possible that the second episode of hepatitis was due to recurrence of infection from the first inoculum and not as a result of the second since a chronic carrier state existed for the agent(s) responsible for NANBH (Alter *et al.*, 1978; Tabor *et al.*, 1978). Bradley *et al.* (1980), in contrast to Shimizu *et al.* (1979), did not observe nuclear changes in hepatocytes

from infected chimpanzee livers in the acute phase of infection but did note the presence of cytoplasmic tubules. Shimizu *et al.* (1979) reported two different types of ultrastructural changes in infected hepatocytes. Chimpanzees were inoculated with plasma obtained from a patient with chronic NANBH (strain F) and from a patient with acute NANBH (strain H). Membranous tubular structures were seen in the cytoplasm of hepatocytes of chimpanzees infected with strain F, whereas strain H infected hepatocytes showed 20-27nm intranuclear particles. These ultrastructural changes were originally thought to represent two distinct NANBH agents, subsequent studies however have identified nuclear and cytoplasmic changes at the same time in experimentally infected chimpanzees (reviewed by Farci & Purcell, 1993). Neither HAV or HBV infected hepatocytes displayed such ultrastructural changes, although cytoplasmic tubular structures analogous to those observed in NANBH have been observed in chimpanzees infected with HDV, in the cytoplasm of mosquito cells infected with the flavivirus St. Louis encephalitis and in lymphocytes of patients with acquired immune deficiency disease, suggesting these ultrastructural changes are not exclusive to NANBH. The feature common to all viruses which cause such ultrastructural changes is an RNA genome (reviewed Bradley, 1990).

1.2.3 Cloning of non-A, non-B hepatitis infectious agent

In 1982, workers at the Chiron Corporation initiated attempts to clone the NANBH genome from infected chimpanzees, reporting the results of this work in 1989 (Choo *et al.*, 1989), they believed that the concentration of NANBH-specific nucleic acid in samples from infectious chimpanzees was below the limit detectable by existing techniques. In order to obtain infectious material with a high enough titre they pooled plasmapheresis samples from a chronically infected chimpanzee obtaining a preparation estimated to contain approximately 10^6 chimpanzee infectious doses per ml (CID/ml). The plasma was centrifuged until a pellet was obtained which would include even the smallest infectious agent. Total nucleic acid was extracted from the pellet, since it was

not known if the genome was RNA or DNA the extracted nucleic acid was denatured resulting in single stranded nucleic acid, cDNA was synthesised using random primers, cloned into the bacteriophage vector λgt11 and expressed in *Escherichia coli.* The cDNA library was screened with serum from a chronic NANBH patient and a single positive clone, 5-1-1, was identified from a total of about 10^6 clones.

To investigate the origin of clone 5-1-1, a larger overlapping 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 DNA derived from NANBH infected chimpanzees in Southern blot analysis, it did however hybridise to RNA extracted from the liver tissue of an NANBH infected chimpanzee, but not to RNA derived from an uninfected control. This suggested that a RNA molecule was associated with NANBH. Further evidence to support this hypothesis came as a result of the observation that the hybridisation signal of total nucleic acid extracted from the pellet fraction of NANBH infected chimpanzee ultracentrifuged plasma to clone 81 was lost following treatment with ribonuclease but not deoxyribonuclease. Clones 5-1-1 and 81 therefore appeared to be derived from the RNA genome of a blood-borne virus associated with NANBH infection. Further analysis showed it to be a positive single stranded RNA genome, approximately 10kb. The virus responsible for NANBH was termed hepatitis C virus (Choo *et al.*, 1989, 1991; Kato *et al.*, 1990; Takamizawa *et al.*, 1991).

The cloning of the NANBH agent, HCV, subsequently led to the development of an assay for HCV antibody detection (Kuo *et al.*, 1989). Using the original HCV clone 5-1-1 as a hybridization probe to the original cDNA library (Choo *et al.*, 1989), three overlapping clones, 81, 36 and 32, were isolated. Clones 5-1-1, 81, 36 and 32 were ligated together to re-construct part of the ORF, producing the clone C100 located in the NS4 region of the genome. C100 was expressed in recombinant yeast as a fusion polypeptide with the human gene encoding human superoxide dismutase (SOD), the resultant fusion polypeptide C100-3 was used to coat the wells of a micro-titre plate to capture circulating HCV antibodies in serum or plasma samples.

1.2.4 Biophysical properties of HCV

The results from cross-challenge experiments in chimpanzees suggested the possibility of two agents being responsible for NANBH. Chimpanzee infectivity studies conducted by Bradley *et al.* (1983) demonstrated the existence of chloroform sensitive and chloroform resistant NANBH agents. The NANBH agent associated with the formation of the cytoplasmic tubules in infected hepatocytes was sensitive to chloroform resulting in loss of infectivity in chimpanzees, this suggested the tubule forming agent (TFA) contained a lipid structure (virus envelope). Microfiltration studies demonstrated the tubule forming, chloroform sensitive agent to have a diameter of <80nm. Plasma from a chronically infected chimpanzee was diluted, pelleted and resuspended in buffer before being sequentially passed through 450-, 200- and 80nm polycarbonate (capillary pore) membrane filters, the resultant filtrate was inoculated into a chimpanzee, which subsequently developed hepatitis demonstrating the TFA had a diameter of <80nm (Bradley *et al.*, 1985). The agent was also recovered from infected plasma by the use of a method commonly used for the purification of small, enveloped RNA viruses (Bradley *et al.*, 1983, 1985). Subsequent studies, one using PCR to detect the virus in the filtrate suggested the diameter was between 30-38nm (He *et al.*, 1987; Yuasa *et al.*, 1991). The existence of a chloroform resistant agent is still to be confirmed.

Very little is known about the morphology of the HCV virion. Electron microscopy studies suggest the HCV virion it is a spherical particle, 55 to 65nm in diameter, with spike-like projections 6nm in length (Kaito *et al.*, 1994), similar to flaviviruses virions. In agreement with these findings are those of Shimizu *et al.* (1996). Using liver tissue from HCV infected chimpanzees and human B-cell and T-cell lines which supported HCV replication they performed electon microscopy studies and detected virion-like particles, approximately 50nm in diameter, within cytoplasmic vesicles. Such particles were not detected in uninfected cells. Putative viral core

particles 30 to 35nm in diameter and icoshedral in shape have also been identified in human plasma (Takahashi *et al.*, 1992a).

Several investigators have determined the buoyant density of HCV virions by gradient ultracentrifugation, followed by testing of gradient fractions for infectivity or the presence of HCV core antigen or RNA. The density of HCV was first investigated by Bradley *et al.* (1991), and estimated to be between 1.09 and 1.11g/ml. Buoyant density was determined by inoculating various dilutions of sucrose density gradient fractions into chimpanzees and monitoring them for the development of hepatitis. Subsequent work has used PCR to demonstrate the presence of HCV RNA in gradient fractions, indicating the density to be 1.08g/ml in sucrose, similar to the density of pestiviruses (1.09 to 1.16g/ml) rather than flaviviruses (1.19 to 1.20g/ml) (Miyamoto *et al.*, 1992). Following treatment with detergents the density increased to 1.25g/ml, corresponding to putative viral core particles. Takahaski *et al.* (1992a, 1992b) indicated HCV had a density of 1.11g/ml and the putative viral core particles 1.115g/ml in potassium bromide. The observed difference in density between these two groups is probably due to the composition and ionic strength of the gradient material used. The density, however, can be affected by the interaction between HCV virions and serum β-lipoprotein or immunoglobulins, varying between 1.03 and 1.20g/ml. (Thomssen *et al.*, 1992; Sato *et al.*, 1993; Hijikata, Minako *et al.*, 1993). Complexes of virus and antibody present in the high density fraction (1.17g/ml) were associated with low infectivity (taken during chronic phase of infection). In contrast such complexes were not found in serum samples with high infectivity (taken during acute phase of infection), when the virions were present in the low density fraction (Hijikata, Minako *et al.*, 1993). HCV virions, banding at a low density were found to coprecipitate with β-lipoprotein and those banding at a high density to coprecipitate with IgG (Thomssen *et al.*, 1992, 1993). It is unclear why HCV binds to β-lipoprotein, Thomssen *et al.* have suggested β–lipoprotein may mask antibody binding sites preventing the binding of neutralising antibodies, resulting in a persistent HCV infection or alternatively it may act as a ligand allowing HCV to infect host cells.

Mehdi *et al.* (1994, 1996) have shown in the case of HBV that recombinant HBsAg binds to a 46kDa protein associated with liver plasma membrane, plasma and lipoproteins. They have identified this protein as apolipoprotein H (apoH), a serum protein often associated with chylomicrons and high density lipoproteins (HDL). Both these lipoproteins are targeted to hepatocytes during normal lipid processing. Medhi *et al.* have proposed that a possible mechanism by which HBV can enter hepatocytes is via a lipoprotein-bound apoH. HBV binds to apoH on the surface of these lipoproteins via it's surface antigen, and is then taken into hepatocytes along with the lipoproteins, via lipoprotein receptors expressed on hepatocytes.

1.3 Viral replication

Evidence from chimpanzee studies suggests viral replication is an early event in HCV infection (Shimizu *et al.*, 1990; Negro *et al.*, 1992). The site of replication was found to be the cytoplasm of infected hepatocytes. In the study by Shimizu *et al.*, blood and liver biopsy samples were obtained from two chimpanzees inoculated with HCV. Sera were examined by RT-PCR and HCV RNA was detected as early as three days post-inoculation after a period during which no HCV-RNA was detected. Cytoplasmic antigen and ultrastructural changes (appearance of tubule structures) were detected in hepatocytes 3 and 6 days respectively, after HCV RNA was first detected in the serum. Circulating anti-HCV appeared between 3 to 33 weeks post-inoculation. The data from this study indicate a very early replicative phase for HCV and, potentially, a long period of infectivity before the appearance of anti-HCV. The results of this study were confirmed by that of Negro *et al.*, who used an *in situ* hybridization assay to detect HCV RNA in hepatocytes. Examination of liver biopsy samples taken from two HCV infected chimpanzees showed the presence of HCV RNA 2 days after both chimpanzees were inoculated with the virus. Because no suitable *in vitro* cell culture

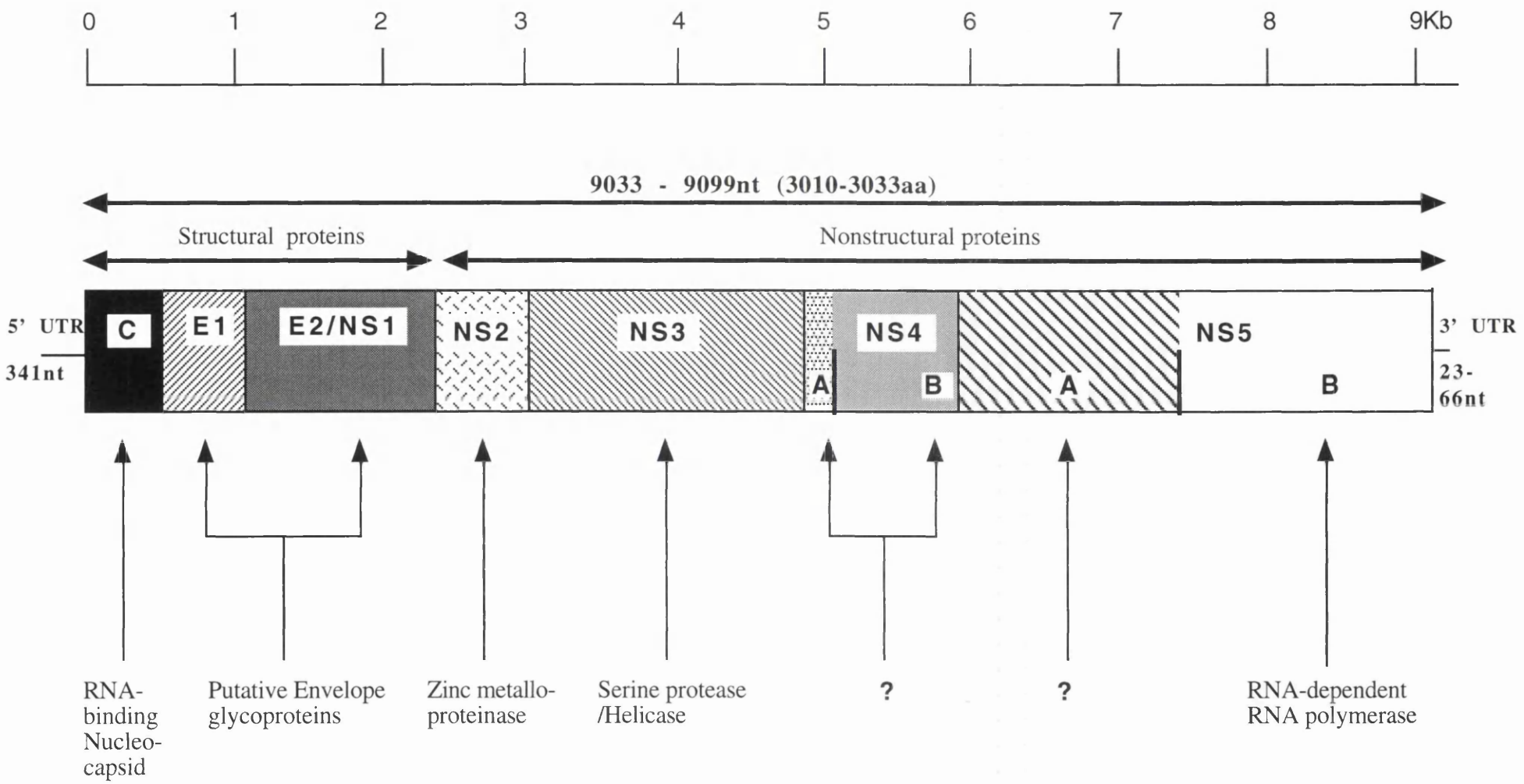
system for HCV exists very little is known about the replication cycle of HCV. HCV has been cultured with limited success in human T-cell lines (Shimizu *et al.*, 1992).

1.4 Hepatitis C virus: Structure, viral proteins and processing of the polyprotein

1.4.1 Genome organisation

HCV contains a positive stranded RNA genome of approximately 9400 nucleotides. A large ORF, spanning almost the entire length of the HCV RNA genome, encodes a precursor viral polyprotein of 3010 to 3030 amino acids which is flanked at the 5' and 3' termini by untranslated regions (UTR) (Choo *et al.*, 1989; Choo *et al.*, 1991; Takamizawa *et al.*, 1991; Okamoto & Mishiro, 1994a). Comparison of the hydrophobicity profiles of the encoded polyproteins of HCV, the flavivirus yellow fever virus (YFV) and the pestivirus bovine viral diarrhoea virus (BVDV) suggest HCV has a similar genetic organisation to flavi- and pestiviruses (Choo *et al.*, 1991; reviewed by Moennig & Plagemann, 1992; reviewed Rice & Strauss, 1990). A schematic representation of the genome organisation of HCV is shown in Figure 1.1. The HCV precursor viral polyprotein, like that of flavi- and pestiviruses, is proteolytically processed to yield both structural and non-structural (NS) proteins. The putative structural proteins (core, E1 and E2/NS1) are located at the 5' terminus of the polyprotein, with the remainder of the polyprotein consisting of the non-structural proteins (NS2, NS3, NS4a, NS4b, NS5a and NS5b).

Figure 1.1 Schematic representation of the HCV genomic organisation. The putative function of each gene is shown.


0
1
2
3
4
5
6
7
8
9Kb
9033 - 9099nt (3010-3033aa)
Structural proteins
Nonstructural proteins
5' UTR
341nt
C
E1
E2/NS1
NS2
NS3
NS4
A
B
NS5
A
B
3' UTR
23-
66nt
RNA-
binding
Nucleo-
capsid
Putative Envelope
glycoproteins
Zinc metallo-
proteinase
Serine protease
/Helicase
?
?
RNA-dependent
RNA polymerase

1.4.2 Classification of HCV as a member of the flaviviridae family

Sequence comparisons between HCV and other known viruses reveals little overall homology with any known viral sequence, however, the size of the precursor polyprotein of HCV is similar to that of flavi- and pestiviruses. Three regions of the HCV polyprotein were found to share amino acid homology with members of the flavivirus, pestivirus and plant potyvirus genera (Miller & Purcell, 1990). The first region of homology, is between amino acids 1230 and 1500 (HCV-1 isolate) and the helicase (NS3) domain of plant potyviruses, flavi- and pestiviruses suggesting this region of the HCV polyprotein encodes a helicase enzyme. The second region of homology lies upstream of the putative helicase region, sharing sequence similarities with the putative trypsin-like proteases thought to be encoded by flavi- and pestiviruses. The third, existed between amino acids 2703 and 2739, contains the amino acid motif Gly-Asp-Asp conserved among all RNA-dependent-RNA polymerases of positive stranded RNA viruses (Miller & Purcell, 1990; Choo *et al.*, 1991; Takamizawa *et al.*, 1991). The similarities displayed between HCV, flavi- and pestiviruses has led to its current classification as a separate genus within the family *Flaviviridae*. (Choo *et al.*, 1991; Han *et al.,* 1991; Takamizawa *et al.*, 1991).

1.4.3 The 5'and 3' untranslated regions

Sequence analysis has revealed the 5'UTR is highly conserved among HCV isolates (Bukh *et al.*, 1992). This suggests a possible functional role for this region in viral replication or gene translation (Han *et al.*, 1991; Houghton *et al.*, 1991; Bukh *et al.*, 1992). The 5'UTR comprises 341 nts or less, depending on the origin of the HCV isolate (Okamoto & Mishiro, 1994a). The precise 5' terminus is difficult to define. The 5'UTR of HCV appears to resemble that of pestiviruses (45-49% homology) rather than that of flaviviruses (Han *et al.*, 1991; Houghton *et al.*, 1991; Bukh *et al.*, 1992). The 5'UTR of flaviviruses is shorter (95 to 132nts) and lacks the small ORFs found in

both HCV and pestiviruses upstream of the initiator AUG codon (Han *et al.*, 1991). The HCV 5'UTR contains up to five small ORFs, that can encode polypeptides of between 12 to 28 amino acids. Different isolates of HCV display different patterns of these small ORFs (Bukh *et al.*, 1992). It is not known if there is a cap structure at the 5' terminus of the genomic RNA.

Experimental evidence indicates the existence of an internal ribosome entry site (IRES) within the 5'UTR proximal to the initiator AUG codon. (Tsukiyama-Kohara *et al.*, 1992; Wang *et al.*, 1993). Initiation of translation seems to occur in a cap-independent manner similar to that observed in picornaviruses. Picornavirus genomic RNAs are uncapped and have long 5'UTR (more than 600nts.) which contain silent AUGs. Tsukiyama-Kohara *et al.* (1992) presented evidence that an IRES element existed between nucleotides 101 and 332 (HCV-1b isolate). Work by Reynolds *et al.* (1994) however, has mapped the 5' end of the IRES to be between nucleotides 42 and 71 and the 3' end to nucleotides 354 and 372; thus extending into the coding region (starts at nucleotide 341) suggesting the involvement of the initial coding sequence of the genome in the function of the IRES. The secondary structures of the 5'UTR of HCV, the pestiviruses BVDV and hog cholera virus (HoChV) have been determined based upon thermodynamic, phylogenetic and biochemical considerations (Brown *et al.*, 1992). The proposed secondary structure of the 5'UTR of HCV is shown in Figure 1.2. The 5'UTR shares a large stem-loop structure (domain III) similar to that seen in the 5'UTR of pestiviruses. This conserved stem-loop structure is located between nucleotides 125 to 323, within the region identified by Tsukyama-Kohara *et al.* (1992) as containing the IRES. A common feature of the IRES of picornaviruses is the presence of pyrimidine-rich tracts. Within these tracts short conserved regions complementary to 18S ribosomal RNA have been identified. A similar region is found within the apical loop of domain III of the HCV 5'UTR (Figure 1.2) but mutational analysis has shown this region to be dispensable for IRES function (Wang *et al.*, 1994). Mutational analysis has also revealed that the primary sequence of the pyrimidine-rich tract (nts 120 to 130) (Figure 1.2) is not important but that the helical

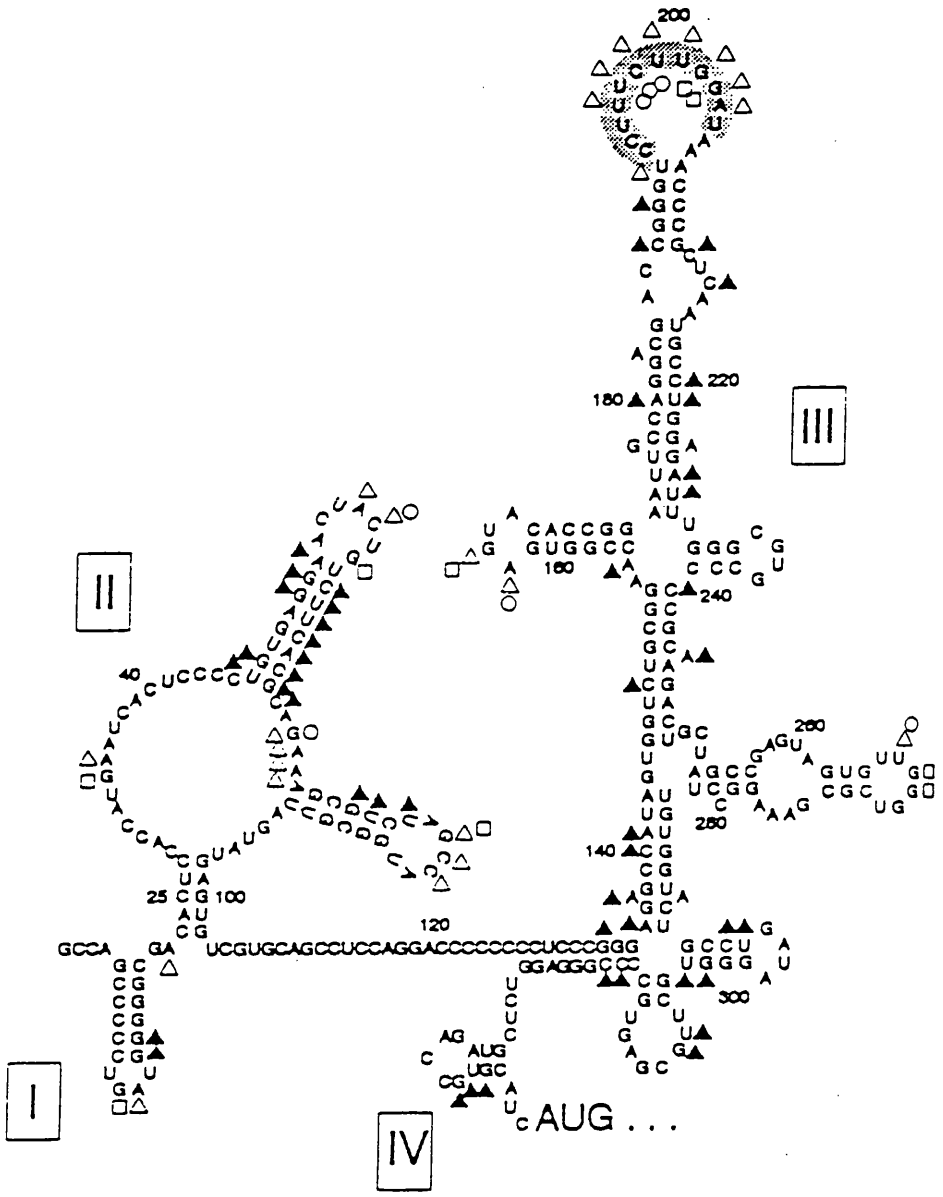


Figure 1.2 Proposed secondary structure of the HCV 5'NCR. Sites of nuclease cleavages with single-stranded specific Rnases are indicated by symbols adajacent to individual nucleotides: = T1, O = T2, Δ = S1. ▲ = V1 indicates double-stranded specific Rnase. The shaded nucleotides represent the pyrimidine-rich tract within the apical loop of domain III which is complementary to 18S ribosomal RNA. Domain III is thought to be essential for internal ribosomal entry. Reproduced by permission of Oxford University Press.

structural element associated with this region is essential for HCV IRES function (Wang *et al.*, 1994), this suggests the folding of the 5'UTR is critical for the function of the IRES. Based upon current knowledge, it seems likely that initiation of translation of the HCV genome occurs in a cap-independent manner via the IRES present within the 5'UTR before the initiator AUG codon. Recently, it has been suggested that translation initiation of the pestivirus BVDV occurs via an IRES contained in the 5'NCR (Poole *et al.*, 1995).

A short untranslated region is located downstream of the stop codon terminating translation of the ORF encoding the precursor viral polyprotein. The 3'UTR is variable among HCV isolates both in length and in sequence (Okamoto & Mishiro, 1994a). All HCV isolates except HCV-1 possess a poly(rU) tail which follows the 3'UTR. Isolate HCV-1 is followed by a poly(rA) tail (Han *et al.*, 1991). Recently Tanaka *et al.* (1996) and Kolykhalov *et al.* (1996) have both described a novel 98 nucleotide sequence downstream from the poly(U) stretch in the HCV genome which appears to represent the authentic 3' terminus. This sequence has been shown to be present in several HCV isolates and appears to be highly conserved amongst the isolates examined thus far. Using computer modeling programs, the 46 bases at the 3' terminus are predicted to form a conserved stem-loop structure. For other positive strand RNA viruses, conserved 3' terminal sequences or structures play a role in intiation of viral replication and in packaging of viral RNAs. This suggests that the HCV 3'UTR may play an critical role in initiation of viral replication.

1.4.4 HCV structural proteins

The structural proteins of HCV like those of the related flavi- and pestiviruses appear to be processed from the amino terminal of the HCV encoded precursor polyprotein (Figure 1.1). The putative structural proteins have been expressed both *in vitro* and *in vivo*. (in cell culture systems) (Takeuchi *et al.*, 1990b; Harada *et al.*, 1991; Hijikata *et*

al., 1991a; Matsuura *et al.*, 1992; Spaete *et al.*, 1992; Grakoui *et al.*, 1993a; Ralston *et al.*, 1993). Three structural proteins have been identified, core, E1 and E2/NS1.

Core: The first 191 amino acids of the precursor polyprotein are relatively well conserved among HCV isolates. Analogy of the HCV polyprotein with flavi- and pestiviruses suggests this region encodes the core protein which is probably the nucleocapsid component of the HCV virion. This region encodes a 21-22 kilodalton (KDa) protein, which lacks N-glycosylation sites and is highly basic (Takeuchi *et al.*, 1990b; Harada *et al.*, 1991; Hijikata *et al.*, 1991a; Grakoui *et al.*, 1993a). The core protein has been observed in the cytoplasm and nucleus of infected cells (Harada *et al.*, 1991; Matsuura *et al.*, 1992; Lanford *et al.*, 1993). The core protein observed in the nucleus of the cell lacked the (20aa) hydrophobic domain present at the C-terminal end and was not expressed with E1, unlike the core protein observed in the cell cytoplasm. Nuclear localisation of core protein has also been reported for flaviviruses. Initially the core protein of flaviviruses is produced in a membrane associated form, subsequent cleavage of the C-terminal hydrophobic domain converts it to mature core protein (Nowak *et al.*, 1989). Thus, the core protein of HCV may be subject to further processing events with the C-terminal hydrophobic domain acting as the signal sequence required for cleavage by host cell signal peptidases (Takeuchi *et al.*, 1990b). Experimental data shows that the core protein binds to ribosomes *in vitro* and appears to have RNA binding activity with the RNA binding domain residing between amino acids 1 to 75 in the highly basic N-terminal region of the protein (Santolini *et al.*, 1994). Antibody against the HCV core protein is found with high frequency in HCV infected patients and appears early in infection, but is does not appear to be neutralising (Harada *et al.*, 1991).

Envelope proteins: Two putative envelope proteins are located downstream from the core protein (Figure 1.1). The E1 gene comprises 192 amino acids and encodes a 31-35 kDa protein (gp35) which has 5 or 6 potential glycosylation sites (Hijikata *et*

al., 1991a; Grakoui *et al.*, 1993a; Ralston *et al.*, 1993). Deglycosylation of the E1 protein results in a size shift of the protein from 35kDa to 21kDa (Hijikata *et al.*, 1991a). Antibodies to gp35 are detected in patient sera suggesting that this protein is synthesised during infection (Matsuura *et al.*, 1992). The HCV gp35 appears to be equivalent to the envelope proteins of flavi- and pestiviruses.

The E2/NS1 gene encodes a 68-72kDa protein (Hijikata *et al.*, 1991a; Grakoui *et al.*, 1993a; Ralston *et al.*, 1993). This protein may represent a second envelope protein equivalent to the gp53/gp55 of the pestiviruses or represent the first non-structural protein equivalent to the flavivirus NS1 protein (Houghton *et al.*, 1991). Comparative sequence analyses suggest that the E2/NS1 gene is more closely related to pestiviruses rather than flaviviruses (Miller & Purcell, 1990; Choo *et al.*, 1991). The E2/NS1 is a glycosylated protein (11 potential glycosylation sites). Half of the mass of this protein is due to the mannose sugar residues, deglycosylation results in a 38kDa protein (Hijikata *et al.*, 1991a; Spaete *et al.*, 1992). A hypervariable region (HVR) is found in the N-terminus of the E2/NS1 protein (Weiner *et al.*, 1991). This region lacks conserved secondary structure and resembles the V3 loop of human immunodeficiency virus (HIV). This region of HCV will be discussed in detail later in this thesis.

In transfected mammalian cells both the E1 and E2/NS1 proteins are cell associated remaining anchored to the lumen of the endoplasmic reticulum (ER), unlike the flavivirus NS1 protein which is secreted. Removal of the C-terminal hydrophobic anchor region of E2/NS1 results in secretion of the protein into the medium (Spaete *et al.*, 1992; Ralston *et al.*, 1993). Amino acids 715 to 730 appear to serve as the membrane anchor (Selby *et al.*, 1994). For reasons of simplicity E2/NS1 will now be referred to as E2.

The expression of E1 and E2 appear to be associated. Monoclonal antibodies to either E1 or E2 are capable of immunoprecipitating both E1 and E2, suggesting the formation of E1:E2 complexes (Grakoui *et al.*, 1993a; Ralston *et al.*, 1993). Grakoui *et al.* (1993a) suggested E1 was associated with E2 via disulphide bonds based upon the reduction in association of the two proteins with dithiothretiol. In contrast to this

report Ralston *et al.* (1993) and Matsuura *et al.* (1994) were unable to demonstrate association via disulphide bonds. It has been reported that formation of E1:E2 complexes can occur via two pathways. The predominant pathway involves noncovalent interactions between E1 and E2 leading to the formation of a stable E1:E2 heterodimer, with only a minority of the E1:E2 complexes formed being stabilised via disulphide bonds. Dubussion *et al.* (1994) suggest that these represent misfolded complexes. E1:E2 complexes have been shown to confer a degree of immunity in experimentally infected chimpanzees (Choo *et al.*, 1994), suggesting vaccines towards HCV should incorporate both E1 and E2.

1.4.5 HCV non-structural proteins

Six proteins presumed to be non-structural (Figure 1.1) all appear to be unglycosylated. The NS2 gene encodes a 21 to 23kDa protein and the NS3 a 70kDa protein (Grakoui *et al.*, 1993a; Mizushima *et al.*, 1994a). Two different protease enzymes are encoded by NS2 and NS3 which are required for processing of the nonstructural region of the precursor polyprotein (see section 1.3.6). The NS3 protein contains a serine-like protease domain at the N-terminal and located immediately downstream is a helicase domain (Miller & Purcell, 1990; Choo *et al.*, 1991; Houghton *et al.*, 1991). The functions of these two domains within NS3 are not separated by proteolytic cleavage. The helicase activity of NS3 is presumed to be responsible for the unwinding of the RNA template during replication and translation. The residues His-1003, Asp-1107 and Ser-1165 found in the N-terminus of NS3 are highly conserved among HCV isolates and correspond to the catalytic triad found in serine type proteases. (Bartenschlager *et al.*, 1993; Grakoui *et al.*, 1993b; Tomei *et al.*, 1993). The N-terminal of the active protease of NS3 appears to reside between amino acids 1050 to 1083, with the C-terminal residing between amino acids 1115 and 1218 (Han *et al.*, 1995). Protease activity is inhibited by chymotrypsin-specific inhibitors but not by trypsin-like protease or protease inhibitors (Hahm *et al.*, 1995; Han *et al.*,

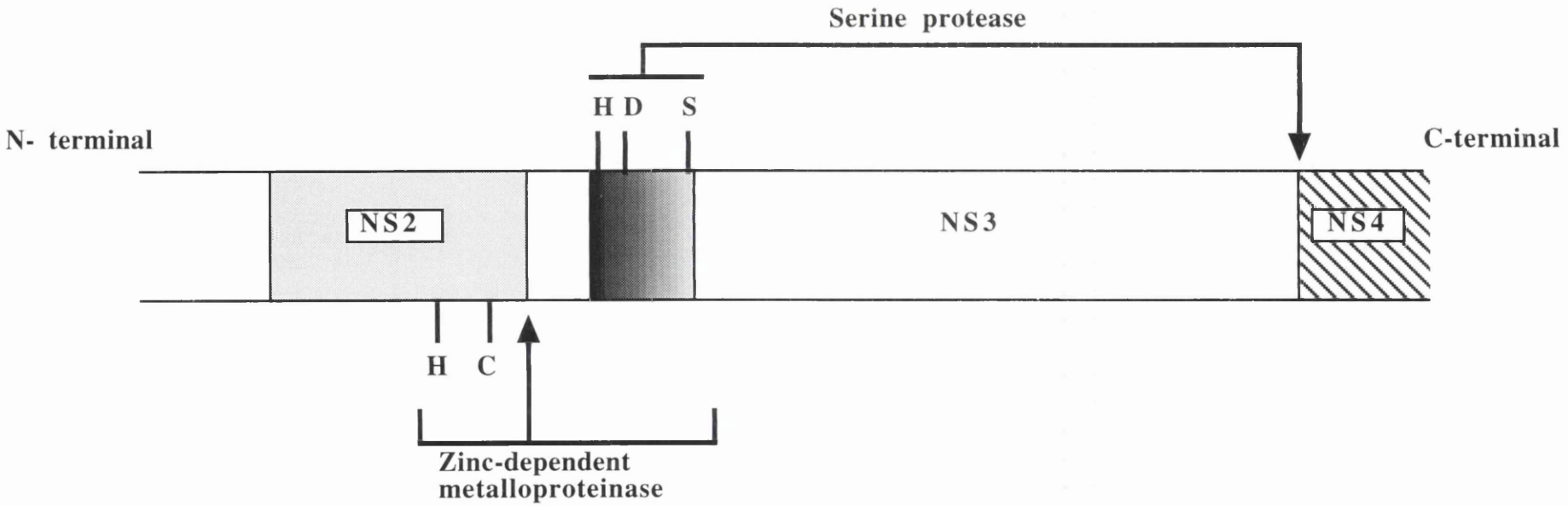
1995). A second protease enzyme exists between amino acids 827 to 1233, which appears to overlap regions of NS2 and NS3. It encompasses the C-terminal of NS2, the NS2/NS3 cleavage site and protease domain of NS3 (Grakoui *et al.*, 1993c; Hijikata *et al.*, 1993a). A schematic diagram showing the position of this protease is shown in Figure. 1.3. It appears to be a zinc-dependent metalloprotease (Hijikata *et al.*, 1993a). The zinc co-ordinate residues are proposed to be His-952 and Cys-993.

Downstream of NS3 are NS4a (4-8kDa) and NS4b (27kDa) (Grakoui *et al.*, 1993a; Hijikata *et al.*, 1993b). The precise functions of NS4a and NS4b are unknown. The NS5A and NS5b encode 56 to 58kDa and 66 to 68kDa proteins, respectively (Grakoui *et al.*, 1993a; Hijikata *et al.*, 1993b). NS5b contains the amino acid sequence motif Gly-Asp-Asp which is highly conserved among RNA dependent RNA polymerases encoded by RNA viruses and is presumably responsible for the replication of the RNA template (Choo *et al.*, 1991). Unlike HCV and pestiviruses, flaviviruses do not process the NS5 region into two distinct domains (reviewed by Collett *et al.*, 1988; reviewed by Rice & Strauss, 1990). Immunofluorescent assays have shown that apart from NS3 and NS5a all the non-structural proteins appear to be associated with membranes of the ER. NS3 and NS5a were detected in the cytoplasm, suggesting that they are soluble proteins (Selby *et al.*, 1993) .

1.4.6 Processing of the HCV polyprotein

The HCV precursor polyprotein appears to be co- and post-translationally processed. The N-termini of E1 and E2 have been sequenced following translation *in vitro* and cleavage sites mapped to Tyr-192 and His-384, respectively (Hijikata *et al.*, 1991a). The regions between amino acids 175-191 and 370-383 are hydrophobic and probably act as signal sequences for the host signal peptidase. The processing of the structural proteins *in vitro* is dependent upon the presence of microsomal membranes which suggests that cleavage of the structural proteins is mediated by host signal peptidases (Hijikata *et al.*, 1991a; Santolini *et al.*, 1994). Mature core protein appears to terminate

Figure 1.3 Schematic representation of the relative position of the two viral encoded proteases involved in processing of the HCV polyprotein. The shaded area represents the putative serine protease encoded by the NS3 gene, the predicted catalytic triad of this protease, His-1083, Asp-1107 and Ser- 1165, is indicated by H, D and S, respectively. The proposed zinc coordinating residues of the putative metalloproteinase, His- 952 and Cys- 993 are indicated by H and C, respectively. Adapted from Grakoui *et al.* (1993c).


aa
703
810
1027
1657
1764
Serine protease
H D S
N- terminal
C-terminal
NS2
NS3
NS4
H C
Zinc-dependent
metalloproteinase

around amino acid 174, however cleavage between core and E1 takes place at amino acid 191/192. A hydrophobic domain exists between amino acids 174 and 191. Experimental data indicates that the core protein is released from the precursor polyprotein by two cotranslational endoproteolytic cleavages mediated by host signal peptidases which generate the C-terminus of the core and the N-terminus of the E1 protein by removing the hydrophobic domain (Santolini *et al.*, 1994). Thus, from the experimental data accumulated so far, the structural proteins would appear to be cleaved by host signal peptidases.

Host signal peptidases are also believed to mediate cleavage between E2 and NS2 (Grakoui *et al.*, 1993a; Lin *et al.*, 1994). Three forms of E2 (36, 41 and 62kDa) were observed after endoglycosidase treatment (Grakoui *et al.*, 1993a). Grakoui *et al.* suggested the 62kDa protein represented E2 complexed with NS2 and the 36 and 41kDa proteins, the deglycosylated forms of E2. Selby *et al.* (1994) provides evidence that three forms of E2 with distinct C-termini occur representing E2 (terminating at amino acid 730), E2 complexed with the N-terminal of NS2 (terminates at amino acid 807) and E2 complexed with NS2 (terminates at amino acid 1026). The E2/NS2 complexes may be precursors for the E2 and NS2 proteins. The E2/NS2 junction was originally estimated to be around amino acid position 740, based upon deletion analysis of E2 (Hijikata *et al.*, 1991a). The cleavage site of E2/NS2 has now been mapped to amino acid 810 (Mizushima *et al.*, 1994a). The difference between the estimated and deduced cleavage site of E2/NS2 suggests that the region between amino acid 740 and 810 encodes a small polypeptide (Grakoui *et al.*, 1993a; Mizushima *et al.*, 1994a). A small hydrophobic 7kDa (p7) protein comprised of about 60 amino acids has been reported to be encoded by this region between E2 and NS2 (Lin *et al.*, 1994; Mizushima *et al.*, 1994b). A hydrophobic domain at the C-terminus of E2 preceding p7 may be involved in the translocation of p7 to the ER membrane thus allowing cleavage to be mediated by host signal peptidase. This theory is supported by the dependence of the cleavage of E2/p7 and p7/NS2 upon the presence of microsomal

membranes in a cell free system (Lin *et al.*, 1994). The function of this protein and whether it is structural or nonstructural is unknown.

The NS2/NS3 cleavage site has been mapped to amino acids 1026/1027 (Grakoui *et al.*, 1993c) (Figure 1.4). Cleavage occurs in *cis* by the zinc-dependent metalloprotease. Substitution of either of the zinc co-ordinate residues, His-952 and Cys-993 (Figure 1.3), with alanine results in cleavage being abolished (Grakoui *et al.*, 1993c; Hijikata *et al.*, 1993a). The NS3 chymoptrypsin-like protease is responsible for cleavage at NS3/NS4a, NS4a/NS4b, NS4b/NS5a and NS5a/NS5b. Cleavage at NS3/NS4a (aa 1657/1658) occurs in *cis* whereas cleavage at NS4a/NS4b (aa 1711/1712), NS4b/NS5a (aa 1972/1973) and NS5a/NS5b (aa 2420/2421) occur in *trans* (Tomei *et al.*, 1993). The cleavage recognition site for the NS3 protease is Asp/Glu -X-X-X-X-Cys/Thr ⇓ Ser/Ala where X is variable (Hijikata *et al.*, 1993b). Polar residues (Cys/Thr) are preferred at position 1 (P1) of the cleavage recognition site. The residues present at this position appear to be critical for recognition by the protease (Bartenschlager *et al.*, 1995). In contrast to this the acidic residue preferred at P6 does not appear to be essential for cleavage to occur. Acidic residues found in close proximity may compensate for the substitution of the acidic residue at P6 with a non-acidic residue (Figure 1.4). Mutational analysis of the NS3-dependent cleavage sites shows the NS3/4a cleavage site to have a higher degree of tolerance for amino acid substitutions within the cleavage recognition site in comparison to the sites cleaved in *trans.*, which suggests cleavage in *cis* is primarily governed by the structure of the polyprotein whereas cleavage in *trans* is governed by the interaction between the protease enzyme and substrate (Bartenschlager *et al.*, 1995a).

NS4a appears to be required for cleavage at the NS3/NS4a, NS4a/NS4b and NS4b/NS5a sites, and enhances cleavage efficiency between NS5a and NS5b (Bartenschlager *et al.*, 1994; Failla *et al.*, 1994). The NS3 protease and NS4a form a stable complex when expressed as a single polyprotein or as separate molecules. Deletion mapping experiments have shown that stable formation of the NS3/NS4a complex requires the N-terminal 22 amino acid residues of NS3 and the central NS4a

Cleavage site		P6	P1	P1'
3/4A	(1658)	**AspXXXX**	**Thr** ⇓	**Ser**
4A/4B	(1712)	**AspXXXX**	**Cys** ⇓	**Ser/Ala**
4B/5A	(1973)	**Glu/AspXXXX**	**Cys** ⇓	**Ser**
5A/5B	(2421)	**Glu/AspXXXX**	**Cys** ⇓	**Ser**

Figure 1.4 HCV NS3 serine protease-dependent cleavage sites. The cleavage recognition site is shown in boldface type. X is variable. The amino acid position of each cleavage site is shown in brackets. Adapted from Grakoui *et al.* (1993b).

domain between between amino acid residues 1675 and 1686 (Bartenschlager *et al.*, 1995b; Satoh *et al.*, 1995). A stable NS3/NS4a interaction is important for cleavage at the NS3/NS4a, NS4a/NS4b and NS4b/NS5a sites. NS4a may aid in the correct folding of the NS3 protease. Alternatively, NS4a may facilitate proteolysis by acting as an anchor for the hydrophilic NS3, bringing it into contact with the HCV proteins located in the membranes of the ER (Bartenschlager *et al.*, 1994).

Based on what is currently known about the mechanisms involved in processing of the non-structural polyprotein precursor the following model is proposed. The NS2/NS3 site is cleaved in *cis* by the NS2/NS3 zinc-dependent metalloprotease. Processing between NS3 and NS4a occurs cotranslationally by interaction of the N-terminal 22 amino acid residues of NS3 with the central domain of NS4a on the nascent polyprotein precursor. Cleavage at the NS3/NS4a site then occurs in *cis*, and a stable NS3/NS4a complex is formed which then mediates cleavage in *trans* at the NS4a/NS4b, NS4b/NS5a and NS5a/NS5b sites.

1.5 Genetic variability of hepatitis C virus

1.5.1 HCV genotypes

Following the discovery of HCV, sequence comparisons of HCV isolates from Japan, HCV-J, HCV-BK (Kato *et al.*, 1990; Takamizawa *et al.*, 1991) to the prototype isolate HCV-1 (Choo *et al.*, 1991) revealed significant differences in nucleotide sequence. HCV-J and HCV-BK showed 91.2% similarity to each other compared to only 78.4% and 77.1% similarity to HCV-1 respectively. This suggested that there were at least two subtypes of HCV. Analysis of full and partial sequences from different HCV isolates indicate HCV could be classified into distinct genotypes, which display significant differences in nucleotide sequence and geographic distribution. Each genotype can be further divided into subtypes, depending upon the degree of similarity

of isolates within the group. To date at least nine HCV genotypes are thought to exist (Bukh *et al.*, 1992; Chan *et al.*, 1992; Cha *et al.*, 1992; Simmonds *et al.*, 1993a; Tokita *et al.*, 1994, 1995). However, the classification of isolates from Vietnam and Thailand as types 7, 8 and 9 has been disputed. Tokita *et al.* (1994, 1995) used the unweighted pair-group method with the arithmetic mean (UPGMA) to construct phylogenetic trees when classifying the HCV isolates from Vietnam and Thailand. This method is not reliable when the nucleotide substitiution rate is not constant over time, as is likely in the case of HCV. Both Mellor *et al.* (1995) and Mizokami *et al.* (1996) suggest that the HCV isolates previously proposed to represent types 7, 8 and 9 should be classified as type 6 subtypes. Their findings are based upon re-analysis of the data from Tokita *et al.* (1994, 1995) using the neighbour-joining and Maximum-likelihood methods to construct phylogenetic trees. Such methods take into account variation in the nucleotide substitution rate in different lineages.

Several authors have proposed systems of nomenclature for classification of HCV isolates (Enomoto *et al.*, 1990; Cha *et al.*, 1992; Chan *et al.*, 1992; Okamoto *et al.*, 1992a; Simmonds *et al.*, 1993a). Several of these proposed systems do not allow for the classification of some variants, namely 4, 5 and 6, thus a new system of nomenclature has been proposed by Simmonds *et al.* (1993b). Table 1.1 displays the proposed systems of HCV nomenclature. Classification of HCV isolates as 2c, 4, 5 and 6 is based upon only partial sequence analysis of either the 5'UTR, core, E1 or NS5 regions of the genome. It is therefore possible that if the full genome of such isolates is sequenced they may be classified differently. Many authors have used the sequence of the 5'UTR successfully to type isolates, however, this region has proved to be unsuitable for subtyping isolates. Genotyping HCV isolates is important in understanding the natural history, diagnosis and clinical course of HCV infection. Evidence exists to suggest infection with different HCV genotypes may lead to a different clinical course of infection. Patients with HCV type 1 tend to have more severe liver disease (chronic hepatitis, cirrhosis or hepatocellular carcinoma) compared to types 2, 3 and 4 (Dusheiko *et al.*, 1994b). Response to interferon treatment appears

Table 1.1 Comparison of the proposed systems of nomenclature for HCV sequences. Several groups have proposed a system for classifying HCV isolates into genotypes and subtypes. NC denotes isolates not classified by proposed system of nomenclature. Adapted from paper by Simmonds *et al.* (1993b).

	Proposed system of nomenclature			
HCV Isolate	**Simmonds**	**Chiron**	**Enomoto**	**Okamoto**
HCV-1	1a	I	K-PT	I
HCV-J	1b	II	K-1	II
HCV-BK				
HC-G9	1c	NC	NC	NC
HC-J6	2a	III	K-2a	III
HC-J8	2b	III	K-2b	IV
ARG-6	2c	III	NC	NC
NZL-1	3a	IV	NC	V
Tb *	3b	IV	NC	VI
EG-16, 29, 33	4	NC	NC	NC
SA3, 4	5	V	NC	NC
HK-1, 2, 3, 4	6	NC	NC	NC

to be better in type 2 infected patients than type 1, especially type 1b (Takada *et al.*, 1992).

1.5.2 Geographic distribution of HCV genotypes

The genotypes 1a, 1b, 2a, 2b and 3a are found in the USA and Europe (reviewed by Dusheiko & Simmonds, 1994). The frequency which each particular genotype is found varies among countries. In the Netherlands type 1b appears to predominate (Cuypers *et al.*, 1991). In Japan and China the genotypes 1b, 2a and 2b predominate. In Japan type 1a is found only in haemophiliacs who received blood or blood products produced in the USA (Okamoto & Mishiro, 1994a). The distribution of genotypes in Taiwan is similar to that found in Japan.

Countries in the Middle East and Africa display a different distribution. On the basis of sequence comparisons in the core, E1 and NS5 regions, genotype 4 predominates in Egypt, Cameroon, Burundi and Zaire, with certain subtypes of genotype 4 being found only in Egypt. Type 5 appears to be mainly restricted to South Africa. Likewise type 6 has been detected in Hong Kong, Macau and Vietnam (Bukh *et al.*, 1993; Simmonds *et al.*, 1993a, 1993b; Stuyver *et al.*, 1993; Mellor *et al.*, 1995).

The different geographic distribution displayed by the genotypes of HCV could perhaps be due to the mode of transmission. Predominantly, HCV is transmitted via blood or blood products but sporadic cases of HCV infection has occurred where transmission by blood has been eliminated.

1.5.3 Sequence variation in the 5'UTR and putative capsid protein

Sequence comparisons between isolates of HCV shows substantial nucleotide sequence variation which is not uniformly distributed across the entire genome. Comparisons of the individual HCV isolates show the 5'UTR to be highly conserved

among different isolates. Table 1.2 shows the degree of nucleotide sequence similarity between the 5'UTR from different HCV isolates.

The 5'UTR of HCV consists of highly conserved domains interspersed between variable domains. The most variable domain is found between nucleotides -167 to -118 (numbering according to Choo *et al.*, 1991) (Bukh *et al.*, 1992). This region forms the highly stable stem loop structure observed in the HCV 5'UTR (Brown *et al.*, 1992; Tsukiyama-Kohara *et al.*, 1992). Nucleotide insertions are observed within this region in some isolates (Bukh *et al.*, 1992). Since these nucleotide insertions are found in the non-based paired terminal loop (Domain III Figure 1.2) they do not interfere with the stem loop structure. Three highly conserved domains were found at positions -263 to -246, -199 to -178 and -65 to -3 in 44 HCV isolates studied by Bukh *et al.* (1992). In HBV, sequence here is important for encapsidation. The HBV encapsidation signal forms a stable stem-loop structure. Two sets of repeated sequences termed R_2 (-CACTCC-) and R_3 (-CCCGGAG-) are also found in the 5'UTR of reported HCV isolates (Inchauspe *et al.*, 1991). These conserved domains and repeated sequences may represent regulatory elements important for viral replication. The putative nucleocapsid protein appears to the most highly conserved protein between members of the same and different genotype (Houghton *et al.*, 1991).

1.5.4 Sequence variation within the envelope region

In contrast to the 5'UTR and nucleocapsid protein, the putative envelope glycoproteins, E1 and E2 display a high degree of variability when HCV isolates are compared. Table 1.3 shows the degree of nucleotide and deduced amino acid sequence similarity between HCV isolates. Even across the putative envelope proteins variation is not uniformly spread, in E1 a moderately variable domain exists between amino acids 215 and 255 and at the N-terminus of E2 a hypervariable region (HVR1) exists (Weiner *et al.*, 1991). The precise numbering of the individual amino acids which constitute HVR1 appears to vary depending on the genotype of the isolate. HVR1 was

Table 1.2 Sequence similarities between the 5'NCR of HCV isolates*

	Percentage of nucleotide similarity						
	HCV-H	**HCV-J**	**HCV-BK**	**HC-G9**	**HC-J6**	**HC-J8**	**NZL-1**
HCV-1	99	98	98	99	93	93	90
HCV-H		98	98	99	94	94	90
HCV-J			99	98	93	93	93
HCV-BK				98	94	92	92
HC-G9					94	94	90
HC-J6						96	88
HC-J8							87

*HCV-1 (Choo *et al.*, 1991). HCV-H (Han *et al.*, 1991). HCV-BK (Takamizawa *et al.*, 1991). HCV-J (Kato *et al.*, 1990). HC-G9 (Okamoto *et al.*, 1994b). HC-J6 (Okamoto *et al.*, 1991). HC-J8 (Okamoto *et al.*, 1992a). NZL-1 (Sakamoto *et al.*, 1994).

Table 1.3 Sequence similarities between the envelope proteins of HCV isolates*

		Percentage of nucleotide (amino acid) similarity						
		HCV-H	**HCV-BK**	**HCV-J**	**HC-G9**	**HC-J6**	**HC-J8**	**NZL-1**
HCV-1	E1	93.9 (97.9)	74.8 (90.6)	75.5 (91.7)	76.6 (91.1)	62.5 (76.6)	55.9 (71.9)	64.6 (81.3)
	E2	93.7 (96.9)	71.2 (89.2)	72.6 (89.2)	75.8 (89.9)	64.7 (82.9)	65.3 (82.4)	64.9 (83.3)
HCV-H	E1		73.9 (90.6)	74.3 (91.7)	76.7 (91.7)	61.3 (77.6)	55.9 (72.4)	64.9 (81.8)
	E2		71.6 (89.4)	73.2 (89.9)	76.1 (89.9)	64.5 (82.6)	65.6 (82.6)	65.6 (82.8)
HCV-BK	E1			90.6 (96.3)	74.5 (89.1)	59.2 (73.9)	54.5 (72.4)	64.8 (83.3)
	E2			87.7 (91.8)	71.5 (88.7)	68.5 (84.0)	65.8 (82.4)	65.3 (81.4)
HCV-J	E1				74.5 (89.0)	59.4 (74.5)	55.4 (72.9)	64.8 (83.6)
	E2				72.5 (88.7)	68.1 (83.8)	65.4 (82.6)	65.3 (82.2)
HC-G9	E1					59.9 (77.1)	55.2 (71.9)	62.7 (79.7)
	E2					65.9 (82.8)	65.0 (83.1)	65.1 (83.1)
HC-J6	E1						69.3 (83.3)	59.5 (74.5)
	E2						72.1 (89.1)	64.3 (83.7)
HC-J8	E1							57.8 (73.4)
	E2							62.7 (82.1)

*HCV-1 (Choo *et al.*, 1991). HCV-H (Han *et al.*, 1991). HCV-BK (Takamizawa *et al.*, 1991). HCV-J (Kato *et al.*, 1990). HC-G9 (Okamoto *et al.*, 1994b). HC-J6 (Okamoto *et al.*, 1991). HC-J8 (Okamoto *et al.*, 1992a). NZL-1 (Sakamoto *et al.*, 1994).

mapped to amino acids 384 to 414 for type 1a (Weiner *et al.*, 1991, 1992), this hypervariable region (HVR1) represents only 8% of the total nucleotide sequence but accounts for between 30 to 47% of the nucleotide changes and 40 to 63% of the amino acid changes observed in the envelope region between HCV isolates (Weiner *et al.*, 1991). For type 1b isolates HVR1 was mapped to amino acids 384 to 410 (Kato *et al.*, 1992a). Hijikata *et al.* (1991b) identified a second hypervariable region (HVR2) between amino acids 474 to 480 in Japanese type 1b isolates. The equivalent region in type 1a isolates seldom shows any sequence variation. The position of the variable and hypervariable regions are shown schematically in Figure 1.5.

Despite its high degree of divergence 26 cysteine residues present in the envelope region are completely conserved, suggesting they may be important in forming inter- and intra- disulphide bonds essential for conformation of the envelope proteins. N-glycosylation sites also appear to be conserved (Kato *et al.*, 1992a; Okamoto *et al.*, 1992a). This conservation of secondary structure does not apply to the hypervariable region which lacks both cysteine residues and potential glycosylation sites. The lack of secondary structure and high degree of variation observed in this region is similar to that found in the V3 loop of HIV envelope gp120, this suggests that HVR1 like the V3 loop of HIV may be the target of the host's humoral immune system (Weiner *et al.*, 1992). Within this hypervariable region certain amino acids appear to be well conserved among isolates. i.e threonine at position 385, glycine at position 406, glutamine at position 409 and leucine at position 413 (Kato *et al.*, 1992a; Lesniewski *et al.*, 1993). It is possible these amino acids play a role in receptor recognition.

1.5.5 Viral Quasispecies

RNA viruses exhibit a high degree of genetic diversity, due to the lack of proof reading activity of the RNA polymerase. RNA viruses therefore replicate with limited fidelity resulting in the generation of genomic variants which may have a phenotypic selective advantage in the host. These variants would then subsequently replicate and become

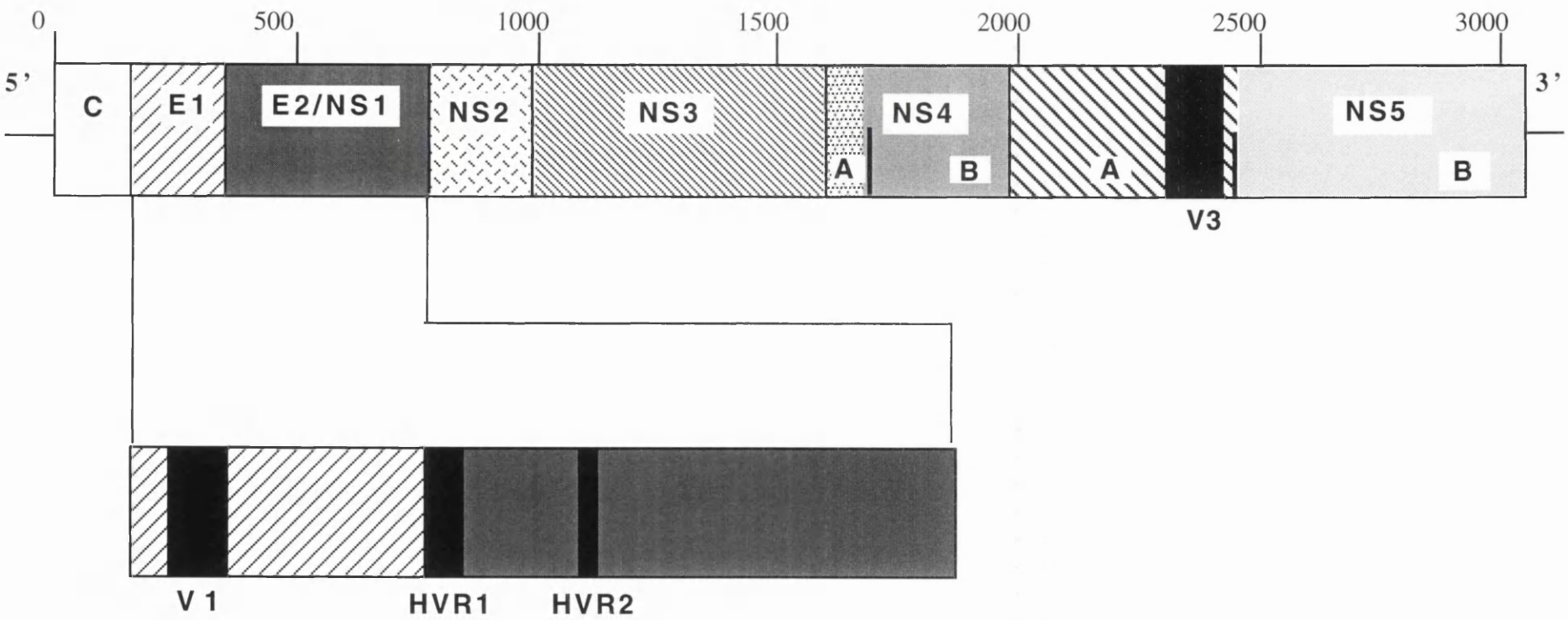


Figure 1.5 Schematic diagram showing the position of the variable and hypervariable regions of the HCV genome according to Weiner *et al.* (1991) and Inchauspe *et al.* (1991). HVR1 is also referred to as V2.

dominant in the virus population. The high mutation frequencies of RNA viruses results in a mixture of variants existing in the virus population at any given time. Evidence indicates that most RNA viruses circulate as a "quasispecies" population which evolves due to the high mutation rate per site in the RNA genome (10^{-3} to 10^{-6}) (reviewed by Holland *et al.*, 1992). The term "quasispecies" refers to the heterogeneous mixture of circulating closely related genomes. In this complex mixture the genomes compete against one another for "survival". The representation of each genome is dependent on the rate at which mutation occurs and the relative "fitness" of that mutation in the viral genome population. Thus, each quasispecies contains a master or predominant sequence (the most frequently represented) and a large spectrum of closely related variants. Circulation of RNA viruses as a population of different but related variants confers a selective advantage. It allows the virus to adapt to host selection pressures by rapid selection of a variant (mutation) with better fitness resulting in a change in the distribution of variants within the quasispecies. The quasispecies will remain in stable equilibrium if conditions remain unchanged. The high degree of sequence variation displayed by HCV, makes it necessary to discuss HCV isolates in terms of quasispecies (Martell *et al.*, 1992).

1.5.6 Sequence variation within an individual

An important feature of HCV infection is that genetic variation exists not only between isolates from different individuals, but between isolates from the same individual (Ogata *et al.*, 1991; Abe *et al.*, 1992a; Okamoto *et al.*, 1992b; Tanaka *et al.*, 1992). The HCV isolate obtained during the acute phase of infection from an experimentally infected chimpanzee was found to differ in 111 of the 9412 nucleotides positions (1.18%) from the isolate obtained during the chronic phase 8.2 years later (Okamoto *et al.*, 1992b). The substitution rate of this isolate was estimated to be about 1.44×10^{-3} base substitutions per site per year. The observed sequence heterogeneity observed in the sample taken during the acute phase of infection can affect the estimation of the rate

of sequence change. Ogata *et al.* (1991) estimated the substitution rate of the HCV genome isolated from a human carrier to be approximately 1.92 x 10^{-3} base substitutions per site per year (over a 13 year period). However, only 50% of the viral genome was analysed. Both groups found amino acids changes tended to cluster within HVR1. The substitution rate of HVR1 has been estimated to be about 4.3 x 10^{-3} base substitutions per site per year (Abe *et al.*, 1992a). Okamoto *et al.* (1992b) analysed both HVR1 and HVR2 at the time of infection and 3.5 and 8.2 years later. A homogenous population was found to be present within these regions at the beginning of HCV infection. When examined 3.5 years later heterogeneous populations were present, the heterogeneity of which was found to have increased when these regions were examined 4.7 years later (8.2 years after infection). The diversification of the HVR sequences, along with the overall sequence heterogeneity displayed throughout the viral genome, suggests variation develops with persistence of infection.

A notable feature of infection with HCV is the high level of chronicity, at least 50% of acutely infected individuals remain infected and subsequently develop chronic hepatitis. Multiple infection of an individual with several HCV isolates may also result in sequence diversity (Oshima *et al.*, 1991; Okada *et al.*, 1992). Up to 9.5 and 7.7% substitutions in the nucleotide and the amino acid sequence respectively, were found between cDNA clones isolated from liver tissue of an infected individual, the cDNA clones representing non-structural domains (of NS4 and NS5). This high degree of variation within a single patient suggested multiple infection (Oshima *et al.*, 1991). This suggests that the immune response to one HCV isolate (type) does not protect the host from re-infection by another isolate or alternatively, the individual was initially infected with more than one HCV isolate.

Several studies have examined sequential HVR1 sequences isolated from infected individuals (Kumar *et al.*, 1993; Kurosaki *et al.*, 1993; Sakamoto *et al.*, 1994). The emergence of E2 HVR1 variants was shown to occur during the development of chronic infection. The HCV quasispecies appears to change with the appearance of HCV isolates with different (new) HVR1 sequence replacing the predominant

population. This suggests alteration of the predominant HVR sequence is caused not only by spontaneous mutation of HCV during replication but also by positive selective forces. The high rate of sequence variability observed in HVR1 of E2 may result from humoral and/or T-cell mediated immune selection, as in the case of the V3 loop of HIV gp120 (Kato *et al.*, 1992b; Weiner *et al.*, 1992).

1.5.7 Mechanisms of variation

The mechanisms responsible for the high substitution rate of HCV are not fully understood. The replacement of one amino acid with another is likely to be dependent upon the size, charge and function of the amino acid residue (i.e. part of an antibody epitope or enzymatic site). The need to conserve secondary structures involved in viral replication or translation (i.e. the stem loop structure within the 5'UTR) may also be important (reviewed by Dusheiko & Simmonds., 1994).

The viral E1/E2 proteins are likely to be on the outside of the virus, exposing them to attack by the host immune system. Thus, the sequence variability observed in the envelope proteins may be humoral immune driven. The E2 HVR may encode the part of the envelope protein involved in cellular recognition and so, be under pressure to mutate from the host's humoral immune response. Since antibodies binding to this region of HCV would prevent it from recognising and therefore, infecting hepatocytes. Antibody-epitope binding studies reveal isolate-specific epitopes located in the E2 HVR (Weiner *et al.*, 1992). A hypothesis proposed by Weiner *et al.* suggests the envelope sequence variability alters the antigenicity of the protein allowing HCV to 'escape' the host humoral immune response. The appearance of a new sequence variant during infection would be followed by the development of antibodies which would specifically recognise and neutralise it (Weiner *et al.*, 1992; Kato *et al.*, 1993a; Taniguchi *et al.*, 1993). However, cross challenge experiments in chimpanzees have shown a lack of protective immunity. Experimentally infected chimpanzees can be

reinfected with the same HCV isolate, although, re-infection was seldom symptomatic in contrast to primary infection (Farci *et al.*, 1992).

1.6 Natural history of hepatitis C virus

1.6.1 Acute hepatitis C

Infection with HCV can be divided into two phases, acute and chronic. The incubation period between infection and hepatitis onset is on average about 6 weeks, during this incubation period HCV RNA can be detected in the serum of infected patients. Farci *et al.* (1991) reported detection of HCV RNA in serum one week after the patient was transfused with infected blood, preceding the initial elevation in ALT levels by 5 weeks. In acute resolving hepatitis C, viraemia is transient and lasts for between 15 to 38 weeks (Farci *et al.*, 1991; Abe *et al.*, 1992b). The length of time for the appearance of antibodies (Abs) to HCV is variable depending on what antibody is measured, Abs to the nucleocapsid usually appear first. The majority of patients with acute HCV infection are asymptomatic or present with mild symptoms.

1.6.2 Chronic hepatitis C

HCV infection is characterised by a high incidence of chronicity. The rate of chronicity is the same in community-acquired sporadic infection as in parentally-transmitted infection (Alter *et al.*, 1992) Chronic hepatitis C may also result in the development of hepatocellular carcinoma (HCC) (Saito *et al.*, 1990; Liang *et al.*, 1993).

Di Bisceglie *et al.* (1991) evaluated hepatitis C infection in patients who received blood during heart surgery. Acute post-transfusion (PT) hepatitis developed in 65/1070 patients (6.1%), becoming chronic in 45 (69%). Antibody to HCV was detected in 54 patients (82%) with PT-hepatitis, thirty-nine of these patients were

followed for between 1 and 24 years cirrhosis developed in 8 patients (20%) between 1.5 and 16 years after transfusion. A characteristic feature of infection with HCV is the pattern of fluctuating ALT levels. The course of HCV infection can be divided into three phases, 1) acute phase, where ALT levels are high 2) silent phase, which lasts for approximately 10 to 15 years, with ALT levels remaining near normal and 3) reactivated phase, where ALT levels begin to rise and fluctuate until severe liver disease develops (Kiyosawa *et al.*, 1994).

Chronic hepatitis follows one of two patterns; persistent viraemia (duration of at least more than one year) or intermittent viraemia (Abe *et al.*, 1992b). In patients with persistent viraemia, HCV RNA is detected throughout and can be associated with either high or near normal ALT levels. In patients with intermittent viraemia HCV RNA is detected during the acute phase becoming undetectable when ALT levels return to normal and detectable again when ALT levels increase as the disease becomes active (Farci *et al.*, 1991; Abe *et al.*, 1992b).

HCV is believed to be associated with hepatocellular carcinoma (HCC). In Japan the incidence of liver carcinoma is higher than in Europe or the USA. Kaneko *et al.* (1994) prospectively followed 158 patients with chronic hepatitis, positive for anti-HCV, for the development of HCC. Twenty-two (14%) of these patients developed HCC during the follow up period (5-17 years). A higher incidence of HCC was found in those patients with chronic active hepatitis (CAH). The development of HCC was found to be independent of age, sex, alcohol abuse etc. indicating hepatitis C infection is associated with HCC in Japan. The mechanism by which HCV causes HCC is unknown.

1.7 Epidemiology and transmission of HCV

HCV has a worldwide distribution. The prevalence of anti-HCV ranges from 0.3% in Canada and northern Europe, to 0.6% in the United States and central Europe, to 1.2% to 1.5% in southern Europe and Japan (reviewed by Alter, 1991). HCV may be transmitted by parenteral or non-parenteral routes.

Parenteral transmission: The most efficient routes of transmission are through blood and blood products, transplantation of organs from infected donors or intravenous drug abuse. The introduction of tests for the detection of HCV has reduced the frequency of HCV transmission via blood transfusion. Blood transfusion is now estimated to account for about 2.1% of cases (reviewed by Lemon & Brown, 1995). Anti-HCV is found in 50-90% of haemophiliacs and 70-92% of intravenous drug abusers. In haemodialysis patients the prevalence of anti-HCV is between 4-20%. The mode of transmission of HCV in renal dialysis units is not yet fully understood since not all haemodialysis patients positive for HCV received blood transfusions or shared equipment (reviewed by Esteban, 1993; Allander *et al.*, 1994).

Non-parenteral transmission: Sexual transmission has been reported as a possible route of infection, however evidence in support of sexual transmission is not convincing. In 50 heterosexual partners of individuals with HCV, none were positive for anti-HCV or HCV RNA (Bresters *et al.*, 1993). If HCV is spread via sexual transmission, the rate of infection is low in comparison to HBV or HIV. Mother to child transmission has been demonstrated for HBV and HIV infection. Maternal transmission of HCV has been documented (Inoue *et al.*, 1992; Ohto *et al.*, 1994). Ohto *et al.* demonstrated vertical transmission of HCV from mother to child and correlated the risk of transmission with the HCV RNA titre of the mother. A higher incidence of transmission was found in mothers with high titres of HCV RNA.

1.8 HCV: Treatment and prevention

1.8.1 Treatment

Treatment for hepatitis C infection is currently unsatisfactory. Interferon (IFN) is used at present to treat hepatitis C infection, but not all patients respond and of those that do, about 50% relapse once treatment is stopped. Response to IFN treatment appears to be determined by HCV genotype, severity of disease at time of treatment and IFN dosage. Individuals infected with genotypes 2 and 3 appear to respond better to IFN treatment than those infected with genotype 1 (Duskeiko *et al.*, 1994). In a study by Hino *et al.* (1994) 74 out of 136 (54%) patients with chronic HCV responded to interferon treatment. When the HCV genotypes of the patients were examined 34 out of 85 (40%) patients infected with genotype 1b responded to treatment compared to 22 out of 26 (85%) patients infected with genotype 2a and 7 out of 10 (70%) patients infected with genotype 2b. Ribavirin, an analogue of the nucleoside guanosine, is an alternative treatment. In a study by Di Bisceglie *et al.* (1992), ribavirin was administered over a period of 6 months, during this period the ALT level decreased in all patients but increased again after treatment was stopped. HCV-RNA was detectable throughout treatment. The advantage of ribavirin is that it can be administered orally. The use of ribavirin as a therapy for hepatitis C requires further evaluation. A pilot study by Brillanti *et al.* (1994, 1995) indicates that combination therapy of ribavirin and IFN induces a sustained response in patients who previously responded to IFN treatment, but who suffered a relapse after the treatment course was completed. Seventy-five percent of IFN relapsers showed sustained normal ALT levels and remained negative for HCV RNA 6 months after receiving combination therapy. Larger patient studies to investigate the combined effect of IFN and ribavirin to treat HCV infection are currently underway.

1.8.2 Prevention

There is no vaccine available for the prevention of HCV infection. Re-infection of an individual with the same or different HCV isolate, results in the development of a second bout of disease which complicates the development of a vaccine (Farci *et al.*, 1991), as does the high rate of persistence of the virus in infected individuals. Clear knowledge of the host's immune response to HCV infection is required for the successful development of a vaccine. One study reported the protection of chimpanzees against infection with HCV (homologous inoculum) following immunisation with recombinant derived HCV envelope glycoproteins (Choo *et al.*, 1994). Until the development of a vaccine against HCV the only ways of reducing the incidence of HCV infection at present are by careful screening of blood donations, adherence to precautions for preventing the spread of blood-borne infections in clinical settings and stopping the practice of needle-sharing amongst drug abusers.

1.9 Aims of this project

The work presented in this thesis had the following objectives :

1. To characterise the emergence of genetic variants of the hypervariable region (HVR1) in the 5' terminus of the putative envelope protein E2/NS1 of HCV and examine the possible association of periods of liver dysfunction with sequence changes in HVR1 in patients on long term dialysis, with the aim of testing the hypothesis that the sequence variants that arose in the region during the course of infection in these patients represented immune escape mutants.

2. To study the effect of serum storage conditions on the detection of HCV RNA by RT-PCR.

3. To construct control HCV RNA transcripts for quantitative RT-PCR analysis in order to investigate the relationship between the level of HCV RNA in patient sera and the emergence of HVR1 genetic variants.

Chapter 2 Materials

2.1 Oligonucleotide synthesis and purification

Acetonitrile, cyanoethy phophoramidites (dA, dG, dC and T), biotin cyanoethyphoramidite, tetrazole/acetonitrile solution, acetic anhydride in lutidine and tetrahydrofuran, N-Methylimidazole in \tetrahydrofuran, 0.1M Iodine in tetrahydrofuran, pyridine and water, 3% Trichloroacetic acid in dichlorometane and 0.2 micromol scale Controlled Pore Glass (CPG) columns (dA, dG, dC and T).	Cruachem Ltd
Acetonitrile (Bioscience grade) and ammonium hydroxide solution (high purity grade).	Fisons
Triethylamine acetate, triflucroacetic acid and oligonucleotide purification cartridges.	Applied Biosystems

2.2 Extraction of RNA

Guandium thiocyanate and ß-m ercaptoethanol.	BDH Ltd
Sarcosine, antifoam-A and transfer-RNA (sheep - fibroblast).	Sigma Chemical Company

2.3 Reverse Transcription

Moloney Murine Leukaemia Virus Reverse Transcriptase (M-MLV RT), 5x First Strand Buffer and 0.1M Dithiothreitol (DTT).	Gibco, BRL Ltd
Random hexamer pd(N)6.	Pharmacia LKB Ltd
Rnasin.	Promega Corp.

2.4 Polymerase Chain Reaction (PCR)

Taq DNA polymerase, 10x PCR Buffer and 50mM magnesium chloride.	Gibco, BRL Ltd

2.5 Electrophoresis

Agarose (catalogue number 9012-36-6).	Sigma Chemical Company
Nuseive Agarose (genetic technology grade).	FMC Bioproducts
Sequagel concentrate, diluent and buffer.	National Diagnostics
Ammonium persulphate, N, N, N'N-tetramethylenediamine (TEMED)	BioRad Laboratories Ltd
DNA molecular weight markers (VI).	BoehringerMannheim

2.6 DNA Sequencing

Dynabeads M-280 Streptavidivn.	Dynal International
Sequenase Version 2.0 DNA Sequencing kit.	United States Biochemical (USB)
a-^{35}S dATP >1000Ci/mmol (10 Ci/ml).	Amersham International PLC

2.7 Human sera

Human sera used in the following study were stored by the Virology Section of the Clinical Microbiology Department of the Western Infirmary, Glasgow. Informed consent for the use of sera taken from patients in the Western Infirmary for diagnostic assays was obtained from Dr J D Briggs and physicians of the Renal Unit. Sera was defined positive by Second generation Hepatitis C antibody ELISA (Abbott Lab.), RIBA (Chiron Corp.) and PCR analysis (5'NCR). Sera stored at -20^{o}C.

Control HCV negative serum was provided by Dr W. Carman, Institute of Virology. This serum was defined negative by antibody and PCR analysis. Repeated sera taken from the individual have remained negative for HCV antibody and RNA to date (October 1996).

2.8 SSCP

Acrylamide (electran), NN'-Methylenebisacrylamide (electran)	BDH Ltd.
Silver nitrate ($AgNO_3$)	Johnson Matthey Ltd.
Formamide, Formaldehyde	Fluka

2.9 Micellaneous Materials

Dideoxynucleotides 5'- triphosphates, ultrapure dNTP set	Pharmacia LKB Ltd
AmpliWax PCR Gem 100	Perkin Elmer Cetus
X-OMAT S Film	Kodak Ltd
3mm Chromatography paper	Whatmann Ltd
Diagnostic Assay kits for Hepatitis B markers, HIV and Hepatitis C antibody	Abbott Laboratories and Murex Diagnostics
Chiron RIBA HCV 3.0 Strip Immunoblot Assay kit	Kindly donated by Othro Diagnostic Systems Inc.
Geneclean II kit	Bio 101 Inc.

Mineral oil	Sigma Chemical Company
Biochemical grade water	BDH Ltd
Restriction enzymes and buffers.	Gibco, BRL Ltd or Boehringer Mannheim

2.10 Chemicals

All chemicals were obtained from either BDH Ltd., Boehringer Mannheim, Sigma Chemical Company. or Prolabo Ltd.

2.11 Commonly used solutions

TE buffer: 10mM Tris-HCL (pH 7.5 or 8.5), 1mM EDTA.

TAE buffer: 40mM Tris-HCL (pH8.0), 1mM EDTA.

TBE buffer: 90mM Tris-HCL, 90mM Boric acid, 1mM EDTA.

Ligase buffer: 300mM Tris-HCl (pH 7.8), 100mM $MgCl_2$, 100mM DTT and 10mM ATP.

Phosphate buffered saline (PBS): 170mM NaCl, 3.4mM KCl, 10mM Na_2HPO_4 (pH 7.2) supplemented with 6.8mM $CaCl_2$ and 4.9mM $MgCl_2$.

Sequencing buffer: 0.9M Tris-HCL, 0.9M Boric acid, 0.02M EDTA.

Solution D: 4M guanidium thiocyanate, 56mM sarcosine, 0.33% antifoam A, 50mM Tris-HCl (pH 7.5), 2mM EDTA and 0.7% β-mercaptoethanol.

STET: 0.1M NaCl, 10mM Tris-HCl (pH 8), 1mM EDTA and 5% Triton X-100.

Dynabead binding and washing buffer: 10mM Tris-HCl (pH 7.5), 1mM EDTA, 2.0MmNaCl (final concentration 1mM).

Gel loading buffer: 50% sucrose, 0.1M EDTA, 0.05% bromophenol blue

Chapter 3 Methods

3.1 Oligonucleotides

3.1.1 Oligonucleotide synthesis

Oligonucleotides were made by cyanoethyl phosphoramidite chemistry, using a Cruachem PS250 synthesiser. This involved the sequential addition of a protected form of cyanoethyl phosphoramidite monomers. Each protected monomer has a labile dimethoxytrityl protecting group which was removed by the detritylating agent, 3% trichloracetic acid in dichloromethane, following the addition of each new base. When synthesising biotin-labelled oligonucleotides, the last base to be added is a biotin-cyanoethyl phosphoramidite monomer, allowing the addition of a biotin molecule at the 5' terminus of the oligonucleotide. The growing DNA molecule is covalently linked to an inert support, Controlled Pore Glass (CPG), which is contained in a reaction column attached to the machine.

3.1.2 Oligonucleotide cleavage and deprotection

After synthesis, the oligonucleotide was removed from the CPG column by cleavage with concentrated ammonia solution. One and a half millilitres of concentrated ammonia solution was taken up in a disposable 5ml syringe which was attached to one end of the column. The ammonia solution was gently pushed into the column until it could be seen at the opposite end, to which a second 5ml disposable syringe was attached. The column was then left for 20 minutes at room temperature, after which a further column volume of ammonia solution was pushed into the column and again the column was left for 20 minutes at room temperature. This action was repeated until all the ammonia solution was passed through the column. After the

final incubation period, the ammonia solution was pushed back and forth through the column several times to mix the solution. The solution was then pulled into one syringe and transferred to a screw cap eppendorf tube.

After cleavage of the oligonucleotide from the CPG column, deprotection of the oligonucleotide to remove the amide protecting groups took place. The microtube containing the cleaved oligonucleotide was placed in a water bath, heated to 55°C for a minimum period of 5 hours. The eppendorf tube was then removed and placed in a refrigerator at 4°C for 10-15 minutes. The deprotected oligonucleotide was then stored at -20°C.

3.1.3 Recovery of deprotected oligonucleotide

The ammonia solution containing the deprotected oligonucleotide was divided between three eppendorf tubes and evaporated in a centrifugal evaporator (Savant "Speedivac" concentrator). The dried oligonucleotide was then resuspended in 100µl of deionized water for immediate use or stored at -20°C. To avoid contamination, oligonucleotides were dried down in a different "Speedivac" to that used for RNA preparations and plasmids.

3.1.4 Purification of oligonucleotides

Oligonucleotides used were either purified or unpurified. Those to be purified by Applied Biosystems Oligonucleotide Purification cartridges were synthesised with their dimethyoxytrityl protecting group still attached to the final base. Deprotected oligonucleotides to be purified were not subjected to evaporation (described in section 3.1.3) instead, after deprotection a $^1/_3$ volume of deionized water was added to the oligonucleotide. A disposable 5ml syringe was attached to the purification cartridge. The cartridge was then flushed with 5ml of HPLC grade acetronitrile followed by 5ml of 2M triethylamine acetate. The syringe was removed from the

cartridge prior to removing the syringe plunger. The deprotected oligonucleotide was then slowly pushed through the cartridge and collected, the flow rate being kept to one or two drops per second. The eluted fraction was then once more pushed slowly through the cartridge. The cartridge was then washed three times with 5ml of 1.5M ammonium hydroxide, followed twice with 5ml of deionized water. The bound oligonucleotide was then detritylated with 5ml of 2% trifluroacetic acid. One millilitre of trifluoracetic acid was pushed through the cartridge and 5 minutes were allowed to elapse before the remaining acid was pushed through. Once more the cartridge was washed twice with 5ml of deionized water. The purified, detritylated oligonucleotide was eluted by slowly washing the cartridge with 1ml of 20% acetonitrile. The optical density at 260nm was then determined.

3.2 Quantitation of nucleic acids

The concentration of nucleic acids was determined by the optical density (O.D) at 260nm using a Beckman DU-62 spectrophotometer.

A value of 1 at OD_{260} is equivalent to: 50mg/ml for double stranded DNA.
40mg/ml for RNA or single stranded DNA.
20mg/ml for oligonucleotides.

The purity of a nucleic acid preparation was assessed from the ratio of absorbance at 260nm/280nm, a pure DNA preparation having a ratio of 1.8 and a pure RNA preparation having a ratio of 2.

The molar concentration of oligonucleotides was determined directly as $O.D_{260}/E_M$ (molar extinction coefficient).

3.3 Preparation of phenol

Phenol was melted at 60°C and aliquoted in 5 or 10ml volumes into universals which were stored at -20°C. When required, an aliquot of phenol was melted at 60°C and an equal volume of deionized water or 1x TE buffer (pH 7.5) added at room temperature and thoroughly mixed with the phenol. The mixture was allowed to settle into two phases. The water saturated or TE saturated phenol solution was stored at 4°C for up to two months.

3.4 Human sera

All human sera were stored at -20°C in small sealed aliquots. When required, the sera were thawed at 4°C. A fresh aliquot of serum was used in each experiment.

3.5 Extraction of RNA from human sera

The method for extraction of RNA from human sera is adapted from the acid guanidium thiocyanate-phenol-chloroform extraction technique devised by Chomczynski and Sacchi (1987).

One hundred microlitres of human serum was gently mixed with 300µl of denaturing solution D and 300µl of water-saturated phenol, to which 1µg/ml of purified carrier transfer RNA (sheep fibroblast) was added. Fifty microlitres of chloroform was added and the mixture incubated on ice for 15 minutes. After centrifugation at 1400g at 4°C for 15 minutes the aqueous phase was removed to a fresh tube and re-extracted with an equal volume of chloroform. The RNA was precipitated with an equal volume of isopropanol at -20°C overnight. RNA was pelleted by centrifugation at 14000g at 4°C for 15-20 minutes. The RNA pellet was

then washed once with 70% ethanol, and centrifuged at 12000g at 4°C for 5 minutes. The RNA pellet was allowed to air dry at room temperature or dried under vacuum for several minutes, before being resuspended in 10µl of deionized water. RNA was stored at -70°C.

3.6 Amplification of RNA

All reagents for reverse transcription and polymerase chain reaction were aliquoted before use and filtered/plugged tips used at all times.

3.6.1 Reverse transcription (RT)

First strand complementary DNA (cDNA) was synthesised from 3µl extracted RNA using 100units of Moloney Murine Leukaemia reverse transciptase (M-MLV RT) in 50mM Tris-HCL (pH 8.3), 75mM KCl, 3mM $MgCl_2$, 10mM DTT, 1mM each of dGTP, dATP, dTTP and dCTP, 40pmol of random hexamers or specific antisense primer and 1unit of RNAsin in a final volume of 20µl. Incubation then took place at 25°C for 10 minutes, followed by 37°C for 55 minutes and 95°C for 5 minutes. After incubation was complete, products reaction were stored at 4°C.

3.6.2 Polymerase chain reaction (PCR)

PCR is a rapid procedure for the *in vitro* enzymatic amplification of a target segment of DNA (Saiki *et al.*, 1985; Mullis & Faloona, 1987). Two oligonucleotide primers, complementary to and flanking the two strands of the target DNA are synthesised, the target DNA denatured and the primers annealed to opposite strands of the DNA, oriented with their 3' ends facing one other. DNA synthesis then occurs by the extension of the primers at their 3' ends by a thermostable DNA polymerase.

Successive cycles of denaturation, primer annealing and extension result in an exponential accumulation of the target DNA.

The technique of 'nested' PCR was used to amplify HCV DNA from HCV cDNA. This involves two rounds of PCR, with product from the first round of amplification being subjected to a second round. The set of oligonucleotide primers used for the second round of amplification is internal to the set used in the first round. The use of nested PCR increases the specificity and amount of HCV DNA amplified. In the first round of PCR the reaction was carried out using 0.5 units of *Taq* DNA polymerase in 20mM Tris-HCl (pH 8.4), 50mM KCl, 1.5mM $MgCl_2$, 0.2mM each of dGTP, dATP, dTTP and dCTP and 40pmol each of the outer pair nested primers in a final volume of 20µl. The reaction mix was then overlaid with 50µl of mineral oil. Amplification was performed over 35 cycles with an initial denaturation step of 95°C for 4 minutes followed by: 94°C/1min.

55°C/2mins.

72°C/3mins.

One microlitre of first round product was then amplified in a second round of PCR, using the same reaction conditions described above except 40pmols each of the inner nested primers was used. Amplification was over 25 cycles. The number of cycles, reaction times and annealing temperature were varied (see results). Ten microlitres of PCR product was run on a 2% agarose gel containing ethidium bromide and visualised under shortwave ultraviolet light (UV).

3.7 Analytical agarose gel electrophoresis of DNA

Electrophoresis of DNA was performed on horizontal slab gel (11 x 14 x 0.5 cm) of 2% (w/v) agarose in 1x TBE or 1x TAE containing 0.1mg/ml ethidium bromide (EtBr). DNA samples were loaded onto the gel in loading buffer (50% sucrose, 0.1M EDTA, 0.05% (w/v) bromophenol blue). Electrophoresis was performed at

100-150 volts in 1x TBE or 1x TAE buffer, except for low melting agarose gels where electrophoresis was performed at 50-80 volts, until the dye front had migrated the required distance. The gel was then visualised under shortwave UV light and photographed.

3.8 Purification of DNA fragments from agarose gels

3.8.1 Extraction of DNA from a low melting agarose gel

Electrophoresis of the DNA of interest was performed using a 2% Nusieve agarose 1x TBE gel at 50-80 volts. The DNA bands were visualised under longwave UV light and the required DNA band excised from the gel. The excised DNA/gel slice was incubated at 60-65°C for 10 minutes until the agarose had fully melted. Two hundred microlitres of 1x TE buffer (pH 7.5) was added and the mixture vortexed. A $^3/_4$ volume of TE saturated phenol (phenol saturated with 1x TE buffer, pH 7.5, see section 3.3) was then added to the gel slice. The mixture was vortexed and the aqueous phase recovered by centrifugation (microcentrifuge, high speed) for 5 minutes and re-extracted twice with an equal volume of 1x TE (pH 7.5) saturated phenol. A further two extractions of the aqueous phase were performed, with equal volumes of chloroform. The DNA was recovered from the aqueous phase by precipitation, by adding a $^1/_{10}$ volume of 3M sodium acetate and 2.5 volumes of ethanol and placing the tube in dry ice/ethanol for 5-10 minutes. The precipitated DNA was then pelleted by centrifugation for 10 minutes, washed with 750µl of 70% ethanol, dried under vacuum and resuspended in the appropriate volume of deionized water. The DNA was stored at -20°C.

3.8.2 Silica matrix adsorption

A commercial kit, "Geneclean", which contains a specially formulated silica matrix called 'Glassmilk', which binds single and double stranded DNA in the presence of sodium iodide (NaI), was used for recovering DNA fragments from agarose gels. An agarose gel slice (in 1x TAE buffer) containing the DNA of interest was mixed with 3 volumes of NaI solution and incubated at 55°C for 5 minutes, until the agarose was completely dissolved. Glassmilk was added at a concentration of 5ml/5mg of DNA and the mixture incubated at 4°C for 10 minutes to allow the DNA to bind. Following a short spin (about 10secs.), the pelleted glassmilk/DNA was washed 3 times with 10-50 volumes of ice-cold 'New Wash' solution (which contains Tris-HCl, EDTA , NaCl and ethanol) by repeated suspension and pelleting. The DNA was eluted from the glassmilk by resuspending the pelleted glassmilk/DNA in half the final volume of deionized water required and incubating at 55°C for 5 minutes. The glassmilk was pelleted by centrifugation (microcentrifuge) and the eluted DNA removed to a fresh tube. A second elution step was then performed. The eluted DNA was stored at -20°C.

3.9 Preparation of DNA for sequencing

3.9.1 Preparation of double stranded DNA templates derived from PCR products

Eighty to one hundred microlitres of second round PCR product was run on a 2% low melting point TBE buffered agarose gel, stained with EtBr and the DNA band of interest excised. DNA was then extracted from the agarose gel slice as described in section 3.8.1 and resuspended in 15-20µl of deionized water. The PCR products were

cleaned prior to sequencing to remove any excess primer and non-specific amplified products.

3.9.2 Generation of single stranded DNA templates derived from PCR products

'Dynabeads' M-280 Streptavidin (Dynal Int.), magnetic beads with streptavidin covalently attached to the bead surface, were used to generate single stranded DNA from PCR products. Essential for the use of Dynabeads is the biotinylation of the PCR oligonucleotide primers. Two 50µl volume second round PCR reactions were performed; in the first PCR reaction, biotinylated sense and unbiotinylated antisense primers were used and in the second, unbiotinylated sense and biotinylated antisense primers were used. PCR reactions were performed using the conditions described in section 3.6.2.

A: Preparation of Dynabeads M-280 Streptavidin.

The Dynabeads M-280 were resuspended by pipetting back and forth, the beads were never vortexed or centrifuged. The appropriate volume of Dynabeads M-280, 20µl per PCR reaction (at concentration of 10mg/ml) was placed into a microtube. The tube was then placed in a 'Dynal' (MPC) magnetic stand for at least 30 seconds and the supernatant removed by aspiration. The tube was then removed and an equal volume of PBS containing 0.1% BSA was added and the beads gently resuspended. Once more the tube was placed in the magnet and the supernatant removed. The beads were then resuspended in 2 volumes of 'Binding and Washing' (B&W) buffer, to a final concentration of 5mg/ml.

B: Immobilisation.

Forty microlitres of pre-washed beads were added to 40µl of amplified PCR product and incubated for 15 minutes at room temperature. The beads were kept resuspended during this incubation period by gently rocking the tube. The tube was then placed in the magnet and the supernatant removed. The beads were then washed with 40µl of B&W buffer.

C: Denaturation.

The tube containing the immobilised product was placed in the magnet and the supernatant removed. Eight microlitres of freshly prepared 0.1M NaOH solution was added to the beads, which were resuspended and incubated for 10 minutes at room temperature. Once more the supernatant was removed by placing the tube containing the beads in the magnet. The beads were washed once with 50µl of 0.1M NaOH, followed by 50µl of B&W buffer and then 50µl of 1x TE buffer (pH 7.5). The beads were finally resuspended in 10-20µl of 1x TE buffer (pH 7.5).

3.10 Dideoxynucleotide sequencing

The dideoxynucleotide chain termination method of Sanger *et al.* (1977), using the Sequenase Version 2.0 DNA Sequencing kit (United States Biochemical) according to a modified version of the manufacturers protocol, was used to determine DNA sequence. In the original procedure of Sanger *et al.* (1977), primer extension was catalysed by the Klenow fragment of *E.coli* DNA ploymerase I. The Sequenase Version 2.0 kit uses a genetic variant of bacteriophage T7 DNA polymerase (Tabor & Richardson, 1989). A list of solutions provided in the kit is provided in Table 3.1.

Table 3.1 List of reagents contained in the Sequenase Version 2.0 DNA Sequencing Kit.

Sequenase Buffer (5X concentrate)	200mM Tris-HCl (pH7.5), 100mM $MgCl_2$, 20mM NaCl
Dithiothreitol (DTT) solution	0.1M
Labelling Mix (dGTP) (5X concentrate)	7.5mM each dGTP, dTTP, dCTP
ddGTP Termination Mix	80mM each dGTP, dATP, dTTP, dCTP, 8mM ddGTP, 50mM NaCl
ddATP Termination Mix	80mM each dGTP, dATP, dTTP, dCTP, 8mM ddATP, 50mM NaCl
ddTTP Termination Mix	80mM each of dGTP, dATP, dTTP, dCTP, 8mM ddTTP, 50mM NaCl
ddCTP Termination Mix	80mM each dGTP, dATP, dTTP, dCTP, 8mM ddCTP, 50mM NaCl
Sequence Extension Mix (dGTP)	180mM each dGTP, dATP, dTTP, dCTP, 50mM NaCl
Stop Solution	95% Formamide, 20mM EDTA 0.05% Bromophenol Blue, 0.05% Xylene Cyanol FF
Enzyme Dilution Buffer	10mM Tris-HCl (pH7.5), 5m DTT, 0.5mg/ml BSA
Sequenase Version 2.0 T7 DNA Polymerase	13 units/ml in 20mM KPO_4 (pH7.4), 1mM DTT, 0.1mM EDTA, 50% Glycerol

A: Annealing reaction.

Seven microlitres of DNA template, either single or doubled stranded DNA, was mixed with 1µl of the appropriate primer (5pmol/µl). For single stranded DNA the annealing reaction was performed at 37°C for 30 minutes. For double stranded DNA, the annealing reaction was preceded by incubation at 95°C for 2 minutes, 70°C for 3 minutes and 95°C for 45 seconds. Annealing was carried out at 37°C for 15 minutes. Incubation was briefly interrupted after 2.5 minutes at 70°C to centrifuge the tubes to ensure all the template/primer mixture was subjected to annealing at 37°C. After annealing was complete the tubes were transferred directly to ice.

B: Labelling reaction.

Before this reaction was carried out, the tubes for the termination reaction were prepared (see section C). The labelling mix, supplied in the Sequenase Version 2.0 kit (Table 3.1), was diluted 10-fold with deionized water. To the tube containing the annealed template/primer, 2µl of reaction buffer, 2µl of 10-fold diluted labelling mix, 1µl of 0.1M DTT and 1µl of a-^{35}S dATP (10mCi) were added and the tube briefly centrifuged to mix contents. Two microlitres of an 8-fold dilution of Sequenase Version 2.0 T7 DNA polymerase was added and the tube incubated at 25°C for 2 minutes.

C: Termination reaction.

Two and a half microlitres (2.5µl) of each termination mix for G, A, T and C tracks (Table 3.1) were added to individual tubes (one set per template/primer reaction). The tubes were pre-incubated at 37°C for at least 1 minute prior to use. Three and a half microlitres (3.5µl) of the completed labelling reaction was added and mixed with each of the four termination mixes. The tubes were incubated at 37°C for 3-5

minutes before termination of the reaction by the addition of 4µl of stop solution (Table 3.1)

The above termination protocol was used when sequencing PCR products 200-500 nucleotides in length. When PCR products greater than 500 nucleotides were to be sequenced, the termination mixes were adjusted. A half microlitre (0.5µl) of dGTP sequence extending mix (Table 3.1) was added to 2µl of each termination mix and the incubation time for the termination reaction was extended to 5-8 minutes.

D: Sequencing gel.

The 'Sequagel' system (National Diagnostics Ltd.) is a 'ready to use' system for the preparation of polyacrylamide gels. For DNA sequence analysis a 5% sequencing gel, containing a final concentration of 8.3M urea in 1x TBE buffer was made from mixing 30ml of Sequagel concentrate with 70ml of Sequagel diluent and 10ml of Sequagel buffer. Polymerisation of the gel was catalysed by adding 800µl of 10% (w/v) ammonium persulphate and 40µl of TEMED for every 100ml of gel casting solution. The BRL model S2 electrophoresis apparatus was used to run the gels. Sequencing gels (30 x 40 x 0.4 cm) were prepared as described by Maniatis *et al.* (1992). The sequencing gel was pre-run at 70 watts for 30-60 minutes in 1x sequencing buffer. Prior to loading onto the gel, single stranded DNA sequencing reactions were heated at 72°C for 3 minutes to elute the DNA from the Dynabeads. Double stranded DNA sequencing reactions were heated at 90°C for 3 minutes to denature the DNA. Two and a half to three and a half microlitres (2.5-3.5µl) of each termination reaction was loaded into the appropriate well and electrophoresis was performed at 70 watts until the dye fronts had migrated the required distance. After electrophoresis was complete, one of the glass plates was carefully removed and the gel fixed in a solution of 10% methanol/10% acetic acid(glacial) for 10-15 minutes. The gel was then transferred to a supporting sheet of Whatman 3MM filter paper,

covered with 'cling film' and dried for 1-2 hours using a vacuum gel drier. The cling film was removed and the gel exposed to X-OMAT S film at room temperature.

3.11 Recombinant Immunoblot Assay (RIBA)

A commercial *in vitro* qualitative enzyme immunoassay for the detection of antibody to HCV, the Chiron RIBA HCV 3.0 Strip Immunoblot Assay, kindly donated by Ortho Diagnostics Systems Inc., was used to test patient serum samples of interest for the presence of HCV antibodies. In the above mentioned commercial kit, HCV recombinant antigens and peptides are immobilised as individual bands on a nitrocellulose strip. Solutions provided in the kit are listed in Table 3.2. During incubation of the strip with a serum sample, antibodies to HCV, if present, will react with the corresponding recombinant antigen and/or peptide band on the strip.

Before beginning the assay, the kit was removed from storage at 4°C and allowed to come to room temperature. The desired number of nitrocellulose strips were removed and 1ml of specimen diluent was added to each strip contained in own individual tube. Twenty microlitres of the appropriate serum sample or control was added. The tubes were put into a rack and rocked for 4 hours at room temperature. The specimen diluent was aspirated and 1ml of fresh specimen diluent added. The tubes were returned to the rocker for a further 30 minutes. Again the diluent was aspirated and 1ml of wash buffer added. The wash buffer and the all the strips were placed in a wash vessel, containing 30ml of wash buffer. The vessel was agitated briefly (5-10mins) and the buffer decanted. A further wash was performed before 1ml of conjugate per strip was added. The wash vessel was then placed on a rotary shaker at 110 ± 5 rpm for 10 minutes at room temperature. The conjugate was removed and the strips washed three times with wash buffer. 1ml of substrate per strip was added. The strips were agitated, on the rotary shaker for 15 minutes, again at room temperature. The substrate was removed by aspiration and the strips washed

Table 3.2 List of reagents contained in the Chiron RIBA HCV 3.0 Immunoblot Assay.

Specimen Diluent detergent	Phosphate buffered saline with protein stabiliers and
Conjugate	Goat anti-human IgG (heavy and light chain specific) conjugated to horseradish peroxide with protein stabilizers
Substrate Solution	4-chloro-1-naphthol in methanol
Substrate Buffer	Buffered hydrogen peroxide
Wash Buffer Concentrate	Buffered detergent solution
Positive Control	Source: Heat and psoralen treated human serum containing anti-HCV and nonreactive for HBsAg and anti-HIV.
Negative Control	Source: Human serum nonreactive for antibody to HCV or HIV and to HBsAg.

twice with deionized water. Using forceps, the strips were transferred to absorbent paper and allowed to air dry in the dark at room temperature. The strips were interpreted within three hours of them drying (see results).

3.12 Restriction enzyme digestion of DNA

DNA was digested with the appropriate restriction enzyme (normally 1 unit/μg DNA) in the appropriate restriction enzyme buffer, according to manufacturers instructions. For the analysis of small scale plasmid DNA preparations, 3μl of DNA was digested in a 20μl volume reaction containing 25μg/ml RNase A at 37°C for 2 hours. DNA fragments were separated by gel electrophoresis.

3.13 Single stranded conformation polymorphism (SSCP)

Single stranded conformation polymorphism (SSCP) analysis is a method for detecting single base differences in PCR products, by detection of mobility differences of single stranded (ss) DNA in non-denaturing polyacrylamide gels during electrophoresis. It was thought that by employing this technique we would have a relatively simple detection method for sequence changes in the HVR of HCV from serum samples from a patient over time, detecting minor variants as well as the predominant variant, thus decreasing the number of samples needing to be sequenced.

PCR reactions on the samples of interest were performed. The PCR products were denatured prior to loading onto the polyacrylamide either by heat denaturation or alkali denaturation. 1μl of loading buffer (0.5% bromophenol blue and 0.5% xylene cyanol FF in formamide) was added to 10μl of PCR product and the products heat denatured at 95°C for 5 minutes and snap-cooled on dry ice/ethanol.

Alternatively, 1µl of 0.5M NaOH and 1µl of 10mM EDTA was added to 10 µl of PCR product and heated to 42°C for 5 minutes before adding 1µl of loading buffer. The PCR products were then loaded onto a non-denaturing polyacrylamide gel, of varying composition (Table 3.3). The polyacrylamide gels were poured using vertical gel apparatus (Protean II kit from Biorad) and run at 15-30 V cm^{-1} (350-600V) for 1.5-2.5 hours and maintained at 20-22°C by circulating cold water. The gels were stained with either by EtBr or silver. Two methods for silver staining the gels were used (Table 3.4).

SSCP was also performed using ^{32}P labelled PCR products. 0.2mM of γ-^{32}P dATP instead of 'cold' dATP was used in the PCR reaction, 2µl of radioactive PCR product was mixed with 8µl of 'cold' PCR product and denatured, by one of the methods described above. The products were run on a gel along with a radiolabelled 100bp marker (GIBCO, BRL) which was labelled according to manufacturers instructions. The gel was dried and exposed to X-OMAT S film at -70°C.

Table 3.3 Composition of polyacrylamide gels used for SSCP analysis.

Gel	% gel (w/v)	Ratio polyacrylamide: bisacrylamide	Buffer	Glycerol (w/v)
1. Stacking gel				
i) Lower gel	6	49:1	0.5x TBE	5
ii)Upper gel*	8	19:1	0.2x TBE	NA
2.	7	49:1	0.5X TBE	5
3.	5	4:1	0.5X TBE	5

NA. Not applicable
* 75% formamide was added to this gel.

Table 3.4 Silver staining methods

Method 1.

Step 1. Fix gel for 2-3hrs in soln. I (25% ethanol, 10% acetic acid).

Step 2. Discard soln. I and leave gel overnight in soln. II (10% ethanol, 0.5% acetic acid).

Step 3. Discard soln. II and add soln. III (0.19% silver nitrate, 99.81%). Leave for 2hrs.

Step 4. Discard soln. III and wash gel with deionized water.

Step 5. Add soln. IV (3% sodium hydroxide, 0.009% sodium borohydride, 0.76% formaldehyd). Shake gel gently over a light box until bands appear. Approx. 5-15mins.

Step 6. Discard soln. IV and fix gel in soln. V (5% acetic acid). Fix for 1hr.

Method 2.

Step 1. Fix gel in soln. I (7.5% acetic acid) for 5-30mins.

Step 2. Wash gel three times with deionized water. 2-5mins.

Step 3. Add soln. II (0.15% silver nitrate, 0.056% formaldehyde) for 10-60mins.

Step 4. Discard soln. II and wash gel with deionized water.

Step 5. Add soln. III (3% sodium carbonate, 0.056% formaldehyde, 0.04% sodium thiosulphate) and gently shake gel over light box until bands appear. Approx. 2-10mins.

Step 6. Discard soln. III and fix gel in soln. IV (7.5% acetic acid) for 5mins.

Chapter 4 Results: Section 1

Sequence evolution of the HVR1 of the putative envelope gene E2 of hepatitis C virus in patients on haemodialysis.

The work presented in this chapter was undertaken to characterise the emergence of sequence variants of the hypervariable region (HVR1) at the 5' end of the gene coding for the putative envelope protein E2 of HCV and to examine the possible association of periods of liver dysfunction with sequence changes in HVR1 in patients on long term dialysis. The E2 HVR1 is defined in the work presented here as between nucleotide positions 1150 and 1233 and amino acid positions 384 and 411 (numbering according to Choo *et al.*, 1991).

A characteristic of the HCV genome is sequence heterogeneity. Comparisons of HCV isolates from different infected individuals showed sequence variation between isolates (Choo *et al.*, 1991; Kato *et al.*, 1990; Takamizawa *et al.*, 1991). Previous studies have shown that the HCV genome is heterogeneous not only between different infected individuals but also within the same individual (Ogata *et al.*, 1991; Okamoto *et al.*, 1992b). The HVR1 domain in particular undergoes diversification of its sequence with time in an individual (Kato *et al.*, 1992a, 1994; Kurosaki *et al.*, 1993; Sakamoto *et al.*, 1994). The high substitution rate observed in the viral envelope proteins, especially the HVR1, may be driven by the immune system since the envelope proteins, located on the surface of the virion, are likely to be the principal target of the host's humoral immune response. The emergence of viral escape variants has, therefore, been proposed as a possible mechanism by which HCV is able to maintain a persistent infection.

A characteristic clinical feature of hepatitis C is the episodic fluctuating pattern of alanine transaminase (ALT) and aspartate transaminase (AST) levels. Periods of elevated ALTs and ASTs, lasting from weeks to months, are interspersed with periods

of normal or near normal levels. Since the levels of the two enzymes generally move in parallel, in this work only the ALT levels have been studied in detail. ALT levels are used as an indicator of liver damage because ALT is primarily found in hepatocytes rather than other cell types. When hepatocytes are damaged, ALT is released into the bloodstream.

The emergence of a new sequence variant of HCV might be expected to be related to a period of liver cell damage in one of two ways, depending on whether damage to hepatocytes is due to a direct cytopathic action of HCV or to the destruction of HCV infected hepatocytes by the host's immune system. If an elevation in ALT was the result of a direct cytopathic action of HCV, the virus would infect hepatocytes and viral replication would take place resulting in hepatocyte damage and the release of infectious virus. This in turn would lead to an increase in the number of hepatocytes infected with HCV. The net result of this process would be liver damage, seen as an elevation in ALT. The appearance of antibodies which could neutralise the infectious virus before it is able to infect an hepatocyte would result in a period of lower viral replication and less liver damage. With the emergence of a new sequence variant which is not recognised by existing antibody there would be an increase in the number of hepatocytes being infected, resulting in an increase in liver damage. In this model the level of viral replication would be high during an elevation in ALT and the new sequence variant would emerge before the ALT peak.

Alternatively, liver damage might result from the destruction of infected hepatocytes by the host's immune system. HCV infected hepatocytes which are recognised by the host's immune system would be destroyed, resulting in an elevation in ALT and fewer infected hepatocytes. Hence, there would be a decrease in the level of viral replication. This would give a selective advantage to other HCV variants which caused the infected hepatocytes to display different epitopes leading to the emergence of a new predominant variant. This leads to an increase in the number of infected hepatocytes and so, an increase in the level of viral replication. In this model, the level

of viral replication would be low during an elevation in ALT and a new sequence variant would appear some time after the ALT peak.

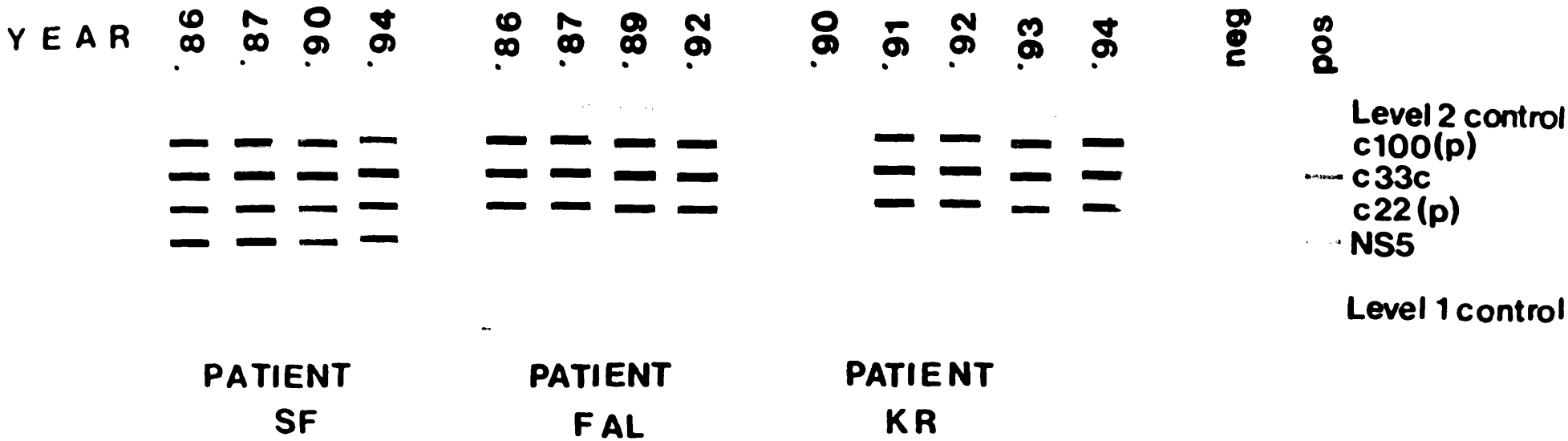
Although elevations of ALT are not uncommon in patients on long-term renal dialysis, HCV infected patients are more likely to have elevated ALT levels than uninfected patients (McIntyre *et al.*, 1994). To characterise the emergence of genetic variants and examine their possible association with periods of liver dysfunction, the HVR1 domain was studied retrospectively in three renal haemodialysis patients.

4.1 Study patients

Three haemodialysis patients chronically infected with hepatitis C were studied. Sera from all three patients were positive by second generation ELISA (Abbott Lab.), confirmed by third generation RIBA (Ortho Diagnostic Systems Inc.) and for HCV RNA by RT-PCR using primers derived from the 5'NCR. Previous work in the laboratory had shown changes in the RIBA pattern over time in renal dialysis patients from either indeterminate result to a positive result or positive to an indeterminate (McIntyre *et al.*, 1994). However, the results of the third generation RIBA for these three patients (Figure 4.1) showed no loss of antibody to any of the expressed antigens over time.

Patient KR :- a 53 year old male (when the first sample was examined). HCV RNA was detected by RT-PCR using primers amplifying the 5'NCR of the genome in July 1990 but no antibody was detected by either ELISA or RIBA until October 1990. He had been on regular haemodialysis since 1989 after two failed transplants and had a third kidney transplant in April 1990 which failed immediately. The source of the virus infecting this patient is not known. Sera from the kidney donor in 1990 and donors of blood cross matched for transfusion between April 1990 and the first RT-PCR positive result were all PCR and antibody negative for HCV.

Figure 4.1 Detection of anti-HCV antibody in serial serum samples from patients SF, FAL and KR using the Chiron RIBA HCV 3.0 Strip Immunoblot Assay. Bands are scored according to the band intensity compared to the Level 1 and 2 controls. Scoring procedure: (1) no visible band on the strip indicates a negative (-) rating; (2) visible band intensity less then Level 1 control indicates are positive/negative rating (+/-); (3) visible band intensity equal to level 1 control a 1+ rating; (4) visible band intensity greater than Level 1 and less than Level 2 control a 2+ rating; (5) visible band intensity equal to Level 2 control a 3+ rating; (6) visible band intensity greater than Level 2 control a 4+ rating. A sample is interpreted as nonreactive if no bands of 1+ or greater reactivity are present; reactive if bands of 1+ or greater reactivity to any two HCV recombinant antigens are present; indeterminate if bands of 1+ or greater reactivity are present but pattern does not meet criteria for reactive antigen band pattern. All patient samples apart from patient KR 1990 (pre-October) sample contained antibody to HCV. Patients FAL and KR had no antibody response to the NS5 antigen.


YEAR
.86
.87
.90
.94
.86
.87
.89
.92
.90
.91
.92
.93
.94
neg
pos
Level 2 control
c100(p)
c33c
c22(p)
NS5
Level 1 control
PATIENT SF
PATIENT FAL
PATIENT KR

Patient SF :- a 43 year old female (when the first sample was examined). The first available serum taken in December 1982 was antibody positive for HCV. Both the time and source of infection for this patient are unknown. SF has been on haemodialysis since 1976 and has had two transplants which both failed.

Patient FAL :- a 43 year old female (when the first sample was examined). The first available serum taken in January 1986 was antibody positive for HCV. Both the time and source of infection are unknown. FAL was on regular haemodialysis from 1976 until she was successfully transplanted in 1992.

Patients KR and FAL are infected with HCV genotype 1a and patient SF with HCV genotype 1b. This information was kindly provided by Dr June Munro and was determined by sequence comparison of the NS5 "V3" region (amino acids 2356 to 2379) with reference sequences (Kato *et al.*, 1990; Choo *et al.*, 1991; Inchauspe *et al.*, 1991; Takamizawa *et al.*, 1991) as part of a molecular epidemiological study.

The serum samples were taken for ordinary diagnostic purposes and not specifically for this study. Because of this, it is likely that the samples, especially the earlier ones, had been thawed several times and may have been subject to adverse conditions e.g. left at room temperature which leads to a reduction in RNA titre (Busch *et al.*, 1992) or exposure to bacterial and fungal contamination. To address this, once the study was underway, aliquots of the sera were stored at -20°C until needed, to reduce the number of freeze-thaw cycles and chance of becoming contaminated.

Serum samples taken at regular intervals were available for the study from 1990 to 1994 for patient KR, from 1986 to 1994 for patient SF and from 1986 to 1992 for patient FAL. Thirteen sera were obtained from patient KR over a three year period from 1991-1994, eleven from patient SF over a six year period from 1988-1994 and eleven sera from patient FAL over a four year period from 1988-1992, after a successful transplant in (August) 1992 no further samples were taken until 1994 when a further

sample was obtained for the purpose of this study, making twelve in total. These samples are listed in Tables 4.1 to 4.3.

4.2 Amplification of the hypervariable region

HVR1 sequence variants which emerged during the course of infection in each patient were detected by RT-PCR. Viral RNA extracted from patient sera was reversed transcribed to give cDNA copies of the virus genomes. For the majority of samples, random hexamers were used for reverse transcription of RNA; for a few samples reverse transcription was performed with antisense primer (c48 or SF4). Table 4.4 shows the sequence and position of primers used. Nested PCR was then performed. (See Chapter 2 section 3.6.2.)

A difficulty arose in that a single primer combination would not amplify the HVR1 from all samples. The original intention was to amplify both envelope genes, E1 and E2 using the primers SF1/SF4 in the first round of amplification and SF3/SF4 in the second, but this proved unsuccessful. It was then decided to amplify only the E2 HVR1 using primers described in the literature (Weiner *et al.*, 1992). The combination of primers of X14/X18 for the first round of amplification and X4/X19 for the second proved unsuccessful. On each occasion the X19 primer was used, no product was amplified. Using SF3/SF4 as the outer set of primers with X14/X18 as the inner set proved successful. It was felt, however, the PCR would be more efficient if the outer and inner set of primers were closer together. For this reason primers based upon those published by Cha *et al.* (1992), c48/c58, were synthesised (Table 4.4) to be used in conjunction with SF3/SF4 instead of X14/X18; however, not all the samples would amplify using this combination. A third primer combination was tried; c48/c58 as the outer primers and X14/X18 as the inner primers. This combination amplified some, but not all, of the samples which the other two primer combinations had failed to amplify. The final primer combination tried was c48/SF4 and X14/X18 which amplified some of

Table 4.1 Sera studied from patient KR

Sample	Diagnostic laboratory Number	Date sample taken
A	127/91	01-01-91
B	1532/91	05-03-91
C	6495/91	10-10-91
D	6110/92	14-08-92
E	8235/92	26-10-92
F	8657/92	04-11-92
G	87/93	06-01-93
H	3216/93	14-04-93
I	3955/93	07-05-93
J	8215/93	10-10-93
K	9122/93	08-11-93
L	631/94	24-01-94
M	2771/94	06-04-94

Table 4.2 Sera studied from patient SF

Sample	Diagnostic laboratory Number	Date sample taken
A	H2070/88	09-06-88
B	H2393/89	07-07-89
C	H3887/89	09-11-89
D	2505/90	04-04-90
E	5020/90	11-07-90
F	173/91	09-01-91
G	6552/91	02-10-91
H	9614/92	02-12-92
I	110/93	06-01-93
J	6494/93	04-08-93
K	636/94	24-01-94

Table 4.3 Sera studied from patient FAL

Sample	Diagnostic laboratory Number	Date sample taken
A	H767/88	02-03-88
B	H2373/88	06-07-88
C	H1599/89	03-05-89
D	3203/90	01-05-90
E	8475/90	04-12-90
F	158/91	08-01-91
G	2198/91	02-04-91
H	6537/91	01-10-91
I	8283/91	08-11-91
J	129/92	07-01-92
K	5111/92	07-07-92
L	No code assigned	12-09-94

Table 4.4 Oligonucleotide primers

Primer	Sequence 5' → 3'	Nucleotide position*	Use
SF1 (s)#	GACGGATCCAACTATGCAACAGGGA	478 to 502	PCR (E1/E2)
SF3 (s)#	GTGGATCCTTCTCTATCTTCCT	512 to 533	PCR (E1/E2)
SF4 (as)#	AAGGATCCGGGGGTAGTGCCAGCA	1456 to 1480	RT/PCR (E1/E2)
X14 (s) †	GGTGCTCACTGGGGAGTCCT	1048 to 1067	PCR (E2)
X4 (s) †	TCCATGGTGGGGAACTGGGC	1087 to 1106	PCR (E2)
X18 (as)†	CATTGCAGTTCA$^{A}/_{G}$GGCCGTGC	271 to 1291	PCR (E2)
X19 (as) †	TGCCAACTGCCGTTGGTGTT	1243 to 1262	PCR (E2)
c48 (s) ‡	ATGATGAACTGGTCCCCTAC	967 to 986	PCR (E2)
c58 (as) ‡	TGATAGG$^{G}/_{A}$CCCCA$^{G}/_{T}$CCCTGG$^{T}/_{G}$CAAA	1393 to 1417	RT/PCR (E2)

s: sense, as: antisense

* Numbering according to Choo *et al.* (1991).

# Denotes primer with an EcoR1 restriction site.

† Based on the Weiner *et al.* (1992) primers.

‡ Based on the Cha *et al.* (1992) primers.

the missing samples. Figure 4.2 shows an example of the products amplified. There were still some samples in which HCV RNA was detected by RT-PCR of the 5'NCR but in which the HVR1 could not be amplified using any of the above primer combinations. Table 4.5 summarises the successful primer combinations and amplification conditions. Different annealing temperatures and magnesium concentrations were also investigated. The conditions used in this study were those found to give the optimal results.

RNA was extracted from two separate aliquots of the original serum samples to ensure there had been no mix up in samples and to detect any potential artefacts that might arise from misincorporation of bases by *Taq* polymerase during amplification. Therefore, two independent sequences for each serum sample were obtained.

Amplification of HVR1 from early serum samples from the patients proved difficult, especially for patient KR around the time he became infected (1990). A previous worker within the laboratory in 1991 had successfully amplified the 5'NCR from sera taken from this patient from July 1990 onwards. Possibly subsequent storage and handling conditions of the serum had affected the stability of the HCV RNA.

4.3 Direct sequencing

The PCR products were purified and directly sequenced as described in Chapter 2 sections 3.8.3, 3.9 and 3.10. The primers used for sequencing were X4, X14, X18, c48 and c58. The majority of samples were sequenced using primers X4 and X18. As each serum was sampled twice, four sequencing reactions were performed; two in the sense and two in the antisense direction. A consensus sequence was then derived. At some positions on the gel, a sequence ambiguity was observed. Direct sequencing can identify both the predominant sequence and minor sequences, resulting in a mixture of bases (i.e. two bases) at the same position on the sequencing gel, see Figure 4.3.

Figure 4.2 2% Agarose gel electrophoresis of HVR1 second round PCR products. Lanes: (1, 7, 15) sample from patient KR; (2, 10, 18) negative control serum; (3, 11, 19) positive control serum; (4, 12, 20) negative control for reverse transcription reaction; (5, 13, 21) negative control for first round PCR reaction; (6, 14, 22) negative control for second round PCR reaction; (8, 16) sample from patient FAL; (9, 17) sample from patient SF. M, molecular weight marker (Boehringer Mannheim DNA molecular weight marker VI).

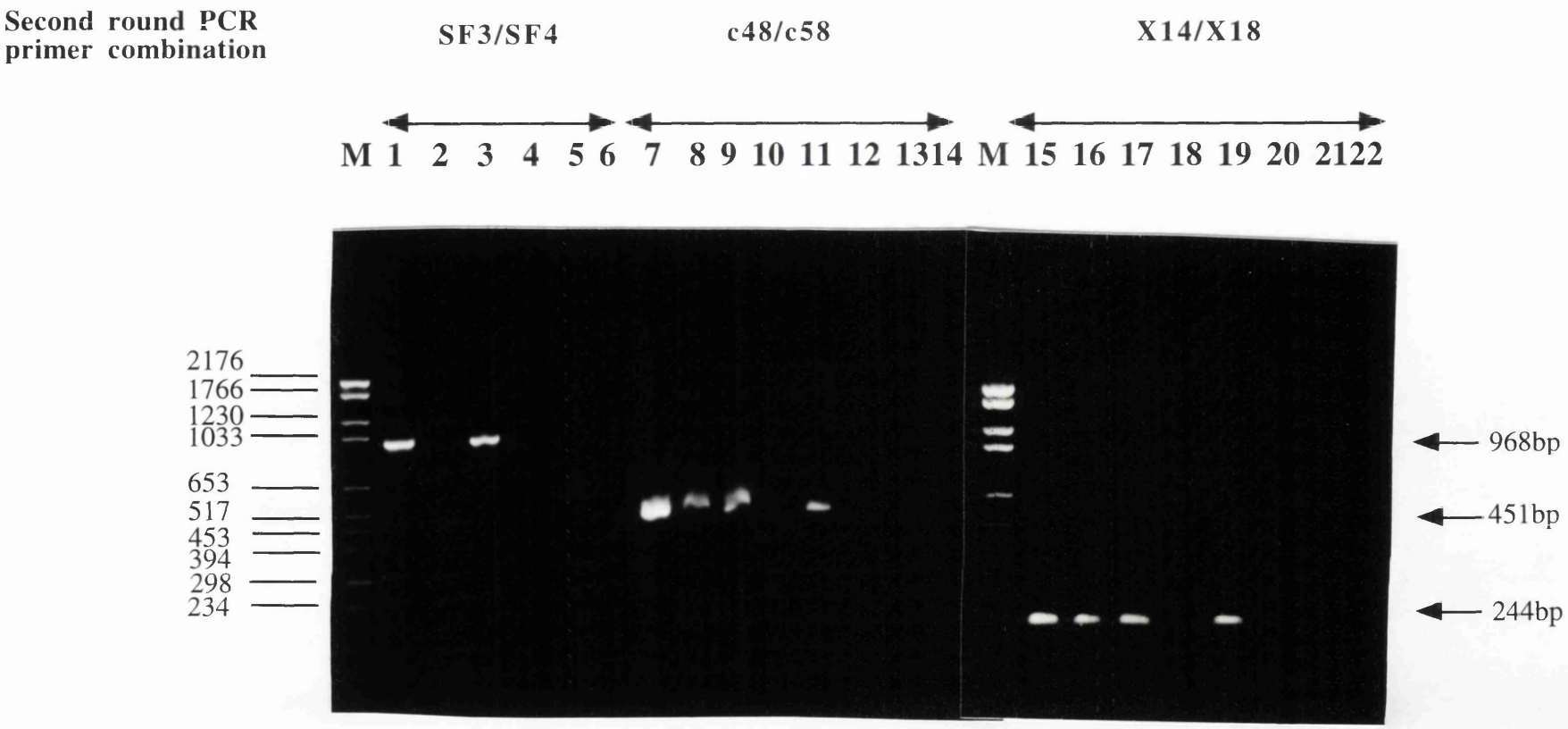

Second round PCR primer combination
SF3/SF4
c48/c58
X14/X18
M 1 2 3 4 5 6 7 8 9 10 11 12 1314 M 15 16 17 18 19 20 2122
2176
1766
1230
1033
653
517
453
394
298
234
968bp
451bp
244bp

Table 4.5 Amplification conditions

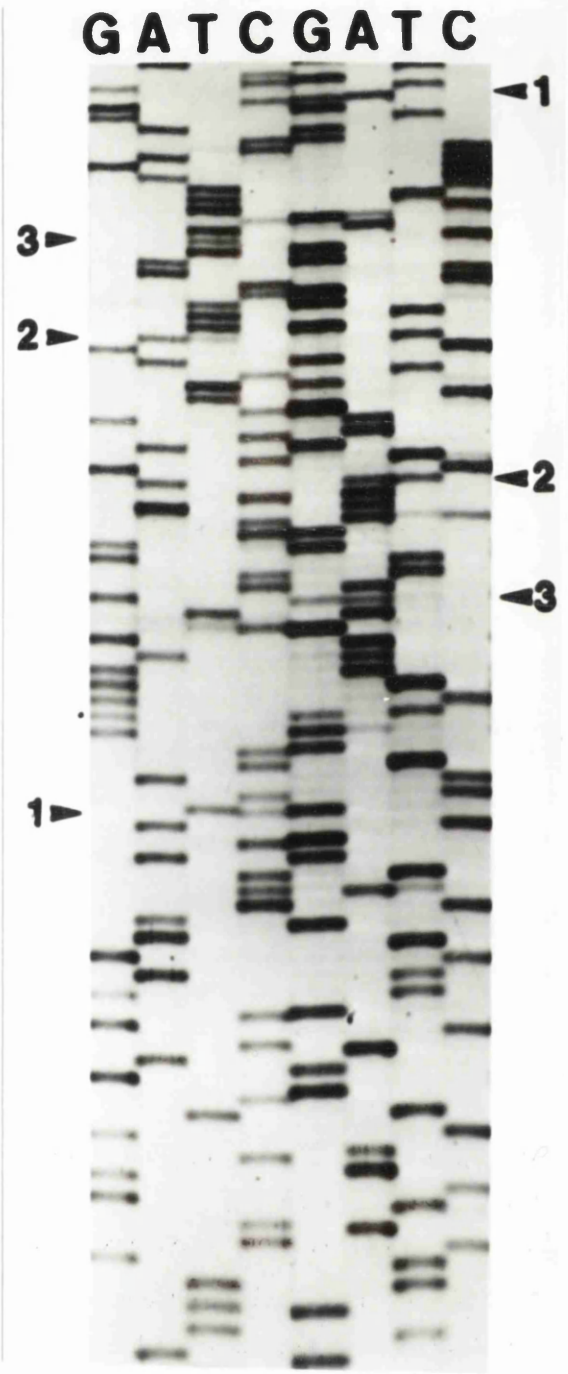
Primer combination	No. amplification cycles	PCR conditions
SF3/SF4	2†	94°C/1 min 45°C/2 mins 72°C/3 mins followed by
	28†	94°C/1 min 55°C/2 mins 72°C/3 mins
X14/X18	25‡	94°C/1 min 60°C/2 mins 72°C/3 mins
c48/c58	35†	94°C/1 min 55°C/2 mins 72°C/3 mins
	25‡	94°C/1 min 55°C/2 mins 72°C/3 mins
c48/SF4	30†	94°C/1 min 55°C/2 mins 72°C/3 mins

† Denotes first round amplification.

‡ Denotes second round amplification.

Figure 4.3 Direct DNA sequencing of a PCR amplified HVR1 cDNA fragment. The gel identified heterogeneity in the sample. Both the sense and antisense sequencing direction are shown. Mixed base positions present in the sample are indicated by arrows 1 to 3. At position 1, a mixture of thymine and cytosine are seen, since the thymine base displays the strongest signal intensity between the two bases in both directions it is recorded as the predominate base at that position; at position 2, a mixture of adenine and thymine are seen, in the sense direction the adenine base displays the strongest signal intensity but in the antisense direction the bases are of equal intensity, this would be recorded as a mixed base position; and at position 3, a mixture of thymine and cytosine of equal signal intensity are seen in the antisense direction whereas only thymine is seen in the sense direction. The thymine at position 3 in the sense direction is seen more clearly on other gels of the same sequence and so, the thymine is recorded as the predominate base at this position.

SENSE ANTI-SENSE



To distinguish between the predominant sequence and minor sequences and derive the consensus sequence, the following criteria were used :

1) If the sequencing gel showed two bases at a given position and one base displayed the strongest signal intensity in at least three of the four sequences it was recorded as the predominant base.

2) If at least two sequences showed a mixture of bases in both the sense and antisense direction of equal signal intensity, a mixed base position was recorded.

3) If two of the four sequences showed a mixture of bases but were the same sense, the predominant base was recorded as the base displayed in the sequence of the opposite strand.

HVR1 sequences were aligned using the programs PILEUP and PRETTY for nucleotide comparisons and TRANSLATE, PILEUP and PRETTY for derived amino acid sequence comparisons. These are programs in the Genetic Computer Group sequence analysis package version 7. The sequence data was also analysed using the computer program PHYLIP version 3.5 (Felsentein., 1993) and the MEGA, Molecular Evolutionary Genetics Analysis, package version 1.01 (Kumar.S *et al.*, 1993).

4.4 Emergence of HVR1 sequence variants in Patient KR

ALT levels were normally measured at least twice a month. Unfortunately due to difficulties experienced with amplifying HCV RNA from patient sera and because samples were not sent to the Virology Laboratory every time ALTs were performed, HCV sequences could not be obtained at each time for which an ALT level is available. The ALT profile of patient KR is shown in Figure 4.4. The graph depicts the highest ALT value recorded each month. The arrows depict the times at which HCV sequences

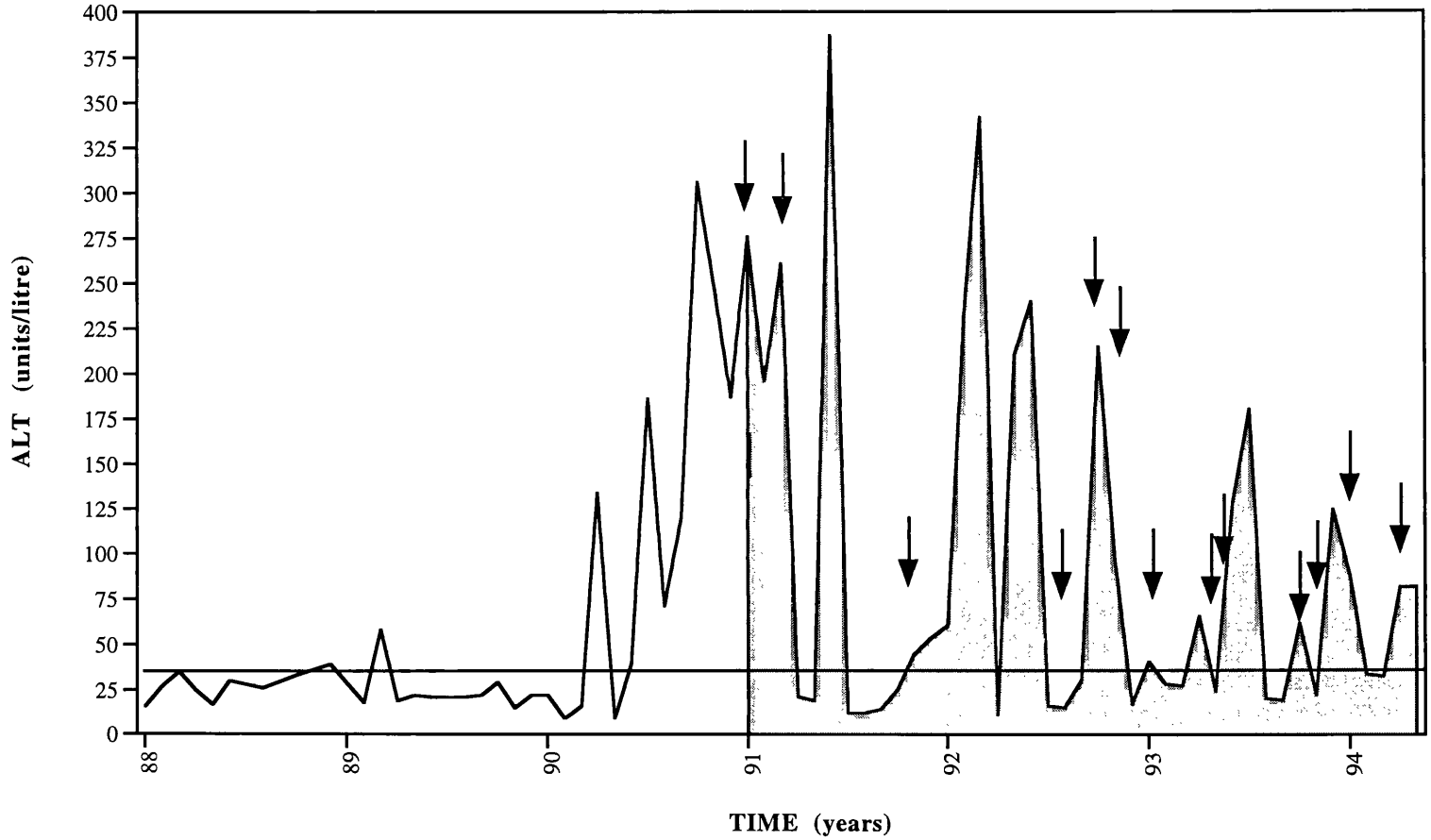


Figure 4.4 Patient KR ALT profile. Arrows denote sampling points. Solid line across graph indicates normal ALT level 40units/litre. Shaded area represents observation period.

were detected. Until 1990, the ALT levels were normal (≤40 U/L). In April of 1990 the ALT level peaked at 134 U/L after KR received his kidney transplant. The ALT level returned to normal before peaking again in July (186 U/L), the month during which HCV RNA was detected for the first time. Between July and October the ALT level was normal or near normal before increasing in October to levels above 200 U/L at which time antibody to HCV (by 3rd generation RIBA) appeared. ALT levels remained elevated (above 100 U/L) until April 1991 after which they fluctuated, with peak values of between 124 to 380 U/L.

The predominant HVR1 nucleotide and deduced amino acid sequences are shown in Figures 4.5 and 4.6, respectively. The predominant HVR1 sequence found on sequencing sample A (Jan'91) changed within a period of 63 days with a new variant emerging, with a T to C transition at position 1204, (Phe-402 to Leu). Discounting mixed based positions, a total of nineteen nucleotide changes occurred in the HVR1 over a period of three years, summarised in Table 4.6. All nineteen nucleotide changes were non-silent and occurred in either the first or second codon position. Transitions were; six G to A, three A to G, one T to C and one C to T. The transversions were, two A to T , two C to A and one each of T to G, C to G, T to A and A to C.

The first double band A/G (mixed base position) appeared at position 1211 in sample G (April 1993). This denoted two populations, one with Lys and one with Arg at position 404. Lys appeared at this position in the predominant HVR1 sequence in sample D (August 1992) changing to Arg in sample E (October 1992). Arg was found at this position in all subsequent sequences until sample M (April 1994), when an Arg to Ser change occurred. All of the mixed based positions resulted in the presence of two amino acids, one of which was present at that position in either a previous or a subsequent HVR1 variant sequence. The exception to this was the appearance of a G/T double band at position 1219 in samples H and I (April and May 1993) denoting Ala and Ser at position 407; Ala was found at this position in all other sequences. The predominant sequence found in sample M (April 1994) had a mixture of two bases (A/G) at nucleotide position 1340 (Figure 4.5) outside HVR1, predicting an asparagine

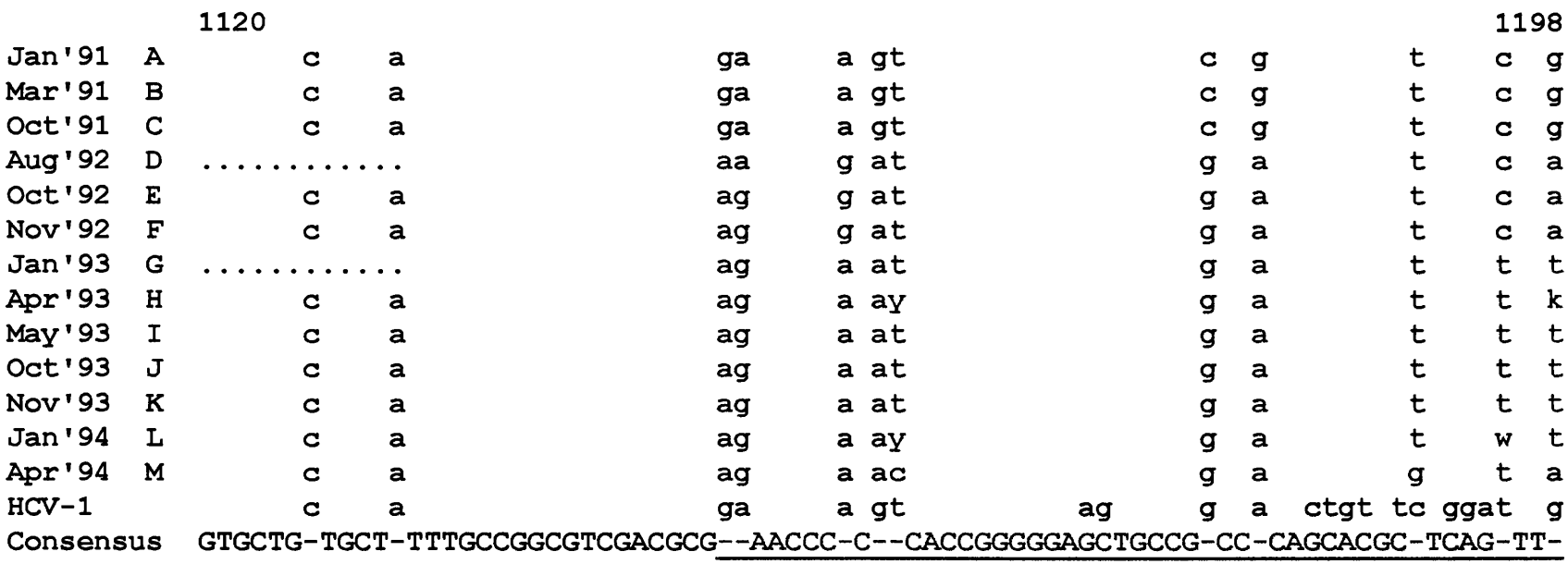


Figure 4.5 Alignment of HCV isolates from patient KR between nucleotide residues 1120 and 1248. Consensus sequence for HVR (residues 1150-1233) shown in red type face and underlined. Green type face denotes a mixed nucleotide position. Blue type face denotes a nucleotide change and magenta type face a second change. *Nucleotide numbering according to Choo *et al.* (1991).

```
          1199                                                    1248
Jan'91  A    g t     caa      g   a              a c
Mar'91  B    g c     caa      g   a              a c
Oct'91  C    g c     caa      g   t              a c
Aug'92  D    a c     aaa      g   t              a g
Oct'92  E    a c     aga      g   t              a g
Nov'92  F    a c     aga      g   t              a g
Jan'93  G    a c     ara      g   t              a g
Apr'93  H    a c     aga      k   t              a g
May'93  I    a c     aga      k   t              a g
Oct'93  J    a a     aga      g   w              a g
Nov'93  K    r y     aga      g   t              a g
Jan'94  L    a y     aga      g   t              a g
Apr'94  M    a c     agc      g   t              r g
HCV-1        t tgcc  c cgca   g   a          c   g c
Consensus    CCA-T-TCTTT---CCAGGC-CCA-GCAGAACGTCCAGTTG-T-AACACC
```

Figure 4.5 Continued.

	374							417
Jan'91 A	e	hv	ar	l	lasf	q	ak	i
Mar'91 B	e	hv	ar	l	lasl	q	ak	i
Oct'91 C	e	hv	ar	l	lasl	q	am	i
Aug'92 D	k	ri	gh	l	ltnl	k	am	m
Oct'92 E	r	ri	gh	l	ltnl	r	am	m
Nov'92 F	r	ri	gh	l	ltnl	r	am	m
Jan'93 G	r	hi	gh	l	fsnl	*k* *r*	am	m
Apr'93 H	r	hi t	gh	l	f*a*snl	r	am *s*	m
May'93 I	r	hi	gh	l	fsnl	r	am *s*	m
Oct'93 J	r	hi	gh	l	fsni	r	am *k*	m
Nov'93 K	r	hi	gh	l	fsn*sf*l	r	am	m
Jan'94 L	r	h*i* t	gh	l	*f*sn*l* *l*	r *f*	am	m
Apr'94 M	r	ht	gh	r	ftnl	s	am	*m* *n*
HCV-1	e	hv	s ghtvsgfvslla				ak	i

Consensus VLLLFAGVDA-T--TGGAA--ST-Q----F-PG--QNVQL-NTN

Figure 4.6 Alignment of the derived amino acid sequences for HCV isolates from patient KR between residues 374 and 417. Consensus sequence for HVR (residues 384-411) shown in red type face and underlined. Blue type face denotes an amino acid change, magenta type face denotes a second and green type face denotes a third change. Italics indicate the possibility of the presence of both amino acids, because of nucleotide sequence ambiguity. *Numbering according to Choo *et al.* (1991).

residue at amino acid position 420 (Figure 4.6) resulting in a potential N-linked glycosylation site. Figure 4.7 shows sequencing gels from isolates from KR. Arrows denote mixed base positions. When the mixed base positions are taken into account the ratio of transitions to transversions was 1.1 to 1; without taking them into account the ratio was 1.4 to 1 (Table 4.6).

4.5 Emergence of HVR1 sequence variants in Patient SF

The ALT profile of patient SF is shown in Figure 4.8. Persistently high ALT levels (up to 272 U/L) are observed from 1983 to 1986, after which normal or near normal levels prevailed apart from the occasional peak. During the observation period the ALT levels ranged between 11 and 49 U/L, apart from two peaks of 99 and 79 U/L in July 1990 and 1991, respectively.

HVR1 nucleotide and amino acid sequences are shown in Figures 4.9 and 4.10, respectively. No changes are observed between June 1988 and April 1990. After this there is continuous change except in December 1992 and January 1993. The first sequence change observed occurs at the point where the ALT level has risen above normal in sample E (July 1990). A double band, C/T appears at position 1214 denoting the presence of Thr and Ile at amino acid position 405; previously only Thr was found at this position. Amino acid position 405 appears to be very variable. At the nucleotide level the changes which occur are A to T to C/T to G to A/C to C to A and back again to C. The result of these changes in terms of amino acid change is Thr to Thr/Ile to Ala to Thr/Pro to Pro to Thr to Pro. The double band, A/G, which appears at position 1167 (sample F) and 1173 (sample G) does not lead to a change in amino acid (Figure 4.9 and 4.10). A further silent mutation, C to T, occurs at position 1158 (sample H and I). The predominant HVR1 sequence obtained in sample J (August 1993) has Arg at position 400; Ser is at this position in all the other samples. Six amino acid changes occur between January and August 1993 with no peak in ALT being observed. The

Figure 4.7 Direct DNA sequencing of the PCR amplified cDNA fragments of the HVR1 from patient KR sera. Arrows (◄) denote the HVR1. The regions of nucleotide changes between the four sequential samples are indicated by arrows 1, 2 and 3. At position 1 (sense strand) the sequence reads AAGAGAA in Jan'91 and Mar' 91 but ATGAGAA in Oct'91 and Aug'92, thymine having replaced adenine; at position 2 (sense strand) the sequence reads TTTCTTT in Jan'91 but TCTCTTT in Mar'91, Oct'91 and Aug'92, cytosine having replaced thymine; at position 3 (sense strand) the sequence reads GGAAA in Jan'91, Mar'91 and Oct'91 but GAAAA in Aug'92, adenine having replaced guanine.

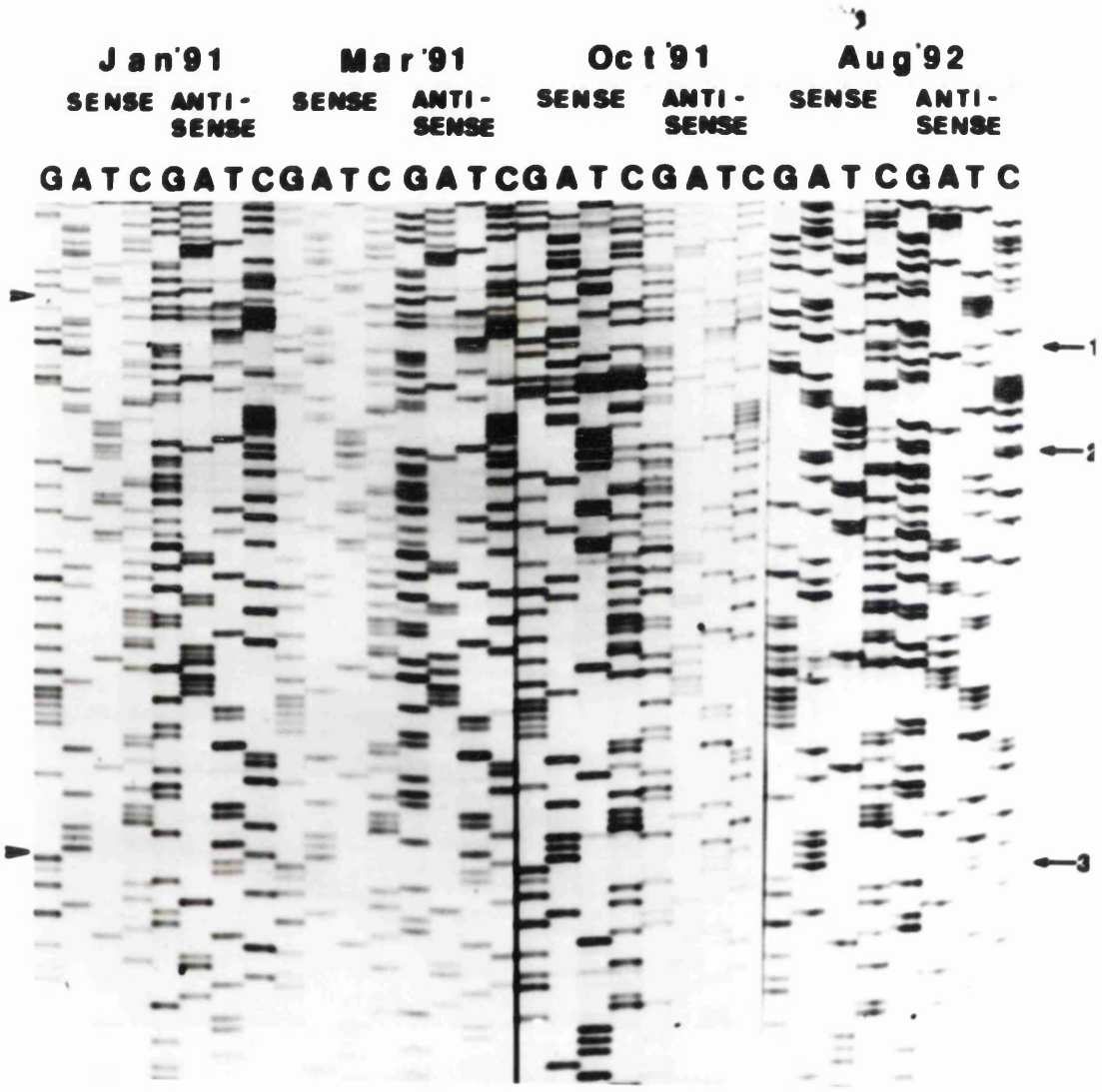

Jan'91
Mar'91
Oct'91
Aug'92
SENSE
ANTI-SENSE
SENSE
ANTI-SENSE
SENSE
ANTI-SENSE
SENSE
ANTI-SENSE
GATCGATCGATCGATCGATCGATCGATCGATC
1
2
3

Table 4.6 Nucleotide and amino acid sequence differences of the HVR between HCV isolates of Patient KR (excluding mixed base positions)

ISOLATES	CHANGES (%)		TRANSITION VS TRANSVERSION	CODON POSITION		
	nucleotide	amino acid		1st	2nd	3rd
Jan'91 vs Mar'91	1/84 (1.2)	1/28 (3.6)	1 : 0	1		
Mar'91 vs Oct'91	1/84 (1.2)	1/28 (3.6)	0 : 1		1	
Oct'91 vs Aug'92	8/84 (9.6)	8/28 (28.6)	6 : 2	3	5	
Aug'92 vs Oct'92	2/84 (2.4)	2/28 (7.2)	2 : 0		2	
Oct'92 vs Nov'92	0/84	NA	NA		NA	
Nov'92 vs Jan'93	3/84 (3.6)	3/28 (10.8)	2 : 1	2	1	
Jan'93 vs Apr'93	0/84	NA	NA		NA	
Apr'93 vs May'93	0/84	NA	NA		NA	
May'93 vs Oct'93	1/84 (1.2)	1/28 (3.6)	0 : 1	1		
Oct'93 vs Nov'93	0/84	NA	NA		NA	
Nov'93 vs Jan'94	0/84	NA	NA		NA	
Jan'94 vs Apr'94	3/84 (3.6)	3/28 (10.8)	0 : 3	1	2	
TOTAL	19/84 (22.6)	19/28 (67.8)	11 : 8	8	11	0

NA denotes not applicable

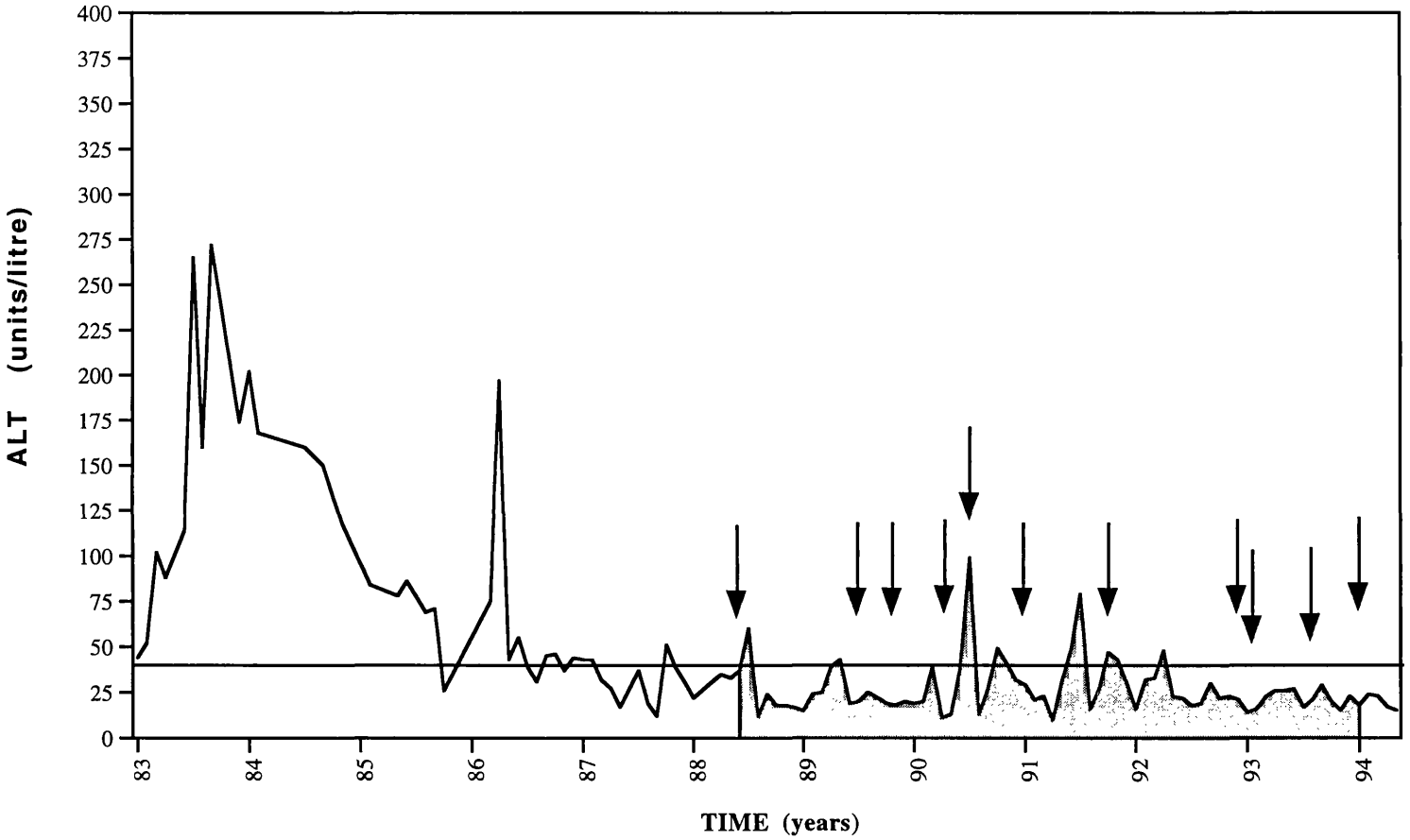


Figure 4.8 Patient SF ALT profile.Solid line across the graph indicates normal ALT level of 40units/litre. Shaded area represents observation period.

Figure 4.9 Alignment of HCV isolates from patient SF between nucleotide residues 1129 and 1251. Consensus sequence for HVR (residues 1150-1233) is shown in red type face and underlined. Green type face denotes a mixed nucleotide position. Blue type face denotes a nucleotide change and magenta type face a second change. *Nucleotide numbering according to Choo *et al.* (1991)

```
          1129                                                                1204
Jun'88  A                                  c c                 a             c g
Jul'89  B                                  c c                 a             c g
Nov'89  C                                  c c                 a             c g
Apr'90  D                                  c c                 a             c g
Jul'90  E                                  c c                 a             c g
Jan'91  F                                  c c           r     a             c g
Oct'91  G                                  c c     r           a             y t
Dec'92  H                                  t t                 g             c g
Jan'93  I                                  t t                 g             c g
Aug'93  J                                  c c                 a             t c   cg
Jan'94  K                                  c c       rr        a             t c      t
HCV-BK      t  t  t  c         gat   c cgtg  a    g    ca  c  aa  cac aa a  c cgt    a
HCV-1       a  t     c  c    c gaa   c cgcc  c    gaagt  c   ca      gtct  at tgttag
Consensus  CTCTTCGCCGGTGTTGACGGGACCACCTA-A-AACGGGGGAGGCGGCAGGCA-CACTGTCCGCGGG-T-ACGTCCC

          1205                                                1251
Jun'88  A       t  a
Jul'89  B       t  a
Nov'89  C       t  a
Apr'90  D       t  a
Jul'90  E       t  ay
Jan'91  F       t  g
Oct'91  G       t  m
Dec'92  H       g  c
Jan'93  I       g  c
Aug'93  J       t  a                    c
Jan'94  K       t  c
HCV-BK      g  cg  ag              a          a    c     t
HCV-1        c cg  c a  cg caag    acg       ga c  c
Consensus  TCTTT-CA-CTGGGCCGTCTCAGAGAATCCAGCTTGTAAATACCAAC
```

```
           378                                 417

Jun'88  A           t  e   n    l     t

Jul'89  B           t  e   n    l     t

Nov'89  C           t  e   n    l     t

Apr'90  D           t  e   n    l     t

Jul'90  E           t  e   n    l     t
                                      i

Jan'91  F           t  e   n    l     a

Oct'91  G           t  e   n    l     t
                                f     p

Dec'92  H           i  e   s    l    ap

Jan'93  I           i  e   s    l    ap

Aug'93  J           t  e   n    f r   t         h

Jan'94  K           t  e   n    f  f  p
                       g

   HCV-BK        d hv  g qak tnrlv   as    k   i

    HCV-1       ae hv  gs  h  s fv  lap ak nv  i

Consensus  FAGVDGTTY-TG-AAG-TVRG-TSLFS-GPSQRIQLVNTN
```

Figure 4.10 Alignment of the derived amino acid sequences of HCV isolates from patient SF between residues 378 and 417. Consensus sequence for HVR (residues 384-411) shown in red type face and underlined. Blue type face denotes an amino acid change, magenta type face denotes a second and green type face denotes a third change. Italics indicate the possibility of the presence of both amino acids, because of nucleotide sequence ambiguity.*Numbering according to Choo *et al.* (1991).

ratio of transitions to transversion is 1.8 to 1 and 1.2 to 1, when mixed base positions are excluded.

4.6 Emergence of HVR1 sequence variants in Patient FAL

The ALT profile of patient FAL is shown in Figure 4.11. FAL displays normal or near normal ALT values apart from the occasional peak, the highest of which is 102 U/L in August 1986. Figures 4.12 and 4.13 show the HVR1 nucleotide and derived amino acid sequences. Between sample A (March 1988) and sample G (April 1991) there are nine amino acid changes and between sample G and sample I (November 1991) there is only one change. At position 1156, a C to C/T change occurs. The resulting amino acid change is His to the combined presence of His and Tyr. At position 1198 a T/A to T change occurs, the result of which is the presence of Thr alone at amino acid position 400. At position 1224 a T to T/G to a T change occurs, denoting the presence of Ser and Arg respectively, at position 408 in sample G. This is the only sequence difference between samples G and I. Base changes at positions 1180 and 1181 result in Arg being replaced by Tyr at amino acid position 394, resulting in a polar charged amino acid being replaced with a polar uncharged amino acid. At position 404 Ala, a hydrophobic amino acid is replaced by Ser, a hydrophilic amino acid. Ser-405 is replaced by Pro, the result of a T to C change at position 1213. Only one silent change occurs, G to G/A to A , at position 1534. When the mixed base positions are taken into account the ratio of transitions to transversions is 1.6 to 1. When they are excluded, the ratio is 6 to 1.

Although the HVR1 from twelve serum samples from patient FAL were amplified, the results of nucleotide sequencing of the PCR products could be read clearly for only the three samples described above. Figure 4.14 shows an example of a sequencing gel for patient FAL. Two regions, denoted by arrows, proved to be the most difficult to read. This inability to read the gels could be explained by the presence *in vivo* of a virus population consisting of a heterogeneous mixture of genetically

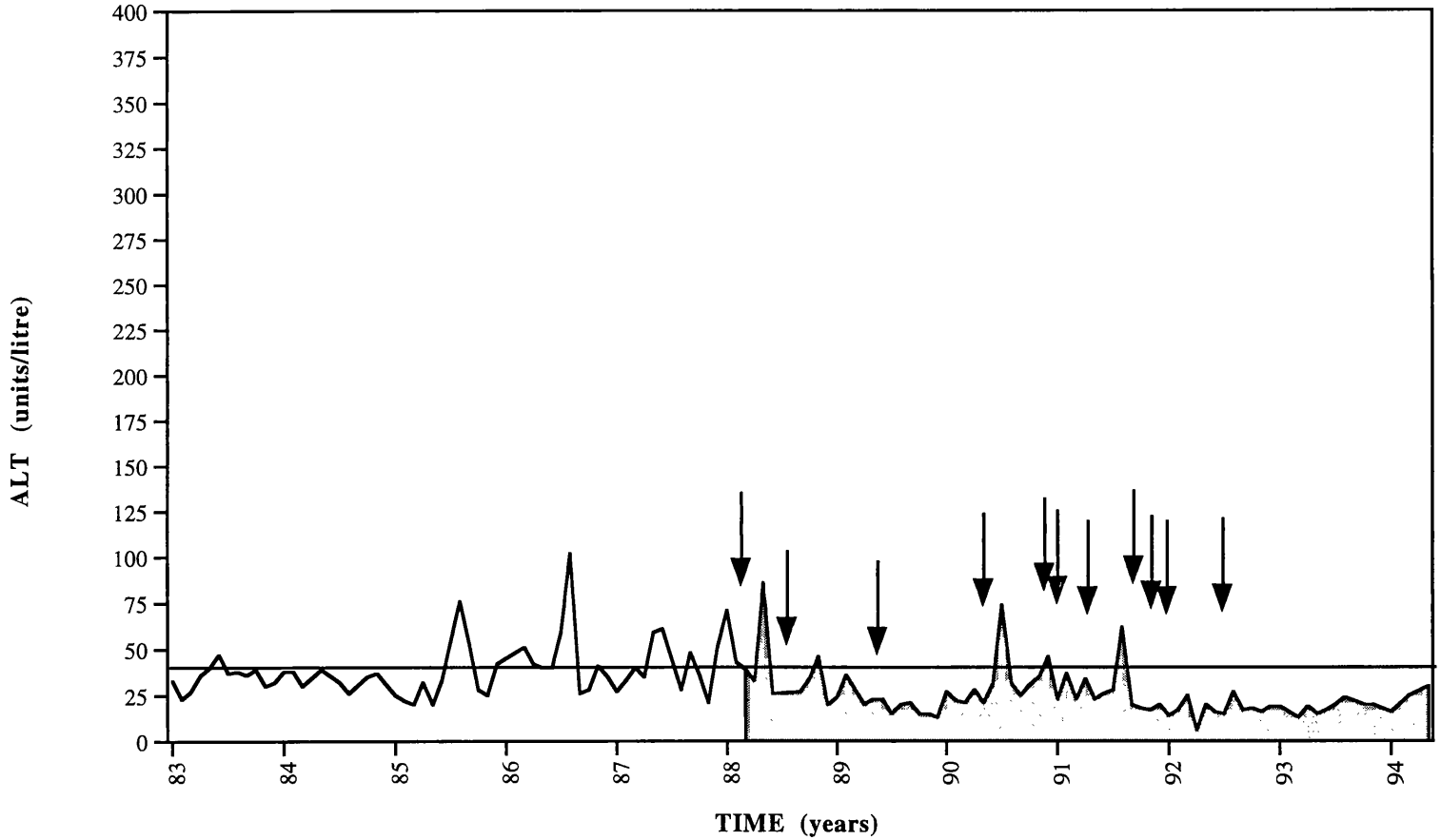


Figure 4.11 Patient FAL ALT profile.Solid line across the graph indicates normal ALT level of 40units/litre. Shaded area represents observation period.

```
            1120                                                1165

Mar'88  A                                   r       c

Apr'91  G                                   g       y

Nov'91  I                                   g       y

HCV-1                                       g       c

Consensus  GTGCTGCTGCTATTTGCCGGCGTCGACGC-GAAACC-ACGTCACCG

            1166                                                1212

Mar'88  A           t     cg                c  w    at       g

Apr'91  G           c     ta                t  a    gk       t

Nov'91  I           c     ta                t  a    gk       t

HCV-1          ag   c  g ca actgt t  g at   gt   gc    c cg

Consensus  GGGGAGCTG-CGCC--CCACACGGCTAGC-TT-CTA--CTCTTT-CA

            1213                                    1258

Mar'88  A  t           t  g             a

Apr'91  G  c           k  r             r

Nov'91  I  c           t  a             a

HCV-1      c         ag  g   g      c   a

Consensus  -CAGGCGCCAG-CA-AACATCCAGTTG-TCAACACCAACG
```

Figure 4.12 Alignment of HCV isolates from patient FAL between nucleotide residues 1120 and 1258. Consensus sequence for HVR (residues 1150-1233) shown in red type face and underlined. Green type face denotes a mixed nucleotide position. Blue type face denotes a nucleotide change. *Nucleotide numbering according to Choo *et al.* (1991).

	374																			417
Mar'88	A				h			v	r		l	*t*	n		as		s			i
												s								
Apr'91	G				*h*			a	y		f	t	*s*		sp		*s*			*i*
					y								*r*				*r*			*v*
Nov'91	I				*h*			a	y		f	t	*s*		sp		s			i
					y								*r*							
HCV-1					h			s	a	g	h	t	v	s	g	f	v	s	l	a	p		k		v	i
Consensus		V	L	L	L	F	A	G	V	D	A	E	T	-	V	T	G	G	A	-	A	-	H	T	A	S	-	-	-	L	F	-	-	G	A	-	Q	N	I	Q	L	-	N	T	N

Figure 4.13 Alignment of the derived amino acid sequences from residues 374 to 417 of HCV isolates from patient FAL. Consensus sequence for HVR (residues 384-411) shown in red type face and underlined. Blue type face denotes first amino acid change. Magenta type face denotes second amino acid change. Italics indicate the possibility of the presence of both amino acids, because of nucleotide sequence ambiguity. *Numbering according to Choo *et al.* (1991).

Figure 4.14 Direct DNA sequencing of the PCR amplified cDNA fragments of the HVR1 from patient FAL sera. Arrows denote the HVR1. The gels identified heterogeneity in patient FAL within this region.


Dec'90
Jan'91
Apr'91
Oct'91
SENSE
ANTI-SENSE
SENSE
ANTI-SENSE
SENSE
ANTI-SENSE
SENSE
ANTI-SENSE
GATCGATCGATCGATCGATCGATCGATCGATC

different but related variants termed 'quasispecies', one of which has an insertion or a deletion in its HVR1 nucleotide sequence. To test this, one would need to either clone the PCR products, perform limiting dilutions or single stranded conformation polymorphism (SSCP) gel electrophoresis before sequencing, to allow detection of both the major population and minor populations present.

4.7 SSCP analysis

To try and resolve the problem presented by a mixed virus population when direct sequencing of PCR products was used, SSCP gel electrophoresis was attempted. This has the advantage of detecting minor HCV quasispecies populations present as well as the major one. The technique of SSCP allows the detection of a single base change in short DNA fragments based on differences in mobility of single-stranded DNA in non-denaturing polyacrylamide gels. Second round PCR products amplified from the HVR1 of the HCV genome from three different serum samples from the same individual were subjected to SSCP analysis. Double-stranded DNA PCR products were denatured prior to electrophoresis in a non-denaturing gel. The products were separated into bands of different mobility in the gel. The mobility of the DNA single strands within the gel is dependent on the conformation adopted by each single strand, which is determined by its nucleotide sequence.

An example of typical SSCP gel is shown in Figure 4.15. Samples were heat denatured and alkali denatured to determine which method gave better separation into single-stranded DNA. The SSCP band profiles (pattern and intensity of the band) was better for heat denatured (lanes 6 to 10) products compared to alkali denatured products (lanes 1 to 5). Since it was not known if the SSCP would be sensitive enough to detect a one or two nucleotide difference in sequence between the three HVR1 PCR products, products amplified from the 5'NCR of the HCV genome were included. The two 5'NCR products each represent a different genotype of HCV, type 1 (lanes 1 and 6)

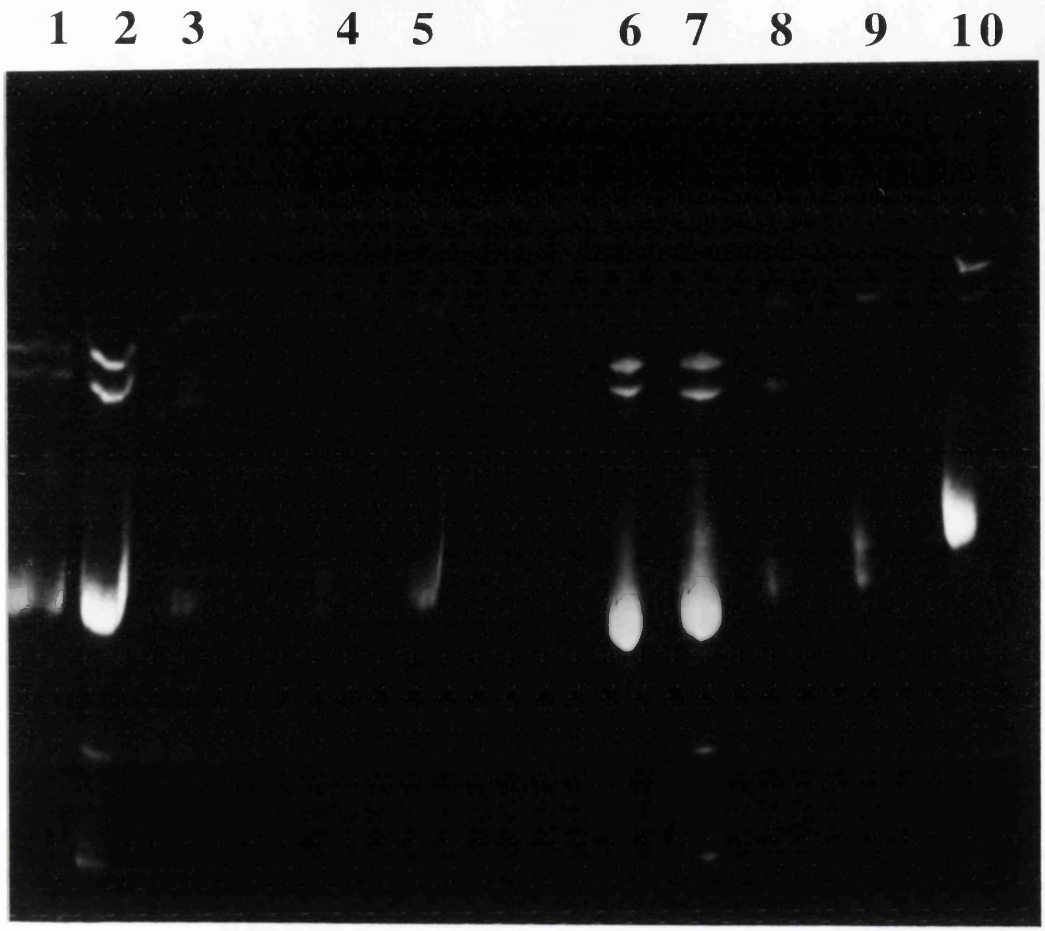


Figure 4.15 SSCP gel. Discontinuous gel electrophoresis was used. DNA was denatured by passage through a 'stacking' gel, 8% (w/v) polyacrylamide:bisacrylamide (19:1 ratio) containing formamide and assumed secondary conformation in a 'resolving' gel, 6% (w/v) polyacrylamide:bisacrylamide (49:1 ratio). 10µl of PCR product was run in each lane. Lanes: (1, 6) type 1 isolate 5'NCR; (2, 7) type 3 isolate 5'NCR; (3, 8) patient KR HVR1 sample C; (4, 9) patient KR HVR1 sample I; (5, 10) patient KR HVR1 sample G. Lanes 1-5 samples alkali denatured and lanes 6-10 samples heat denatured. Gel run at 600 volts.

and type 3 (lanes 2 and 7), respectively. The number of nucleotide differences between these two genotypes in the 5'NCR is at least twenty. However, both the type 1 and type 3 isolates migrate with the same mobility in the gel; despite base differences in their nucleotide sequence. The two bands seen in each lane represent the sense and antisense strands. A difference in the mobility of one or both of the single strands reflects sequence differences. The double stranded DNA fluoresces more intensely than the single strands. An undenatured double stranded control should have been included, giving an indication of product size. The resolution (separation) of the 5'NCR products was better than that of the HVR1 products. Instead of being clear like the 5'NCR bands, the bands are smeared and difficult to see. The HVR1 products appear to migrate with different mobility, indicating the presence of different quasispecies. Due to poor resolution, the number of coexisting quasispecies cannot be determined.

4.8 Phylogenetic analysis of patient HVR1 sequences

The MEGA computer package was originally used to construct phylogenetic trees for each of the three patients under study. However, the MEGA package does not recognise ambiguities in IUPAC single-letter codes, thus a mixed base position is not recognised. Because of this, two or more sequences have to be entered to represent each base present at a mixed base position. Therefore, patient sequences with more than two mixed base positions cannot be entered. To overcome this problem the PHYLIP computer package was used instead. This programme will accommodate mixed base positions as given in the IUB code (Nomenclature Committe, 1985). Phylogenetic trees were constructed from Maximum Likelihood distances using the neighbour joining method for each of the three patients under study. Evolutionary relations of the HVR1 sequences in each individual patient are shown in Figures 4.16 to 4.18. Horizontal lines indicate the distance between each HVR1 sequence. Analysis of the trees suggest that in patient KR the 1991 HVR1 sequences are from a different lineage to those that

Figure 4.16 Phylogenetic analysis of HVR1 sequences from Patient KR. Sequentially recovered predominant HCV HVR1 sequences were analysed for their evolutionary relationship. The sequence of isolate HCV-1 (Choo *et al.*, 1991) was used as a reference group. Horizontal lines represent the evolutionary distances between the HVR1 sequences. The numbers at the forks are the bootstrap values (%). The HVR1 sequences are Kr91a (Jan'91), Kr91b (Mar'91), Kr91c (Oct'91), Kr92a (Aug'92), Kr92b (Oct'92), Kr92c (Nov'92), Kr93a (Jan'93), Kr93b (Apr'93), Kr93c (May'93), Kr93d (Oct'93), Kr93e (Nov'93), Kr94a (Jan'94) and Kr94b (Apr'94).

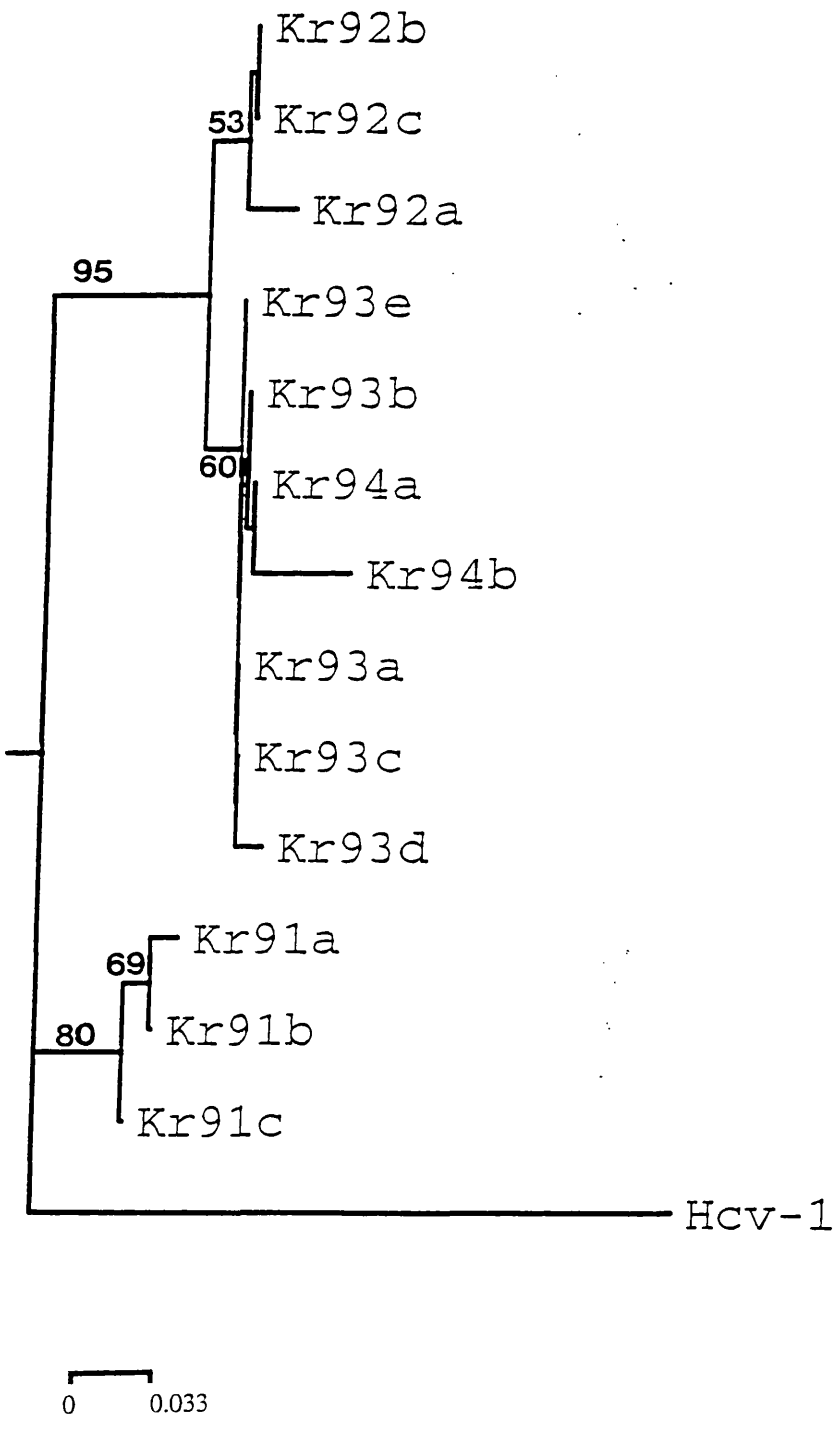

Kr92b
Kr92c
53
Kr92a
95
Kr93e
Kr93b
60
Kr94a
Kr94b
Kr93a
Kr93c
Kr93d
Kr91a
69
Kr91b
80
Kr91c
Hcv-1
0
0.033

Figure 4.17 Phylogenetic analysis of HVR1 sequences from Patient SF. Sequentially recovered predominant HCV HVR1 sequences were analysed for their evolutionary relationship. The sequence of isolate HCV-1 (Choo *et al.*, 1991) was used as a reference group. Horizontal lines represent the evolutionary distances between the HVR1 sequences. The numbers at the forks are the bootstrap values (%). The HVR1 sequences are Sf88 (Jun'88), Sf89a (Jul'89), Sf89b (Nov'89), Sf90a (Apr'90), Sf90b (Jul'90), Sf91a (Jan'91), Sf91b (Oct'91), Sf92 (Dec'92), Sf93a (Jan'93), Sf93b (Aug'93) and Sf94 (Jan'94).

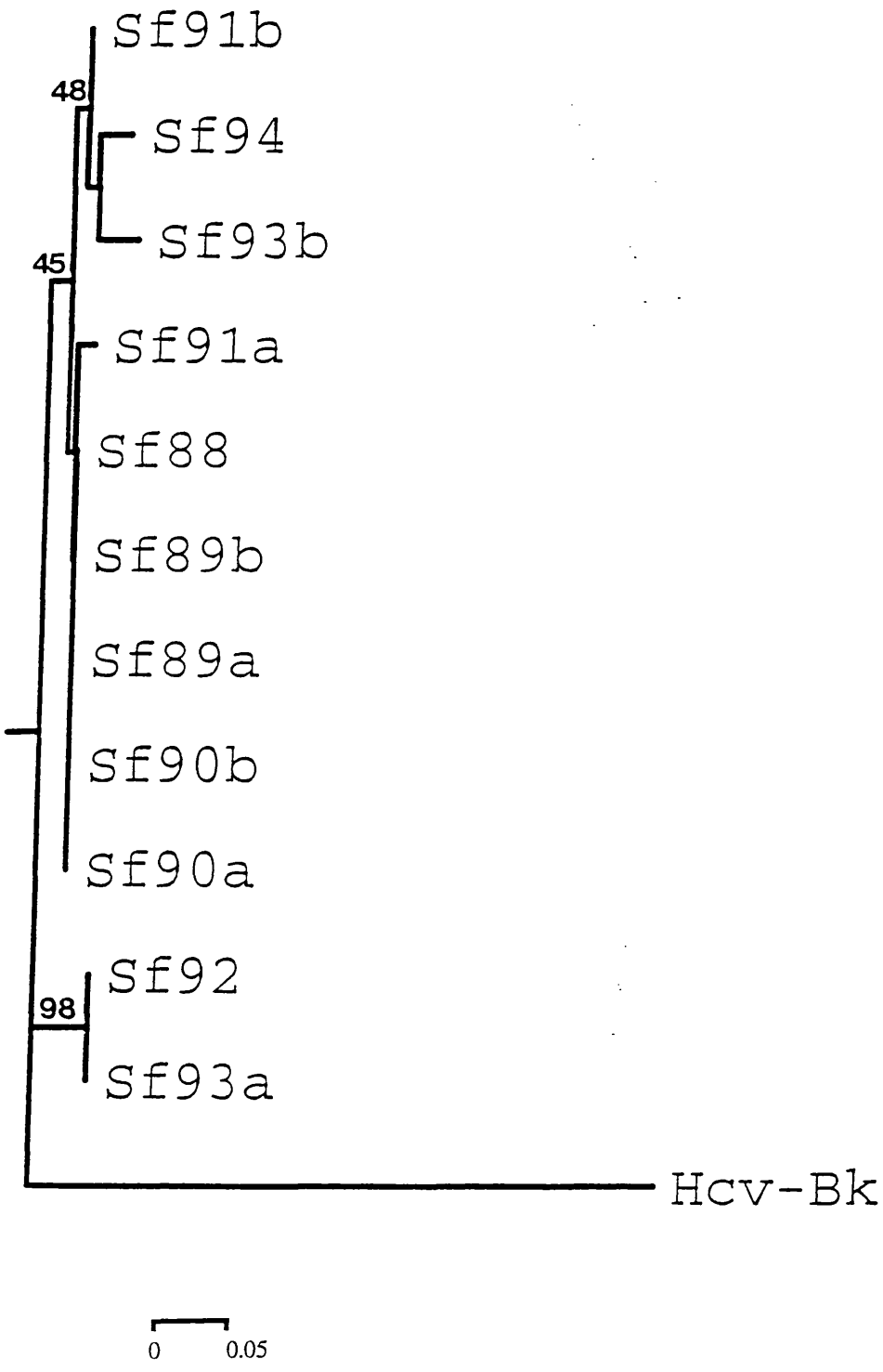
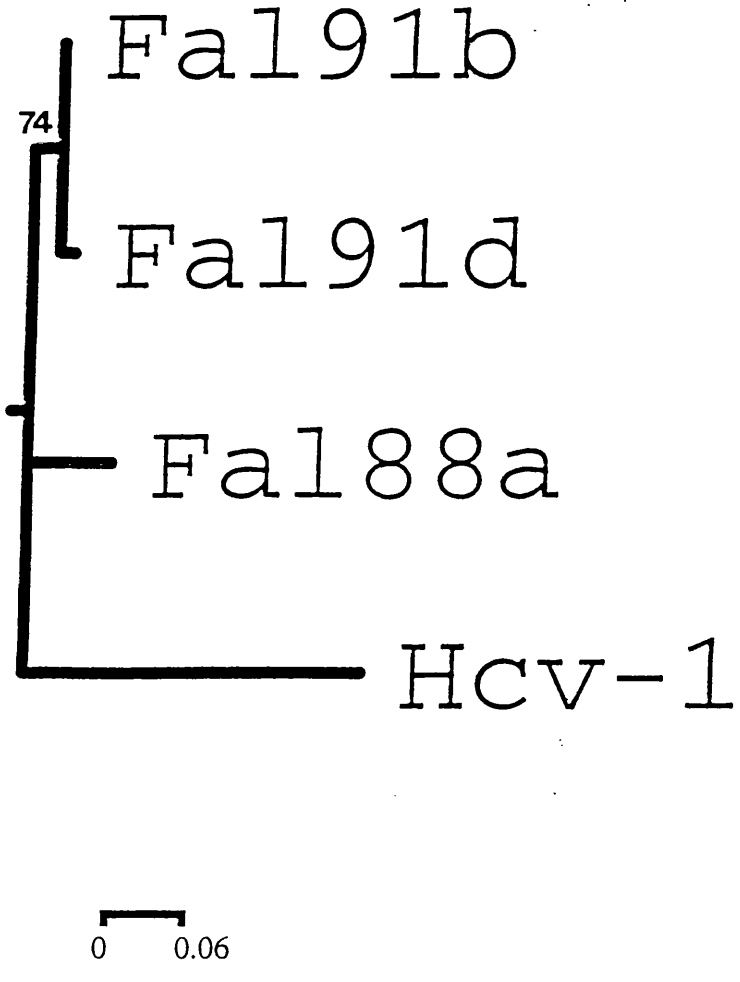

Sf91b
48
Sf94
Sf93b
45
Sf91a
Sf88
Sf89b
Sf89a
Sf90b
Sf90a
Sf92
98
Sf93a
Hcv-Bk
0
0.05

Figure 4.18 Phylogenetic analysis of HVR1 sequences from Patient FAL. Sequentially recovered predominant HCV HVR1 sequences were analysed for their evolutionary relationship. The sequence of isolate HCV-1 (Choo *et al.*, 1991) was used as a reference group. Horizontal lines represent the evolutionary distances between the HVR1 sequences. The numbers at the forks are the bootstrap values (%). The HVR1 sequences are Fal88a (Mar'88), Fal91b (Apr'91) and Fal91d (Nov'91).


Fal91b
74
Fal91d
Fal88a
Hcv-1
0
0.06

emerge in and after 1992. Thus, it appears that the strains which emerge and become predominate in 1992 do not arise by continuous mutation of the 1991 strains. Bootstrap testing of the tree shows that the divergence between the 1991 and 1992 HVR1 sequences is significant, the bootstrap value was 95%. The HVR1 sequences detected after 1991 appear to be continuously evolving, with newly emerging HVR1 variants being similar to previously dominant populations. Although the 1993 and 1994 HVR1 sequences appear to diverge from the 1992 sequences, the bootstrap value was only 53%, therefore this divergence is not significant. The 1993 variants could have equally evolved from the 1992 variants. In patient FAL the divergence shown between the HVR1 sequence detected in 1988 and that detected in 1991 was significant, the bootstrap value was 74%. The distances observed in patient SF were not significant for the majority of sequences, the highest bootstrap value was being 57%. The exception was the 1992 and 1993 (sample I) samples where the bootstrap value was 98%. This suggests that the HVR1 sequence variant detected in December 1992 and January 1993 was previously present as a minor population in the viral quasispecies and due disequilbrium it became predominant. Interpretation of such results is controversial, since for one to truly know whether continuous or discontinuous evolution has taken place it will be necessary to analysis other regions of the genome.

4.9 Discussion

4.9.1 The effectiveness of direct sequencing in analysing HCV sequence variants

To analyse the nucleotide sequence change in the E2 HVR1 of renal dialysis patients infected with HCV, direct sequencing of amplified PCR products without limiting end point dilution was employed. Direct sequencing was used in preference to sequencing cloned PCR products because there is the advantage that PCR derived errors due to the misincorporation of nucleotides by *Taq* polymerase (estimated to incorporate an inaccurate nucleotide once every 9000 nucleotides polymerised) during amplification of the product are not detected. Such errors usually occur infrequently and at random positions in the amplified product, therefore, sequencing a large population of molecules present in the amplified product reduces the chances of detecting such an error. When sequencing cloned PCR products several recombinants have to be sequenced to generate a consensus sequence to eliminate any errors which may have been introduced during amplification, since each recombinant is the product of a single DNA molecule from the PCR product. The other disadvantage of generating sequence data by this approach is that more manipulation steps are required to produce the recombinants for sequencing. Cloning can be a relatively inefficient technique, therefore the frequency at which a particular recombinant is found may not relate to the frequency at which that sequence is present in the population in the amplified product.

RNA viruses are prone to mutations due to the lack of proof-reading activity of RNA-dependent RNA polymerases. As a result, RNA viruses form quasispecies populations that contain a predominant or master sequence and a swarm of closely related genetic variants (reviewed by Steinhauer and Holland (1987); Holland *et al.*, 1992). Possible consequences of a quasispecies virus population are antigenic change allowing evasion of the host immune system and the development of resistance to antiviral drugs through genomic change. Martell *et al.* (1992) described the

quasispecies nature of HCV and suggested that it may account for the tendency of infection with this virus to become persistent.

In the work presented here, the nucleotide sequence of the HVR1 of HCV genomes are those detected in the serum from each of the three patients studied. It could be argued that this sequence may not be the same as that causing disease in the liver though Sakamoto *et al.* (1995) has provided evidence that the HCV quasispecies detected in the plasma and liver of patients with chronic hepatitis C at the same time point are the same in terms of population and diversity. Plasma and liver samples from each patient were obtained simultaneously. By nucleotide sequencing and SSCP analysis they showed that in all eight patients studied, the predominant HVR1 sequence and quasispecies detected in the plasma was identical to that detected in the liver. The study only examined plasma and liver samples at a single time point in each patient. To provide evidence that a change in the quasispecies population in the plasma (or serum) is reflected in that infecting the liver and causing damage, it is necessary to examine serial samples from each patient. To do this would have required several liver biopsies being obtained from each individual patient, but in performing a liver biopsy there is an element of risk to the patient. Therefore, from an ethical viewpoint such a study may not be feasible.

From sequence data obtained from the three patients in this study, the quasispecies nature of HCV was observed by the presence of mixed base positions on the sequencing gels. Direct sequencing distinguishes the predominant sequence from minor sequences according to the relative intensity of each base signal (band) on the sequencing gel. However, the sensitivity of detection of minor species is poor. So, although the predominant sequence at the second or third time point was not demonstrated previously, this does not mean it did not pre-exist as a minor population. The HVR1 variant sequence found to be predominant could have evolved directly from the previous predominant sequence through continuous mutation or could have been present previously as a minor population, outgrowing the previously predominant sequence because it had a selective advantage.

The use of SSCP gel electrophoresis analysis allows the detection of minor HCV quasispecies populations as well as the major one. SSCP analysis is based on the principle that the electrophoretic mobility of a molecule in a gel is sensitive to the size, charge and shape of the molecule. In non-denaturing conditions single-stranded DNA has a folded structure that is determined by the intramolecular interactions and hence, its sequence. The technique allows the detection of a single base change in short DNA fragments, since any change in sequence which affects the folding of the DNA is detected as a mobility difference on electrophoresis. A quasispecies virus population can therefore be separated into different bands by SSCP with each band representing a different quasispecies. The usefulness of SSCP in the analysis of HCV quasispecies has been shown in papers by Enomoto *et al.* (1994) and Kurosaki *et al.* (1994). In both experiments the authors used asymmetric PCR to generate single stranded products, avoiding the need to denature double stranded products.

When SSCP was attempted in this work it was unsuccessful. The HVR1 products analysed appeared to migrate with different mobilities but due to poor resolution of the bands, the number of coexisting quasispecies in each sample could not be determined. The sensitivity of SSCP can be affected by several factors, such as polyacrylamide concentration, degree of gel cross-linking ratios, temperature at which gel electrophoresis is performed and the presence or absence of glycerol. Time did not allow for experimentation to discover optimal conditions for SSCP analysis of HVR1 PCR products. Interestingly, when the 5'NCR products were subjected to SSCP the single stranded DNA bands migrated with the same mobility for both the type 1 and type 3 isolates. One would have expected them to migrate with different mobilities due to nucleotide differences in their sequence. However, the mobility of single-stranded DNA is dependent upon the folded structure adopted by it in non-denaturing conditions. Therefore, if both type 1 and type 3 isolates adopt the same folded structure, regardless of differences in nucleotide sequence, they should migrate with the same mobility in the gel. The 5'NCR of HCV is predicted to form an extended stem-loop structure for the region between nucleotides -208 and -58 which is conserved

among the HCV genotypes (Brown *et al.*, 1992; Smith *et al.*, 1995). It is this region of the 5'NCR which has been amplified. An alternative explanation for the results obtained is that the method failed to work because the optimum conditions were not used.

4.9.2 Emergence of HVR1 sequence variants

The results from this study showed sequence variation of the HVR1 predominant virus population during the natural course of infection in all three renal dialysis patients. The substitution rate of HVR1 was found to vary between the three patients and between different sampling points within the same patient. The substitution rate represents the degree of nucleotide difference between the predominant virus populations obtained at different time intervals. The substitution rates displayed by patient KR ranged from 0.21 to 7.9 x 10^{-2} nucleotide change/site/year. Those for patient SF ranged from 0.29 to 6.3 x 10^{-2} nucleotide change/site/year. For patient FAL the substitution rate calculated for between March 1988 and April 1991 was 4.4 x 10^{-2} nucleotide change/site/year, and that between April 1991 and November 1991 was 1.4 x 10^{-2}. These substitution rates are lower than those reported by Kurosaki *et al.* (1993). They found substitution rates ranging from 1.54 to 2.24 x 10^{-1} nucleotide change/site/year for patients displaying fluctuations in their ALT levels ("flare-up" type patients). In patients whose ALT levels remained within normal limits with little or no fluctuation ("quiescent" type patients) substitution rates ranged from 0.13 to 1.21 x 10^{-1} nucleotide change/site/year. The difference in substitution rate between these two groups was significant according to the Mann-Whitney U test ($p<0.01$).

Adopting the categorisation method of Kurosaki *et al.* (1993) patient KR would be classified as having "flare-up" type hepatitis and patient SF as having "quiescent" type hepatitis. However, no significant difference was found in substitution rates between these two patients, according to the Mann-Whitney U test ($p=0.05$). This discrepancy in results between this study and those of Kurosaki *et al* is probably due to

sampling differences. Kurosaki *et al* examined four patients with "flare-up" type hepatitis and eight with "quiescent" type hepatitis. Each patient was examined at two time points, the time interval between samples ranging from 38 to 50 weeks. Although only one patient from each of the above mentioned categories was examined in the study presented here, each patient was sampled at several time points. The time between samples ranged from 4 to 44 weeks for patient KR and from 4 to 56 weeks for patient SF. A larger sample size examined at several time points would be necessary to confirm that the substitution rate between patients displaying fluctuating ALTs and those displaying normal or near normal ALTs is not significantly different. Differences in the substitution rate of the HVR1 domain amongst patients in the acute phase of infection have also been observed; rates ranged from 0.14 to 0.91 nucleotide change/site/year (Sakamoto *et al.*, 1994). These authors suggested substitution rates increased with progression of infection. However, the findings of this study and those of Kurosaki *et al.* (1993) are contrary to this suggestion.

The data from this study show that during the chronic stage of infection the substitution rate varies. Periods of variation are interspersed with periods of no change, lasting up to several months. Not only does the substitution rate for the HVR1 differ between patients and isolates but other regions of the genome mutate at different rates (Ogata *et al.*, 1991; Okamoto *et al.*, 1992b). The differences in substitution rate between patients and different regions of the viral genome may reflect different error rates of the HCV RNA-dependent RNA polymerase or different selective pressures. It is interesting that analysis of part of the NS5a gene showed the substitution rate to range from 0.12 to 1.62 x 10^{-2} to 2.68 to 3.11 x 10^{-2} nucleotide change/site/year for patients KR and FAL, respectively for three isolates (Data for the NS5 gene were kindly given by Dr June Munro).

The low transition to transversion ratio and the preference for nonsynonymous substitutions in all three patients examined suggests that the HCV HVR1 domain is under selective pressure to mutate. Nucleotide variation occurred predominantly in the first and second codon positions. The majority of nucleotide changes resulted in a

change in the deduced amino acid sequence. The average ratio of synonymous to nonsynonymous substitutions for patients KR, SF and FAL are 0.05, 1.43 and 0.28, respectively (results are summarised in Table 4.7). These are comparable to those found by Kurosaki *et al.* (1993) who found values between 0 and 0.86 for "flare-up" type hepatitis patients and between 0.41 and 1.43 for "quiescent" type hepatitis patients. These values are comparable with the value of 0.67 found for the gp120 V3 domain of the HIV-1 (Simmonds *et al.*, 1990b). A low ratio implies selective pressure for change. Overall, mutations in the HVR1 which survive result in a change in amino acid. The ratio of transitions to transversions was 1.1 for patient KR and 1.8 and 1.6, respectively for patients SF and FAL. Tanaka *et al.* (1992) also found the ratio of transitions to transversions to be low in the HVR1, less than 1.0. Normally transitions are much more frequent than transversions, between closely related sequences transition to transversion ratios of 5 and 6 were observed for the E1 and NS5B regions of the genome, respectively (Smith *et al.*, 1997). The low transition to transversion ratio also suggests increased infidelity of the RNA polymerase at this region during replication, the HVR1 domain may fold in such a way that the RNA polymerase cannot read across it and so jumping occurs. The E2 HVR1 domain may be under no structural constraints to remain conserved since envelope proteins are on the surface of the virus. Thus, nucleotide changes within this region are more likely to be tolerated. In patients KR and SF the majority of amino acid changes are equivalent in terms of shape and charge of the amino acid. However, there are some which are not. In patient KR between samples C and D, three of the eight amino acid changes resulted in a change in charge, one resulted in the replacement of a larger amino acid with a smaller one (Ala-393 to Gly) and one (Ala-400 to Thr) resulted in the replacement of a hydrophobic amino acid with a hydrophilic amino acid. Unlike hydrophobic amino acids, hydrophilic amino acids are found at the surface of the protein, where they can form binding sites for charged molecules. In patient SF amino acid position 405 is variable among isolates. The amino acid changes which occur at this position are Thr to Thr/Ile to Ala to Thr/Pro to Pro to Thr to Pro. Unlike all other amino acids proline has a cyclic

Table 4.7 Nucleotide diversity of HVR1

Patient	Number of codons compared	Sample size	SYN[a]	NON[b]	Ratio NON : SYN
KR	28	34[c]	0.0037 (± 0.0072)	0.0711 (± 0.0152)	1 : 0.05
SF	28	25[c]	0.0599 (± 0.0285)	0.0418 (± 0.0143)	1 : 1.43
FAL	28	6[d]	0.0255 (± 0.0258)	0.0900 (± 0.0290)	1 : 0.28

[a] SYN denotes synonymous. The average number of synonymous substitutions per synonymous site and it's standard error were calculated using MEGA computer analysis package.

[b] NON denotes nonsynonymous. The average number of nonsynonymous substitutions per nonsynonymous site and it's standard error were calculated using MEGA computer analysis package.

[c] Analysis included sequences representing all the mixed base pair positions of the HCV isolates.

[d] Sequences from April 1991 were excluded from the analysis. The analysis included sequences representing all mixed base pair positions from the two remaining isolates.

ring structure, which creates a kink in the polypeptide chain. Using computer analysis packages Taniguchi *et al.* (1993) predicted the secondary structure of the E2 protein. Change in the secondary structure of the HVR1 domain was found to occur during infection, no change in secondary structure occurred in the remainder of the E2 protein despite amino acid substitutions.

A specific viral protein or domain of a protein may tolerate substitutions more than others for several reasons. The substitutions may be equivalent in terms of shape and charge of the amino acid or the protein may require the amino acid only as a spacer. Thus, the substitution does not affect the overall conformation or function of the protein. Alternatively, the substitutions may be beneficial. They may allow escape from immune recognition or they may confer an advantage on the virus by changing a feature of the protein, which in turn may result in the virus persisting in the host.

The results from patient KR show the possible creation of a new potential N-linked glycosylation site in sample M (April 1994) at amino acid position 414 just outside HVR1. Since a mixture of two bases is observed at this position it is necessary that either cloning or limiting dilution be performed on this sample. This will allow the base that predominates at this position to be determined with a degree of clarity. Surface proteins, like the E2 of HCV, are normally glycosylated. The glycosyl groups may be important for physical properties of the protein such as solubility or they may be involved in interactions with other cells, immunoglobulins or cell surface receptors. In the case of HCV, the creation of new potential glycosylation sites could be one mechanism by which the virus persists. The attachment of a glycosyl group to a protein may result in the masking of an HVR1 epitope by preventing the binding of a host neutralising antibody.

On initial infection with HCV in July 1990, patient KR displayed a rise in ALT which subsequently declined to within normal limits only to rise again in October. It is possible that this second elevation in ALT coincides with the recognition of the virus by the patient's immune system, since antibody to HCV first became detectable in October. It is possible that on infection the virus population is relatively homogenous, becoming

heterogeneous after seroconversion. This has been shown for HIV-1. On infection with HIV-1 and until seroconversion the virus population appears to be relatively homogenous in the V3 domain of the envelope protein gp120. Strong selective pressure is believed to account for this homogeneity (Simmonds *et al.*, 1990b; Zhang *et al.*, 1993). Unfortunately the HVR1 from 1990 to 1991 sera could not be amplified to test this hypothesis, despite the 5'NCR having been amplified in earlier testing by a previous worker in 1992. Difficulty in amplifying other regions of the HCV genome extracted from these samples was also experienced. The reasons for this are not known. It was originally thought that serum storage conditions may have affected the stability of HCV RNA but the results of work described in Chapter 5 make this unlikely.

The data show periods during which the predominant strain of HCV remains unchanged i.e. no sequence variant emerges. This is most notable in patient SF. No change in the HVR1 predominant sequence was detected over a 25 month period (June 1988 to July 1990). This suggests that HVR1 sequence variation is not necessary for persistence, as had previously been noted by Kumar *et al.* (1994) and van Doorn *et al.* (1995). However, in contrast to these studies, in patient SF, such persistence was noted in the presence of a detectable humoral immune response. Kumar *et al.* examined the HVR1 domain of an agammaglobulinemic patient infected with chronic hepatitis C at three time points over a 30 month period. No changes were detected in the predominant HVR1 sequence. In the study by van Doorn *et al.*, the HVR1 predominant sequence from a chronically infected chimpanzee remained unchanged for six years. During this period the chimpanzee was negative for HCV antibodies. With the appearance of anti-HCV antibody a HVR1 sequence variant emerged. The conclusion drawn from these studies was that the lack of sequence change in the HVR1 domain was associated with the absence of a humoral immune response. Therefore, sequence variation of the HVR1 is driven by humoral immune selection pressure. However, patient SF displayed a period of no change in the HVR1 predominant sequence comparable to that found by Kumar *et al.* This observation in an immunocompetent

individual casts doubt on their conclusion that the selective pressure driving HVR1 variation is the host's humoral immune response, and indeed, several studies discussed in the next section have documented variation in the absence of immune selection.

4.9.3 Variation in the absence of immune selection

Antigenic variation and the emergence of immunological escape mutants as a strategy for viral persistence, has been proposed for a several viruses, notably influenza virus and the lentiviruses. Influenza A virus has the ability to undergo both antigenic shift and drift resulting in the emergence of distinct viruses which are not recognised by a previously infected host. Mutations in the surface proteins haemagglutinin (HA) and neuraminidase (NA) appear to account for this antigenic variability. Antigenic shift is responsible for influenza virus pandemics. A new influenza virus appears in the population with different surface antigens. The first influenza A virus isolated in 1933 was of the subtype H1N1, in 1957 a new subtype H2N2 (Asian flu) emerged and in 1968 a third subtype H3N2 (Hong Kong flu) emerged, with the subtype H1N1 (Russian flu) reappearing in 1977. Antigenic drift is the result of accumulation of point mutations in the viral surface antigens (HA and NA) and appears to be due to immune antibody selection (reviewed by Wiley & Skehel., 1987; Webster *et al.*, 1992).

Similar findings to those of influenza virus have been reported for the lentiviruses, Visna virus, equine infectious anaemia virus (EIAV) and HIV-1 (reviewed by Clements *et al.*, 1988). The emergence of variants of these viruses occurs in a single host rather than in a population. EIAV infection is characterised by sequential episodes of viraemia associated with haemolytic crises during persistent infection. Neutralisation assay studies have demonstrated that virus isolates recovered from different episodes of viraemia were not susceptible to neutralisation by serum samples taken prior to their appearance. Each episode of viraemia was associated with a unique predominant virus population (Salinovich *et al.*, 1986). The extensive sequence variation displayed by HIV-1 may also be the result of immune selection pressure exerted by the host against

the V3 domain of gp120. This is supported by the generation of neutralisation escape mutants of HIV-1 (Nara *et al.*, 1990; Montefiori *et al.*, 1991).

However, evidence now exists from several systems suggesting antigenic variation is not necessarily the result of immune selection. Immunological escape mutants can occur in the absence of immune pressure. This was first demonstrated by work on foot-and-mouth disease (FMVD) (Diez *et al.*, 1989; Borrego *et al.*, 1993; Domingo *et al.*, 1993) Serial passage of plaque purified FMDV in BHK-21 cells resulted in the emergence of viruses with amino acid substitutions in antigenically important sites, as revealed by its altered reactivity with monoclonal antibodies (mAb). Since the virus was propagated in the absence of anti-FMDV antibodies, antigenic variation of FMDV occurred in the absence of immune selection (Diez *et al.*, 1989). This work was extended by Borrego *et al.* (1993). The same FMDV variants were passaged in the presence of a limited amount of neutralising polyclonal antibodies directed to the major antigenic sites of FMDV. Substitution of antigenically critical amino acids resulting in the loss of several epitopes within site A of the capsid protein VP1 was only observed in FMDV passaged in the presence of antibody. In the presence of neutralising antibody there is selection of a variant in the viral quasispecies with antigenically critical amino acid substitutions because it confers a selective advantage, however, in the absence of immune selection pressure such a variant would have a selective disadvantage and so would be kept at low levels in the quasispecies. Antigenic variation in the absence of immune selection pressure arises randomly by the selection of variants with substitutions in an antigenic site for properties unrelated to their antigenicity. The results suggest that antigenic variants of FMDV can emerge in the presence or absence of immune selection pressure.

The emergence of antigenic variants in the absence of immune selection pressure has not only been observed in FMDV. Serial passage of HIV-1 in cell culture in the absence of anti-HIV-1 antibodies resulted in amino acid substitutions in the V3 domain (Sanchez-Palomino *et al.*, 1993). Similarly, sequential influenza type A virus isolates recovered from a severe combined immunodeficient (SCID) child persistently infected

with the virus, revealed multiple point mutations in the HA1 domain of the HA glycoprotein. The observed amino acid changes in this domain occurred in or close to antigenic sites (Rocha *et al.*, 1991). Such observations mean that the emergence of an experimentally identifiable escape variant is not evidence per se of humoral selection pressure. This is further supported by the observation in EIAV that the sequential appearance of two antigenically distinct viral isolates was independent of the development of variant-specific neutralising antibody (Carpenter *et al.*, 1987). EIAV variation may be the result of cellular or nonimmunological selection pressures rather then humoral immune pressure. Therefore, it is possible that this may also be true for HCV variation.

4.9.4 Evidence for immune selection of HCV

The low transition to transversion ratio and nonsynonymous to synonymous ratio observed in the HVR1 of chronically infected HCV renal dialysis patients suggest this region of the genome is under pressure to mutate. If so, then what pressure is driving this change ? Several groups have suggested humoral immune selection pressure (Weiner *et al.*, 1992; Kato *et al.*, 1993a, 1994; Taniguchi *et al.*, 1993; Kojima *et al.*, 1994; van Doorn *et al.*, 1995), but the results from patient SF suggest that other selection pressures may be responsible. Patient SF was sampled at eleven time points over a six year period from 1988-1994. The majority of other studies have examined only about three samples over a similar period, therefore, periods of no change may have gone undetected. According to the immune selection hypothesis, antibodies to epitopes contained in the HVR1 of E2 cause positive selection of escape mutants. An antibody escape mutant would be expected to be a variant which emerges in the viral population and is then maintained in the absence of antibodies which recognise that particular sequence.

The HVR1 has been shown to contain linear B-cell epitopes, in particular between amino acids 394 and 407. Evidence for antibody escape mutants has come from studies

showing that antibodies present in patient sera to a specific HVR1 predominant variant early in the course of HCV infection, fail to bind to the HVR1 variant which subsequently emerges as the predominant viral population (Weiner *et al.*, 1992; Lesniewski *et al.*, 1993; Taniguchi *et al.*, 1993; van Doorn *et al.*, 1995). All these studies have used solid-phase HVR1-peptide binding assays because of the relative ease with which antibody reactivity to them can be determined. It is possible that the peptide may bind to the plate in such a way that key antigenic amino acids at the plate-peptide interface are no longer recognised by antibodies present in patient sera. Indeed, studies with HIV-1 have revealed limitations in use of such assays. It was found that reactivity of antibody to a solid-phase V3-peptide does not ensure reactivity to that V3 sequence in its native structure either because binding is blocked or the peptide sequence assumes a nonreactive conformation in the native form. Conversely, antibodies that react to conformation-dependent epitopes may not react to the corresponding linear peptide (Moore *et al.*, 1994). For these reasons the results from solid-phase peptide binding assays using HCV HVR1 derived peptides on their own cannot be considered as conclusive evidence in favour of humoral immune selection. Moreover, data from studies in which a branched peptide based solid-phase binding assay was employed provide evidence that sequence variants which arise during HIV-1 and HCV infection may not represent antibody escape mutants. All the afore mentioned studies used monomeric peptides for their assay.

Work performed by Dr Carol Robertson (Ph.D.thesis) and Dr Howard Marsden, Institute of Virology, using a branched peptide based assay showed that high titres of antibody were found against most of the V3 sequence variants that emerged in an HIV-1 infected haemophiliac patient before the variant was detected in the viral population. Despite the presence of antibodies that recognised the sequence of V3 variants, the variants were found to persist within the patient. If the V3 variants which emerged during infection represented antibody escape mutants, one would expect antibody against them to be present only in sera taken after they emerged. This has now also been observed for HCV (Scarselli *et al.*, 1995). Using a branched peptide based assay,

antibodies recognising a specific HVR1 sequence variant which emerged in a chronically infected individual were found to precede its appearance. Coexistence in the bloodstream of the variant in the presence of increasing amounts of specific antibodies for seven months was also observed. Unfortunately, due to problems experienced in amplifying the HVR1 from HCV RNA extracted from patient sera, there was insufficient time to undertake a similar study using a branched peptide based assay with regards to examining the HVR1 sequence variants which emerged during natural infection in the three renal dialysis patients described here. Such a study would give a better indication of the role of humoral immune selection in HVR1 sequence variation because it would allow more samples to be looked at.

A different approach was adopted by Kato *et al.* (1993a, 1994) to examine the humoral immune response to HVR1. They examined the reactivities of anti-HVR1 antibodies in an infected patient to substituted amino acids sequences within two B-cell epitopes. They found that HVR1 variants in both epitopes identified within the HVR1 escaped from anti-HVR1 antibodies that were pre-existing in the patient's serum. To test for reactivities of anti-HVR1 antibodies, an expression plasmid was constructed and used to express a fusion protein with HVR1 variants and the dihydrofolate reductase gene derived from *Escherichia coli* by *in vitro* transcription and translation. Immunoprecipitation of the expression fusion proteins with patient sera collected at various times was then performed. The expression system used does not allow for post-translational modification of the protein, which can have a major effect on the structure or immunogenicity of the protein. Therefore, the expressed HVR1 domain may not show antigenicity similar to that in the native E2 protein. It could be argued that because the HVR1 domain itself does not appear to contain potential glycosylation sites and lacks conserved secondary structure motifs, change may not interfere with conformation (Taniguchi *et al.*, 1993). So, the expression solely of the HVR1 domain is adequate to test for reactivities of anti-HVR1 antibodies. However, secondary structure predictions using computer programs may not take into account the effect of post-translational modifications or contacts made with other proteins (or as dimers,

trimers or other polyproteins). Studies on HIV-1 have shown that antigenicity of the oligomeric viral envelope protein are distinct from the monomer (Broder *et al.*, 1994). In its native form, the E2 protein may adopt a conformation that prevents the binding of antibodies to epitopes in the HVR1 domain which are seen to bind to the HVR1 when it is expressed as a "bare" polypeptide. The optimal way to demonstrate that variation in the HVR1 is the result of selection by the humoral immune system would be to detect the presence of antibodies using as the antigen E1/E2 protein complexes containing the appropriate variations of HVR1 and processed in an appropriate system, reflecting what would occur in an infected hepatocyte. To determine the antibody binding sites on the E1/E2 proteins, a series of expressed proteins increasing in size should be generated.

Kojima *et al.* (1994) studied the influence of humoral immune selection on the emergence of HVR1 sequence by passaging a human plasma infected with HCV through eight chimpanzees in three generations and examining the HVR1 in acute-phase plasma from each chimpanzee. They found that one clone (clone A) detected in the original inoculum was selected for in chimpanzees during three passages. Antibody to clone A was detected in only one chimpanzee, whereas antibody to clones B and C were detected in two chimpanzees. Plasma from all three infected chimps was pooled to create inoculum three, which contained antibody to all three clones. In the two chimpanzees who received this inoculum only clone A was found. The authors suggest that because antibody to clone A was not present in sufficient levels to neutralise it, clone A was selected, indicating positive selection by the humoral immune system. However, it is possible that clone A was selected because it was stable on passage of the virus. During passage, mutations in other regions of the genome may render the other clones in the viral population avirulent. This could explain why they were not detected in the two chimpanzees infected on the third passage.

Such studies as those discussed above do not directly prove the variation observed in the HVR1 domain is driven by humoral immune selection or that the sequence variants which emerge represent antibody escape mutants. Moreover the study by Scarselli *et al.* provides evidence in support of the hypothesis that HVR1

sequence variants do not represent antibody escape mutants. Further studies will need to be conducted to settle this contentious issue.

4.9.5 Mechanisms of liver damage and persistence in chronic hepatitis C

The mechanisms responsible for liver damage in chronic hepatitis C are unclear. The virus may be directly cytopathic or hepatocyte damage may be immune mediated. If damage is due to a direct cytopathic action of the virus, a new genetic variant would emerge before an elevation in ALT. Alternatively, if damage was as a result of the destruction of infected hepatocytes by the host's immune system, a new genetic variant would appear some time after an elevation in ALT.

Throughout the period of observation patient KR's ALT levels fluctuated markedly and new genetic variants were seen to emerge. The first HVR1 sequence variant to be detected was in sample B (March 1991), which emerged during a period when the ALT levels were elevated (> 196U/L). In contrast the sequence variants detected in samples C (October 1991) and D (August 1992) emerged during periods when the ALT levels were within normal limits. The sequence variant detected in sample G (January 1993) emerged after an elevation in ALT in October/November 1992 (>97U/L) whereas those detected in samples J (October 1993) and K (November 1993) emerged before an elevation in ALT in December 1993 (124U/L). The results from patients SF and FAL are similar. In patient SF the sequence variant detected in sample E (July 1990) emerged during an elevation in ALT (99U/L). After the elevation in ALT in July 1991 a sequence variant was detected in sample G (October 1991). Apart from these two sequence variants all the others are detected during periods when the ALT levels are within normal limits. It is possible that the sequence variant detected in sample A (March 1988) from patient FAL is a new genetic variant which has emerged before the elevation in ALT in May 1988 (86U/L). The sequence variants detected in samples G (April 1991) and I (November 1991) emerge before and after a small elevation in ALT (62U/L), respectively. However, it is possible that the

predominant sequence detected in these two samples is the same. HCV isolates from patient FAL proved difficult to sequence by the direct method used here so only three isolates were analysed. The sequence ambiguities detected in the gels from patient FAL and also from patients KR and SF, may be resolved by performing SSCP electrophoresis or limiting dilution on the amplified products.

The sequencing gels of patient FAL suggest that within the quasispecies population there exists one which contains a deletion(s) or insertion(s). This could result in the generation of a defective genome. Cocirculation of defective viral genomes in HCV infection has been documented (Martell *et al.*, 1992; Higashi *et al.*, 1993). This may be another mechanism utilised by HCV to allow persistence within the host since defective viruses are usually poorly cytopathic. Alternatively, the presence of a insertion and /or deletion may result in a frame shift mutation being generated, the result of which is an antigenic escape variant. Antigenic escape mutants in the G protein of respiratory syncytical virus, which is involved in virus binding to target cells, have been shown to arise by frame shift mutations (Garcia-Barreno *et al.*, 1994).

In the study presented here, no direct association could be demonstrated between the emergence of new sequence variants and either mechanism of liver damage. The results from patient KR show that new genetic variants were detected before, during and after elevations in ALT. They also showed that variants emerged during periods when the ALT level was within normal limits. Patient SF in particular displayed this. By using ALT levels as a predictor of liver damage it is difficult to predict from the results whether the direct cytopathic nature of the virus or the host immune response to it is the mechanism responsible for liver damage. Indeed, the results suggest that the emergence of variants may not be related to a period of liver disease on the basis of using ALT levels as a marker of damage. However, it is possible that ALT levels are not a satisfactory predictor of liver damage.

If damage is due to a direct cytopathic nature of the virus, a high titre of neutralising antibodies (those that inactivate the virus and thus arrest the disease process) against the virus would be expected to be found to prevent infection of

hepatocytes. HCV infection has been shown to elicit neutralising antibodies in both humans and chimpanzees (Farci *et al.*, 1994; Shimizu *et al.*, 1994). Since in many viral infections neutralising antibodies are directed against envelope proteins, it is presumed that the type specificity of the neutralising antibodies may be directed against the HVR1 domain. In one study addressing this point antibodies specific to HVR1 variants were shown to block viral attachment to cells in an *in vitro* binding assay (Zibert *et al.*, 1995). Because the neutralisation assay employed in this study was based on the use of HVR1 specific fusion proteins expressed in *Escherichia coli an*d the detection of bound virus by PCR (PCR was also employed by Shimizu *et al.*), careful consideration must be given to the results, since PCR is a technique subject to variability within and between runs. A recent study has shown that neutralising antibodies directed towards regions of E2 protein result in protection from infection in vaccinated chimpanzees. Antibodies directed towards the HVR1 domain are found to be neutralising but are not necessary for protection. No difference in titre of neutralising antibodies was found between chronically infected patients and "healthy seropositives". Thus, HCV infection appears to elicit no or only a low titre of neutralising antibody (Rosa *et al.*, 1996). These findings question whether neutralisation antibody constitute a strong selection pressure. Thus, the HCV sequence variation, especially in the HVR1 domain, may not be driven by the humoral immune system of the host. They also suggest that liver damage in HCV infection may be due to the host's cellular immune response. Because neutralising antibodies are not produced or are only produced in low numbers, hepatocytes will not be protected from infection. Once the virus has infected the cell, it is immune to attack from antibodies but not from cytotoxic T-lymphocytes (CTLs). Persistence of the virus may be due to CTL escape mutants and not antibody escape mutants.

HCV-specific CTL within the liver lymphocytes of individuals infected with chronic hepatitis C have been observed (Battegay *et al.*, 1995; Shirai *et al.*, 1995). The CTL responses observed in chronic HCV infection are recognised in the context of multiple human leukocyte antigens (HLAs). CTLs are restricted by distinct HLAs,

therefore, they recognise different epitopes. CTL response is multi-specific, involving epitopes in most viral proteins that are presented by class I major histocompatability complex (MHC) allotypes. There appears to be no particular clustering of epitopes within a given protein. CTL epitopes been identified in core, E1, E2, NS2, NS3 and NS5b regions (reviewed by Walker, 1996). Yet, despite strong CTL activity, persistent infection can occur. It is possible that CTLs are not present in high enough numbers to eliminate HCV but are capable of causing damage via lysis of hepatocytes expressing HCV proteins, or CTL escape variants exist. Mutations within crucial CTL epitopes may affect MHC binding or T-cell antigen receptor (TCR) recognition. A CTL escape variant has been observed in a chronically infected chimpanzee. CTLs in the chimpanzee liver were able to recognise a conserved epitope in the NS3 protein. The same CTLs did not recognise the HCV quasispecies present at 16 weeks post-infection or at later time points. An aspartic to glutamic acid substitution at amino acid 1449 abrogated recognition (Weiner *et al.*, 1995). CTL escape may be important in persistence since the virus cannot integrate into the host genome. Direct evidence for the emergence of CTL escape mutants in humans infected with HCV is still lacking. CTL escape has been documented in HIV infection. Sequence variation in HIV *gag* CTL epitopes in HIV positive haemophiliac donors was found to lead to loss of CTL recognition (Phillips *et al.*, 1991). However, the emergence of CTL escape variants alone may not be sufficient to sustain HCV persistence. They may be of limited importance for persistence because CTL response is often directed against multiple epitopes in both variable and conserved regions of the viral protein. It is possible that suppression of presentation of MHC class I molecules may be involved in persistence, a mechanism of persistence employed by adenovirus (Oldstone, 1989).

Alternatively, cytokine unresponsiveness would be a more effective mechanism by which the virus could escape the host's immune response, since it would abolish the effect of the entire CTL population instead of just the effect of the CTL response generated against an individual CTL epitope. Recognition of the antigen/MHC class I complex is not sufficient to generate a T-cell response, cytokines are needed to induce

and amplify the T-cell response. If an antigen is presented to CTL by MHC class I molecules but the receptor for cytokines are not or are inefficiently expressed then T-cells are not able to proliferate, instead they will be unresponsive and so, the virus will be allowed to persist. In two patients with chronic HBV infection, the virus was found to mutate in such a way that it could still bind to the MHC class I molecule and the CTL but failed to activate the T-cell and thus, did not induce lysis of the infected hepatocyte (Bertoletti *et al.*, 1994). This has also been documented for HIV (Klenerman *et al.*, 1994; Meier *et al.*, 1995). Such mutations occur within epitopes which result in an alteration in the TCR contact residue on the infected cell. The variant epitopes can also act as TCR antagonists by inhibiting the CTL response to the wild-type epitope, since they retain the ability to bind to CTLs.

The results from patient SF and those of other studies suggest more than one mechanism may be involved in the persistence of HCV within the host. The viral envelope proteins, E1 and E2 of HCV are heavily glycosylated. Viral glycosylation sites have been shown to be conserved among different isolates despite mutation of surrounding sites (Ogata *et al.*, 1992; Okamoto *et al.*, 1992b). It is possible the heavy glycosylation of the envelope proteins masks the viral epitopes thus inhibiting the binding of host antibody. Site-specific antibodies to a peptide sequence of the influenza virus HA molecule containing a site of glycosylation improved in reactivity after removal of the carbohydrate moiety (Alexander & Elder, 1984). Thus, the carbohydrate portion of the glycoprotein can effectively block the interaction of antibodies with the underlying polypeptide regions. Alternatively, the association of HCV with low density lipoprotein (LDL) (Thomssen *et al.*, 1992, 1993) may inhibit or delay binding of host antibodies. Since HCV is an enveloped virus, one presumes that it buds from the liver cell during its life cycle, so it is possible that on budding from the liver cell some host LDL is incorporated into the virus envelope. Thus, the host immune system recognises the virus as "self" and does not mount an immune response, which may account for the low levels of neutralising antibodies in HCV infection. HCV density gradient analysis has shown circulating HCV in chronically infected patients to be found in both a low

and high density fraction. The low density fraction contains free infectious virus while the high density fraction contains virus complexed to antibody (Kanto *et al.*, 1995). With progression of liver disease there appears to be a shift from low to high density dominance. Such virus/antibody complexes may be a consequence of HCV neutralisation. Alternatively, such virus/antibody complexes may represent binding of the virus to non-neutralising antibodies. This would prevent the binding of neutralising antibodies and allow persistence of the virus. In a persistent infection such virus/antibody complexes may deposit over a long period of time producing chronic inflammation.

4.9.6 Serum ALT levels as a marker of liver damage

How "good" a marker at predicting liver damage are serum ALT levels ? A limitation in using serum ALT levels as a predicator of liver damage due to HCV is that an elevation in ALT is not a specific marker of HCV infection. ALT levels are a marker of liver damage in general and thus an elevation may occur for several reasons. Studies have shown that not all HCV infected individuals have elevated ALTs, many individuals displaying normal ALT levels have chronic hepatitis C (Bruno *et al.*, 1994; Prieto *et al.*, 1995; Shindo *et al.*, 1995). More than half of all viraemic patients have normal ALT levels, so the normalisation of ALT levels does not necessarily mean that no liver damage is occurring. Serum ALT levels are also poor markers of the severity of liver disease associated with chronic HCV infection. At present performing a liver biopsy is the only way to evaluate the severity of liver disease (Lemon & Brown, 1995). Due to the associated risks associated with liver biopsies it is not feasible to perform sequential liver biopsies to monitor liver damage. None of the three patients used in the study presented here had liver biopsies.

4.9.7 Future Studies

To provide clear evidence that sequence variants that arise in the HVR1 domain do not represent antibody escape mutants and that variation is not driven by humoral immune selection, it is important that this study be extended to examine the humoral immune response of the three patients. The study should also be extended to examine a larger cohort of patients. Before the humoral immune response of the patients can be examined, minor virus populations and the frequency at which they are detected at each time point will need to be determined either by SSCP or limiting dilution analysis. After this has been performed branched peptides corresponding to the HVR1 variants found during the natural course of infection in each patient can be synthesised to test the hypothesis that the sequence variation in the E2 HVR1 is driven by humoral immune selection. Alternatively, the humoral immune response to the E2 protein could be determined by developing an ELISA system which uses E2 protein which has retained its native form and contains the appropriate sequence variations of HVR1 as the antigen. To do this would require the envelope protein to be expressed in a system reflecting what occurs in an infected hepatocyte. By doing this antibodies to conformational as well as linear epitopes would be detected.

It is possible that sequence variation in the HVR1 domain is accompanied by variation in other regions of the E2 protein and possibly the E1 protein also, and that interaction with other regions of the protein are important for immune recognition. To investigate this, it would be necessary to sequence, clone and express the entire E1 and E2 proteins from each individual viral isolate and test it against autologous sera.

The amino acid changes observed in the HVR1 domain in each of the three patients could represent mutants escaping from CTL responses or such mutants could affect the MHC binding or TCR contact sites, thereby inhibiting CTL function. Indeed the HVR1 domain contains CTL epitopes (reviewed by Walker., 1996). Therefore, it is important not only to investigate the humoral response of each patient but also the CTL

response. To do this would require the generation of CTL lines from the liver of infected individuals.

Chapter 5 Results: Section 2

The Effect of Serum Storage Conditions on the Detection of Hepatitis C Virus by the Polymerase Chain Reaction.

The currently available commercial assays for detection of HCV rely on detecting antibodies to recombinant cloned antigens. First-generation assays detected antibodies to the c100-3, a fusion polypeptide representing nearly all of the NS4 region of the HCV genome (Kuo *et al.*, 1989). The presence of antibodies to C100 proved to be a marker of infection (Alter *et al.*, 1989) but it became evident that this assay lacked specificity and sensitivity. False-positive results were common when screening low risk groups (McFarlane *et al.*, 1990) The development of second- and third-generation assays, which have incorporated HCV recombinant antigens representing other regions of the genome, have increased the specificity and sensitivity of these serological assays but limitations in their use still exist. These assays are unable to distinguish between an ongoing or a resolved infection and there exists a considerable period of time between infection with the virus and appearance of detectable anti-HCV antibodies (Alter *et al.*, 1989; Farci *et al.*, 1991). Since no universally successful ELISA to detect HCV antigens has yet been developed, detection of the virus itself is based upon detection of HCV RNA.

The direct detection of HCV RNA in serum or plasma is the most sensitive and specific marker for HCV infection in both symptomatic and asymptomatic patients. HCV RNA can be detected in the blood within a few days of infection, weeks before an increase in liver enzyme levels or viral antibody levels are detected (Shimizu *et al.*, 1990). Because of the low titre of HCV in blood (10^2-10^7 particles per ml) (Simmonds *et al.*, 1990; Ulrich *et al.*, 1990) amplification of HCV RNA is necessary by the reverse transcription-polymerase chain reaction (RT-PCR). However, RNA is easily

degraded by contaminating RNases and therefore handling of specimens could greatly influence the results obtained by RT-PCR.

The ability to identify HCV infection by anti-HCV assays and detection of HCV RNA has resulted in numerous studies to evaluate the natural course of HCV infection. Most published studies are retrospective, relying on stored sera but few authors comment on the storage conditions. Problems experienced in attempting to amplify the hypervariable region of HCV from patient sera suggested that the storage conditions, especially the number of freeze-thaw cycles, may have affected the stability of the HCV RNA and hence the sensitivity of detection. The aim of this work was to study the effect of serum storage conditions on the detection of HCV RNA by RT-PCR.

5.1 Processing of serum samples

Fresh blood samples were collected from a known HCV positive renal transplant patient, infected with the virus since the early 1980's and an individual repeatedly negative for HCV RNA and anti-HCV antibodies. The blood was allowed to clot for 2hrs at 4°C before being centrifuged at room temperature. The serum was removed and aliquoted into 800µl volumes into three tubes. Tube 1 was placed at 4°C, tube 2 at -20°C and tube 3 was further aliquoted into 100µl volumes which were stored at -20°C for use in future experiments. Total time from collection of blood samples to obtaining serum was 2 hours 45 minutes. To estimate the titre of HCV RNA present, serial 10-fold dilutions of the serum in PBS were performed. RNA was extracted from 100µl of sera using the acid guanidium thiocyanate method described in Chapter 3, section 3.5. Tube 1 was sampled at 14, 17, 19, 25 and 27 weeks and tube 2 after 1, 3, 5, 7 and 10 successive freeze-thaw cycles (-20°C to room temperature).

5.2 HCV RNA RT-PCR Analysis

The extracted HCV RNA was reversed transcribed using random primer (pd(N)6) at 25°C for 10 minutes, followed by 37°C for 55 minutes, with a final 95°C incubation for 5 minutes to inactivate the reverse transcriptase enzyme. A 'nested' PCR reaction was performed, using two primer pairs from the highly conserved 5' non coding region of the HCV genome (outer pair, HCV21 and EHCV24 and inner pair, HCV23 and NCR4). Table 5.1 shows the sequence and position of the primers selected. Negative and positive controls were included in each round. Amplification was performed with an initial denaturation step of 4 minutes at 94°C, followed by 30 cycles of denaturation for 30 seconds at 94°C, annealing of primers for 45 seconds at 55°C and extension for 2.5 minutes at 70°C, followed by a final extension for 7 minutes at 70°C. Amplified products were viewed under UV light on a 2% (w/v) agarose gel, stained with ethidium bromide.

5.3 Effect of serum storage conditions on the detection of HCV RNA

To assess the effect of storage at 4°C and successive freeze-thawing of serum on the detection of HCV RNA, serial 10-fold dilutions of serum were performed and the viral RNA titre detectable by RT-PCR was determined. In the case of serum stored at 4°C the titre of viral RNA detectable was determined after only one run because the amount of serum stored was insufficient to perform a second run. For serum stored at -20°C and subjected to successive freeze-thawing the titre of viral RNA detectable was determined after two runs. Table 5.2 shows the results for the serum stored at 4°C. The titre of viral RNA detectable after 14, 17 and 27 weeks was 10^{-3}, 10^{-2} and $\geq 10^{-2}$ respectively, a 10-fold reduction in titre of HCV RNA detectable having occurred between 14 and 17 weeks, as shown in Figure 5.1. The detection of HCV RNA at a titre of $\geq 10^{-2}$ after 27 weeks was unexpected, so the serum was left for a further 30

Table 5.1 Oligonucleotide primers

Primer	Sequence 5' ⟶ 3'	Nucleotide position*
HCV21 (outer sense)	CGACACTCCACCATGAATCAC	-322 to -302
HCV23 (inner sense)	TCACTCCCCTGTGAGGAACT	-305 to -286
EHCV24 (outer antisense)	CATGGTGCACGGTCTACGAGACC	-20 to 3
NCR4 (inner antisense)	GCACCCTATCAGGCAGT	-54 to -38

* Numbering according to Choo *et al.* (1991).

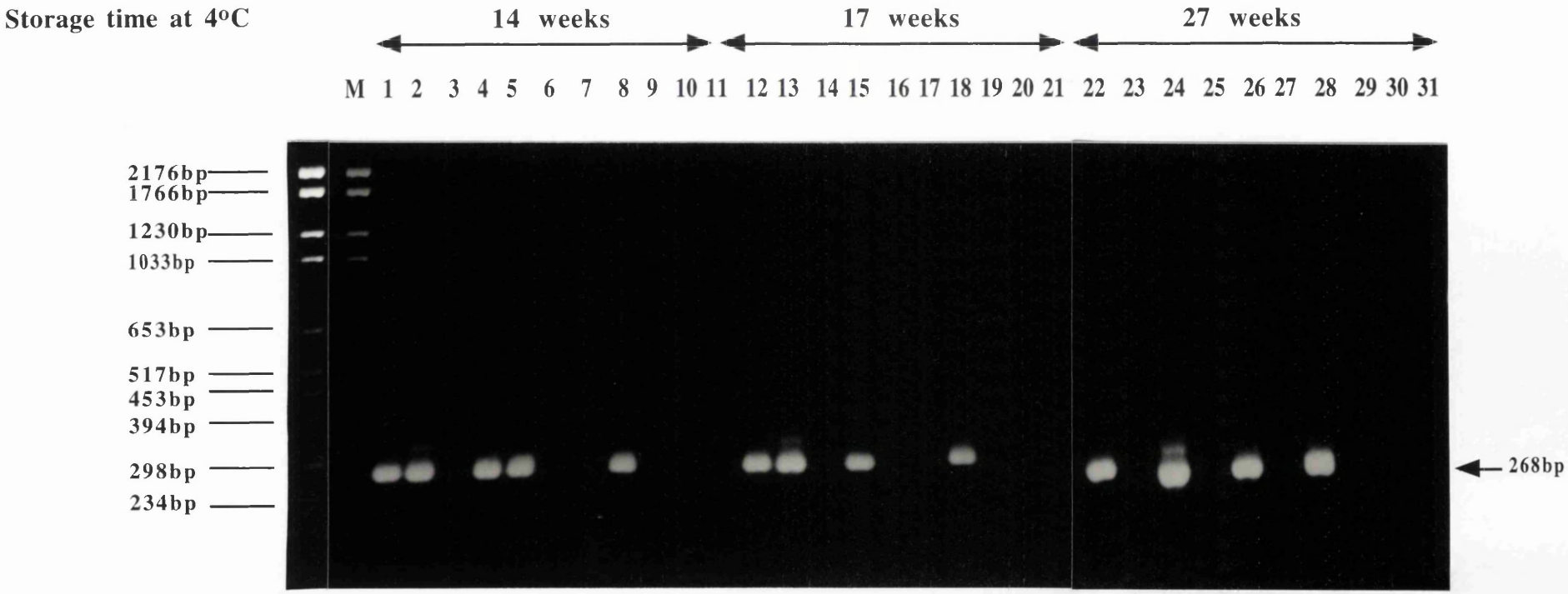
Table 5.2 HCV RNA Titres of serum samples stored at 4°C.

Storage time at 4°C	Titre
14 weeks	10^{-3}
17 weeks	10^{-2}
19 weeks	10^{-2}
25 weeks*	$\geq 10^{-2}$
27 weeks*	$\geq 10^{-2}$
57 weeks*	UD

* Denotes sample was not tested at a dilution higher than 10^{-2} because serum samples stored at 4°C for 17 and 19 weeks tested negative for HCV RNA at 10^{-3}. In hindsight such samples should have been tested at dilutions between 10^{-2} and 10^{-3} to determine the end-point dilution titre.

UD denotes undiluted serum sample.

Figure 5.1 Gel electrophoresis of PCR amplified cDNA from the 5'NCR of serum stored at 4°C. Serum was sampled at 14, 17 and 27 weeks. HCV RNA was extracted from serially diluted serum, reverse transcribed and amplified by PCR. 10µl of each product was run on a 2% agarose gel. Lanes 1, 12 and 22, undiluted serum sample; lanes 2, 13 and 24, serum sample diluted 1:10; lanes 3, 14, 23 and 25, extraction carry-over control; lanes 4, 15 and 26, serum sample diluted 1:100; lanes 5 and 16, serum sample diluted 1:1000; lane 6, serum sample diluted 1:10,000; lanes 7, 17 and 27, negative control serum; lanes 8, 18 and 28, positive control serum; lanes 9, 19 and 29, negative control for reverse transcription reaction; lanes 10, 20 and 30, negative controls for first round PCR reaction; lanes 11, 21 and 31, negative controls for second round PCR reaction. M, molecular weight marker (Boehringer Mannheim DNA molecular weight marker VI). An over-exposed molecular weight marker track is shown. Arrow denotes size of product.


Storage time at 4°C
14 weeks
17 weeks
27 weeks
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
2176bp
1766bp
1230bp
1033bp
653bp
517bp
453bp
394bp
298bp
234bp
268bp

weeks. HCV RNA in this serum sample, stored at 4°C for a total period of 57 weeks could be detected only in the undiluted sample (Figure 5.2), a 100-fold reduction in titre. In contrast a titre of 10^{-3} was detectable in a serum sample stored at -20°C for 57 weeks, a 1000-fold difference in titre (Figure 5.2).

Each dilution of the serum stored at -20°C and subjected to successive freeze-thaw cycles was performed in duplicate (i.e. two separate, extraction and RT-PCR runs were performed). Table 5.3 shows the titre of RNA detectable after 1, 3, 5, 7 and 10 freeze-thaw cycles. The titre detectable after 10 cycles was $\geq 10^{-2}$. An indeterminate result (positive/negative) was observed for the dilution of 10^{-4} after one freeze-thaw cycle (Figure 5.3). Serum from tube 3 (which had not been previously thawed) was therefore diluted 1:100, 1:250, 1:500, 1:750 and 1:1000 and subjected to 1 and 10 successive freeze-thaw cycles. Figure 5.4 shows the results of this work. The endpoint dilution after 10 freeze-thaw cycles was found to be 1:500 (approximately $10^{-2.3}$). Thus, there is at least a 2-fold reduction in titre between the first and tenth freeze-thaw cycle.

Figure 5.2 Gel electrophoresis of PCR amplified cDNA from the 5'NCR of serum stored at 4°C and -20°C for a period of 57 weeks. HCV RNA was extracted from serially diluted serum, reverse transcribed and amplified by PCR. 10µl of each product was run on a 2% agarose gel. Lanes 1, 2 and 4, serum sample stored at 4°C undiluted, diluted 1:10 and 1:100, respectively; lanes 3, 6 and 9, extraction carry-over control; lanes 5, 7, 8, 10 and 11, serum sample stored at -20°C undiluted, diluted 1:10, 1:100, 1:1000 and 1:10,000, respectively; lane 12, negative control serum; lane 13, positive control serum; lane 14, negative control for reverse transcription reaction; lanes 15 and 16, negative controls for first and second round PCR reactions. M, molecular weight marker (Boehringer Mannheim DNA molecular weight marker VI). An over-exposed molecular weight marker track is shown. Arrow denotes size of product.

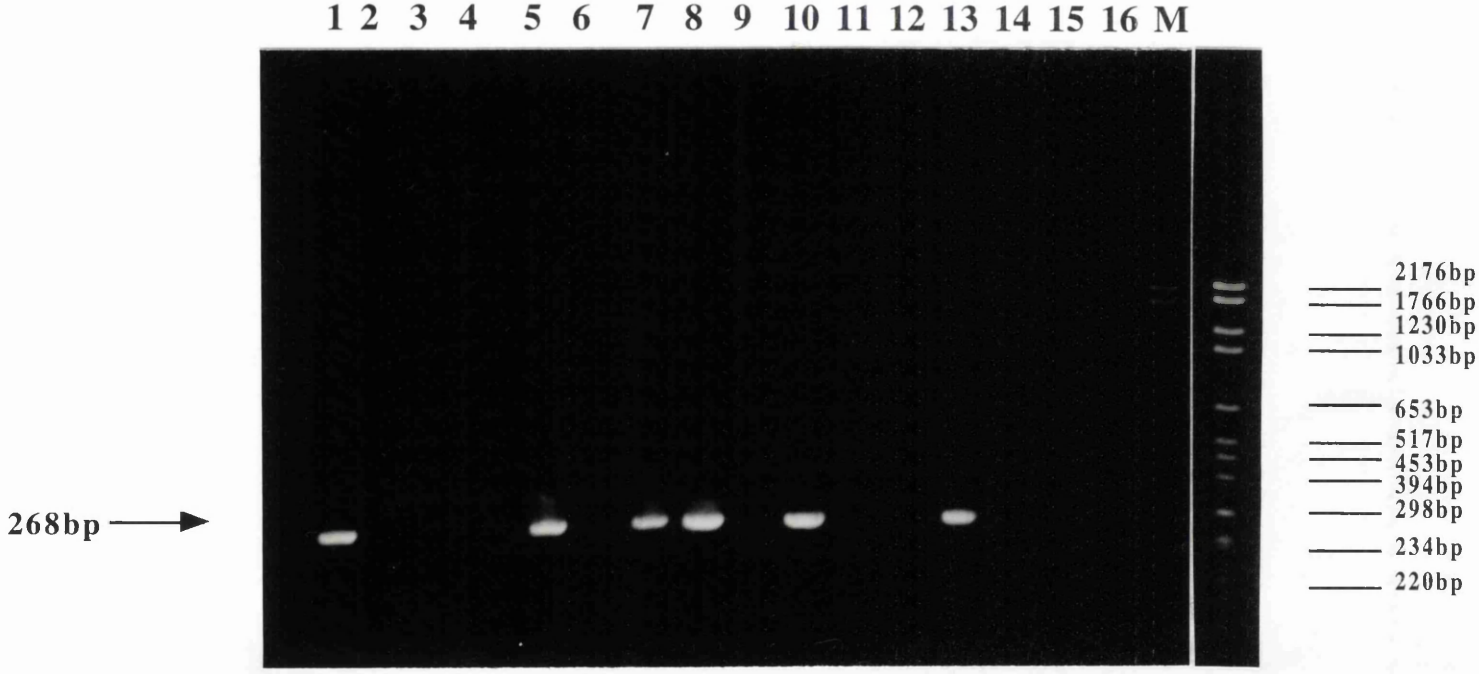

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 M
268bp
2176bp
1766bp
1230bp
1033bp
653bp
517bp
453bp
394bp
298bp
234bp
220bp

Table 5.3 HCV RNA Titres of serum samples subjected to freeze-thawing.

No. Freeze-Thaw Cycles	PCR Results +/+	+/-	-/-	Titre
1	10^{-3}	10^{-4}	10^{-5}	10^{-3}-10^{-4}
3	10^{-2}	NS	10^{-3}	10^{-2}
5	10^{-2}	NS	10^{-3}	10^{-2}
7*	10^{-2}	ND	ND	$\geq 10^{-2}$
10*	10^{-2}	ND	ND	$\geq 10^{-2}$

* Denotes sample was not tested at a dilution higher than 10^{-2} because serum samples tested negative for HCV RNA at 10^{-3} after 3 and 5 freeze-thaw cycles. In hindsight such samples should have been tested at dilutions between 10^{-2} and 10^{-3} to determine the end-point dilution titre.

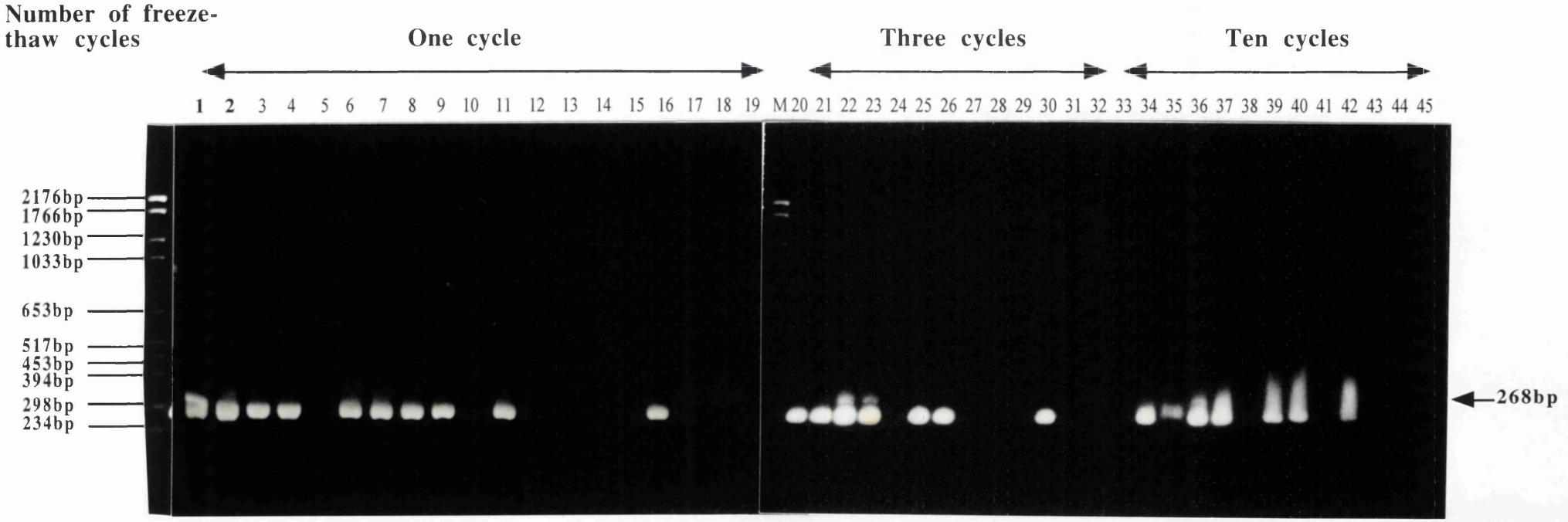
+/+ indicates both samples positive; +/-, one sample positive and the other negative;

-/- both samples negative.

NS denotes not seen.

ND denotes not done.

Figure 5.3 Gel electrophoresis of PCR amplified cDNA from the 5'NCR of serum stored subjected in duplicate to one, three and ten freeze-thaw cycles (-20°C to room temperature). HCV RNA was extracted from serially diluted serum, reverse transcribed and amplified by PCR. 10µl of each product was run on a 2% agarose gel. Lanes 1, 2, 20, 21, 34 and 35, undiluted serum sample; lanes 3, 4, 22, 23, 36 and 37, serum sample diluted 1:10; lanes 5, 10, 24 and 38, extraction carry-over control; lanes 6, 7, 25, 26, 39 and 40, serum sample diluted 1:100; lanes 8, 9, 27 and 28, serum sample diluted 1:1000; lanes 11 and 12, serum sample diluted 1:10,000; lanes 13 and 14, serum sample diluted 1:100,000; lanes 15, 29 and 41, negative control serum; lanes 16, 30 and 42, positive control serum; lanes 17, 31 and 43, negative control for reverse transcription reaction; lanes 18, 32 and 44, negative control for first round PCR reaction; lanes 19, 33 and 45, negative control for second round PCR reaction. M, molecular weight marker (Boehringer Mannheim DNA molecular weight marker VI). An over-exposed molecular weight marker track is shown. Arrow denotes size of product.


Number of freeze-thaw cycles
One cycle
Three cycles
Ten cycles
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 M 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
2176bp
1766bp
1230bp
1033bp
653bp
517bp
453bp
394bp
298bp
234bp
268bp

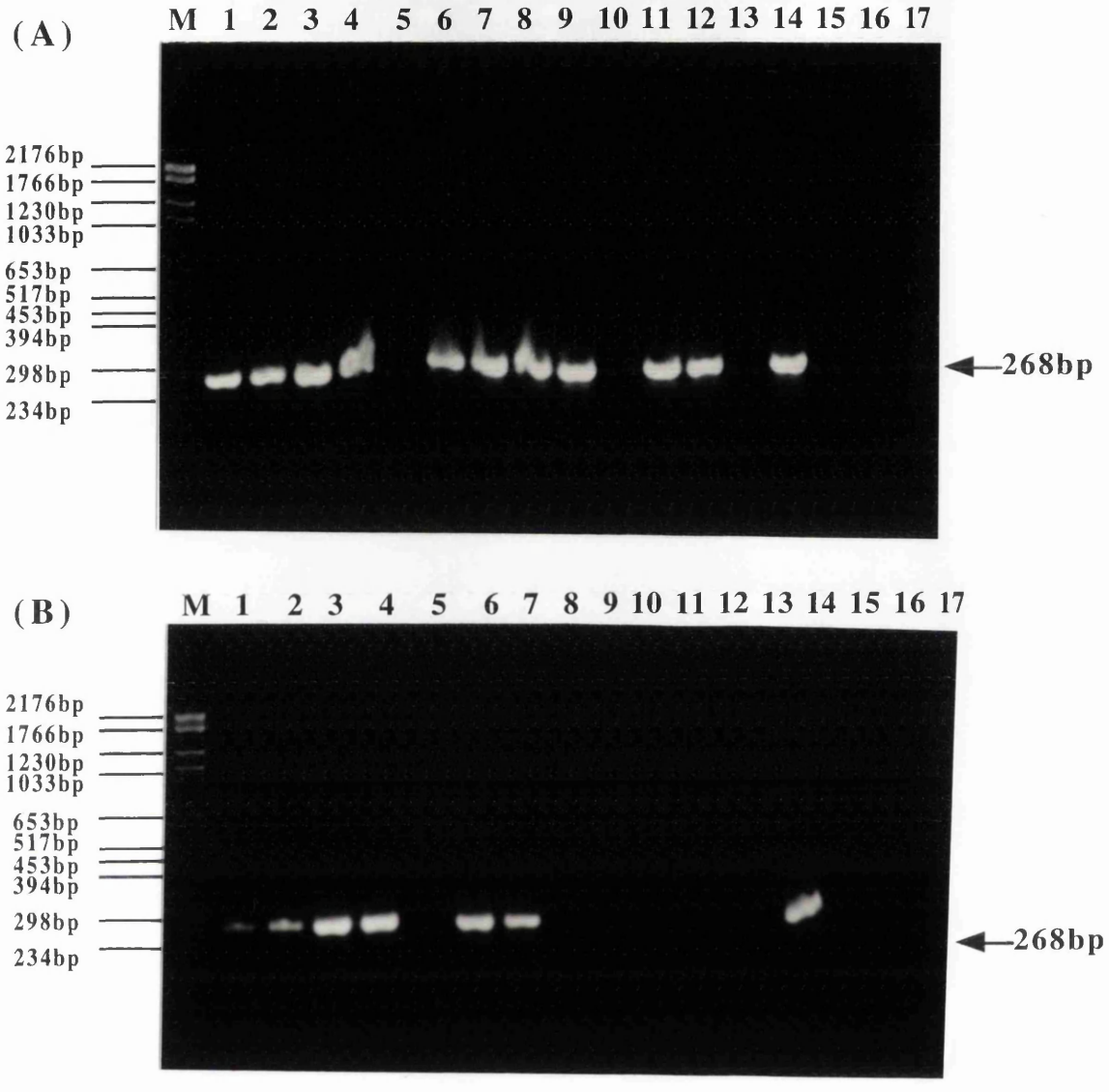


Figure 5.4 Gel electrophoresis of PCR amplified cDNA from the 5'NCR of serum subjected in duplicate to one freeze-thaw cycle **(A)** and 10 freeze-thaw cycles **(B).** HCV RNA was extracted from serially diluted serum, reverse transcribed and amplified by PCR. 10µl of each product was run on a 2% agarose gel. Lanes: (1, 2) serum diluted 1:100; (3,4) serum diluted 1:250; (5, 10) extraction carry-over control; (6,7) serum diluted 1:500; (8, 9) serum diluted 1:750; (11, 12) serum diluted 1:1000; (13) negative serum control; (14) positive serum control (15) negative control for reverse transcription reaction; (16, 17) negative control for first and second round PCR reactions. M, molecular weight marker (Boehringer Mannheim DNA molecular weight marker VI). Arrow denotes size of product. Mis-loading of product occurred in lanes 1 and 2 **(B).**

5.4 Discussion

The work presented in this chapter was undertaken to assess the effect of serum storage conditions on the detection of HCV RNA by RT-PCR. The storage conditions used in this study are common in diagnostic laboratories. It is therefore important to assess how such conditions affect the stability of HCV RNA. Unfortunately, there were several flaws in the study design and since it took more than a year, there was insufficient time to repeat it. Only one serum obtained from a single chronic HCV infected patient was used. Originally, blood from two chronic HCV infected patients was received but HCV RNA could not be consistently amplified from serum obtained from the second patient. From the work presented in this chapter, the actual loss in detectable levels of HCV RNA due to serum storage conditions cannot be fully assessed. No steps were taken to avoid possible bacterial contamination of the stored serum, which could influence the results obtained by RT-PCR. An inhibitor of bacterial growth such as sodium azide should have been added to the serum. The end-point dilution titre of the serum should have been determined before the serum was stored at either 4°C or -20°C and at all sampling points. Due to an insufficient amount of serum stored at 4°C each dilution could not be performed in duplicate. In hindsight, titration of the cDNA could have been done to make up for the insufficient amount of serum.

The results of the study presented in this chapter, are in agreement with those of Busch *et al.* (1992), Cuypers *et al.* (1992), Wang *et al.* (1992) and Fong *et al.* (1993), with minimal loss in the level of HCV RNA detected in serum stored at 4°C, at -20°C and subjected to repeated freeze-thawing. The longest period of time sera were stored at 4°C in the aforementioned studies was 14 days when the end-point dilutions were $10^{-3.3}$ to 10^{-5} (Cuypers *et al.*, 1992). In the study presented here, serum was stored at 4°C for 14 weeks before being sampled and the end-point dilution detectable by RT-PCR was 10^{-3}. This indicates HCV RNA is relatively stable when serum is stored at 4°C. However, over 57 weeks the end-point dilution detectable declined to 10^{o}

compared to 10^{-3} for serum stored at -20°C over the same time period. Therefore, for long term storage, sera should be kept at -20°C. Minimal or no reduction in the level of HCV RNA detectable following serum storage at 4°C for 48 hours, 5 days and 7 days was observed in studies by Wang *et al.* (1992), Fong *et al.* (1993) and Busch *et al.* (1992), respectively. In the study by Wang *et al.* all sera had been stored at -70°C prior to use, therefore the effect of storage conditions was not compared to freshly extracted RNA samples and in that by Busch *et al.* it is unclear if freshly extracted RNA from serum was used.

Repeated freeze-thawing of serum does not appear to affect the stability of HCV RNA. The detectable HCV RNA titre after 10 freeze-thaw cycles was $10^{-2.7}$ approximately (1:500 dilution) compared to 10^{-3} (1:1000 dilution) after one-freeze-thaw cycle, a two-fold reduction in titre. The indeterminate result observed at 10^{-4} after one freeze-thaw cycle could have been followed up by testing serial dilutions between 1:1000 and 1:10000 to obtain a more accurate end-point. The results obtained are in agreement with those of Fong *et al.* (1993) who observed end-point dilutions of 10^{-1} to 10^{-3} in five sets of patient sera subjected to five freeze-thaw cycles. An end-point dilution of 10^{-2} was observed after five freeze-thaw cycles in the study presented here. Despite methodological problems in this study, it is clear HCV RNA appears to be stable despite repeated freeze-thawing.

Serum stored continuously at -20°C had a HCV RNA titre of between 10^{-3} to 10^{-4}. Titres of 10^{-1} to 10^{-4} were detected by Fong et al. (1993).

The effect on the stability of HCV RNA when serum is stored at room temperature was not addressed. Busch *et al.* (1992) found a reduction in detectable HCV RNA titre of 2 log units in serum stored at room temperature for 7 days. A 3 to 4 log unit reduction was found in serum and whole EDTA-blood stored at room temperature after 8 to 14 days by Cuypers *et al.* (1992). In contrast to these two studies, Wang *et al.* (1992) and Fong *et al.* (1993) observed no reduction in HCV RNA titre after storage at room temperature for 48 hours and 5 days, respectively. The results of a study by Davis *et al.* (1994) using the branched-DNA (b-DNA) assay,

which determines quantitatively levels of HCV RNA present in serum, to assess the optimal conditions for storage agree with those of Busch *et al.* and Cuypers *et al.* Their results showed a 4.1% and 6.9% loss of HCV RNA from serum stored for 2 hours at room temperature and 4°C, respectively. After storage for 12 hours at room temperature this loss increased to 16.6% but no additional loss was observed at 4°C.

The results of this study and those from other research groups show that the sensitivity of RT-PCR is not significantly affected by serum storage conditions. Repeated freeze-thawing of serum or storage at 4°C does not result in a substantial reduction in HCV RNA levels detectable by RT-PCR compared to continuous storage at -20°C, indicating that HCV RNA present in serum is relatively resistant to degradation. But it is important to resolve the differences observed in studies concerning storage of serum at room temperature. Any reduction in detection of HCV RNA by RT-PCR must be considered when interpreting results from studies using stored serum where the investigators failed to state whether all samples were handled and stored under uniform conditions. Qualitatively, samples with a low HCV RNA titre may be more sensitive to serum storage conditions. Therefore, serum samples collected for confirmation of HCV infection or to be stored for future research purposes should be separated within 2-3 hours of blood collection and stored at -20°C.

Chapter 6 Results: Section 3

Construction of a control HCV RNA transcript for quantitative analysis of HCV infection.

The technique of RT-PCR is the method of choice for the amplification of HCV RNA, since it requires only small amounts of nucleic acid. However, PCR is qualitative not quantitative, indicating the presence or absence of target sequence only. Because the amount of PCR product increases exponentially with each cycle of amplification until it reaches a plateau, a difference in any of the variables that effect the efficiency of amplification no matter how small, can dramatically alter product yield. The need to measure the level of HCV RNA present is important for studies on the relationship between HCV RNA levels present in serum, plasma or liver tissue and the natural course of acute and chronic HCV infection and for evaluation of response of HCV infected individuals to anti-HCV therapies.

At present quantitation of HCV RNA is done either by limiting dilution analysis or competitive PCR. Limiting dilution analysis determines the end point titre of HCV RNA by performing RT-PCR on serial dilutions of extracted HCV RNA. To exclude variation in sensitivity between RT-PCR assays, calibration of each assay with a purified and quantitated control sample is performed (Kobayashi *et al.*, 1993). With competitive PCR, the amount of HCV RNA in serum or plasma is quantitated in a RT-PCR assay by co-amplification of the target HCV RNA with known amounts of synthetic control HCV RNA (Kaneko *et al.*, 1992; Hagiwara *et al.*, 1993; Kato *et al.*, 1993b; Naito *et al.*, 1994). The synthetic control HCV RNA differs in some way from the target sequence; it may differ in length, contain a unique restriction enzyme site or unique sequence to distinguish it from the target HCV RNA. The same primers are used to amplify the target and synthetic control HCV RNA allowing both target and synthetic control HCV RNA to be amplified with equal efficiency, therefore the ratio of

the amplified products reflects the initial amount of target HCV RNA versus that of the known added amount of synthetic control HCV RNA.

The work described in this chapter was undertaken to provide control HCV RNA to allow quantitative analysis to be performed. This control HCV RNA would allow the sensitivity and efficiency of the RNA extraction, RT and PCR methods employed in our laboratory to be calculated and studies on the natural course of HCV infection in renal dialysis patients to be undertaken. To synthesise the control HCV RNA, PCR products amplified from the 5' non coding and core region of HCV were used as template to introduce a unique 28 nucleotide scramble sequence between nucleotide position -194 to -171, numbering according to Choo *et al.* (1991), using PCR. This region of the virus was chosen because it is highly conserved among HCV isolates. The primers Econ11 and Econ12 were designed with a unique sequence of bases towards the 5' end of each primer, see Figure 6.1. Within the unique sequence an EcoR1 restriction enzyme site was incorporated. The first 10 bases of the unique sequence of primer Econ11 is complementary to the first 10 bases of Econ12 to allow products amplified with Econ11 and Econ12 to anneal together during amplification, resulting in amplification of a product incorporating the unique sequence of bases, Figure 6.2 outlines how the control HCV RNA would be synthesised. The resultant PCR product would then be purified to remove PCR reaction components, primers and other artefacts. A subsequent round of PCR would then be performed using primers containing bacteriophage T7 promoter sequences, to incorporate the T7 promoter sequences into the PCR product. Internal control HCV RNA would then be transcribed.

HCV sequence

5'- GAGAGCCATAGTGGTCTGCGCAACCGGTGAGTACACCGGAATTGCC -3'

CTCTCGGTATCACCAGCAATTCCGGTCTCCTTAAGCG Econ 12, antisense primer

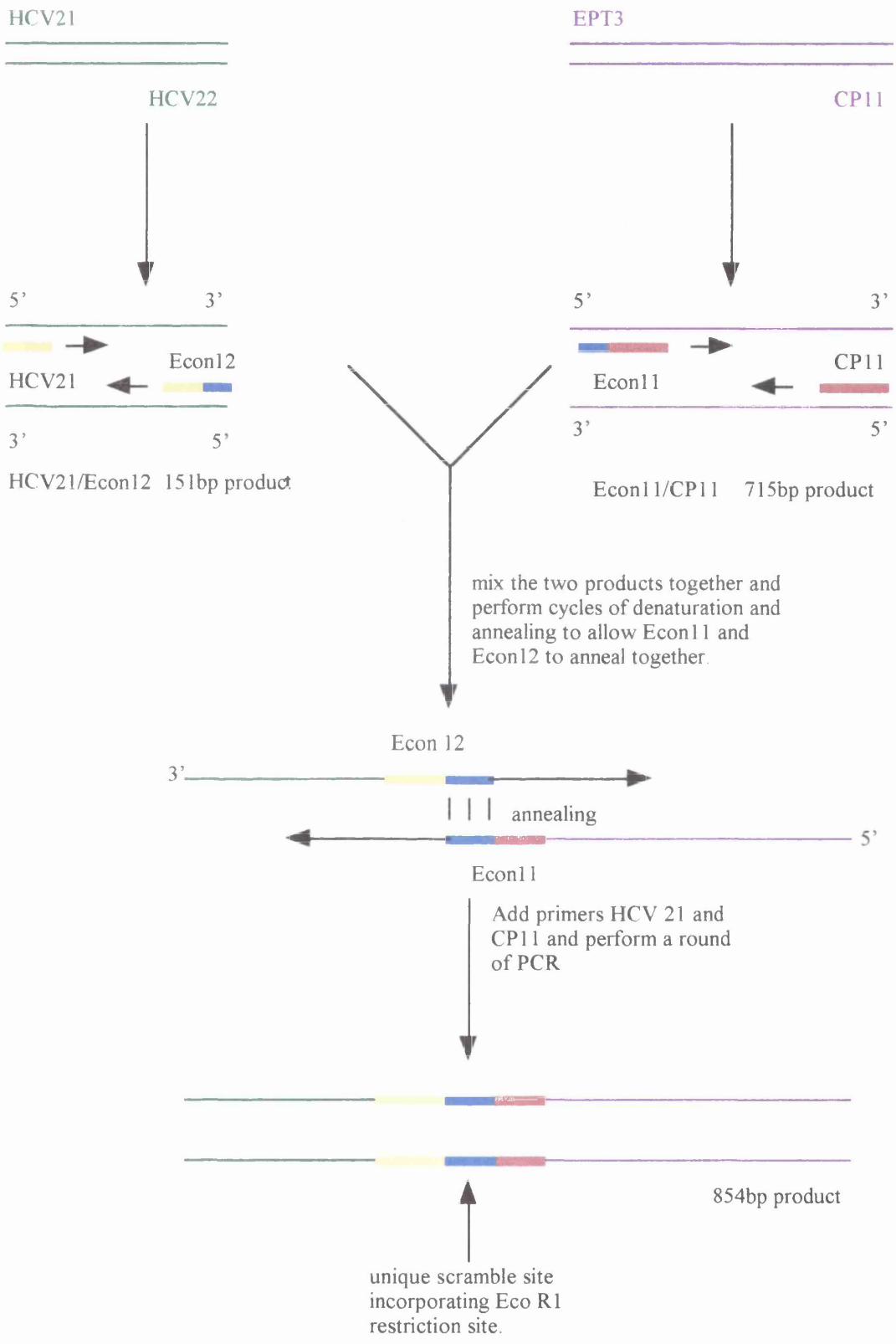
Econ 11, sense primer AGGAATTCGCCAGGCGTCCAGGAGGACCGGGT

3'- ACAGACGCCTTGGCCACTCATGCGCCCTTAACGGTCCTGCTGGCCCA -5'

HCV sequence

Figure 6.1 Shows the primers Econ 11 and Econ 12. The bases in red are those complementary to HCV sequence, those in blue are the unique scramble bases and those underlined are the 10 bases of each primer complentary to the other.

Figure 6.2 Schematic diagram outlining the method used to construct control HCV RNA transcripts.


HCV21
HCV22
5'
3'
Econ12
HCV21
3'
5'
HCV21/Econ12 151bp product
EPT3
CP11
5'
3'
CP11
Econ11
3'
5'
Econ11/CP11 715bp product
mix the two products together and perform cycles of denaturation and annealing to allow Econ11 and Econ12 to anneal together.
Econ 12
3'
annealing
5'
Econ11
Add primers HCV 21 and CP11 and perform a round of PCR
854bp product
unique scramble site incorporating Eco R1 restriction site.

6.1 Oligonucleotide primers

Oligonucleotide primers were derived from the 5' non coding and core region of the HCV isolate HCV-1 (Choo *et al* ., 1991). The following primers were used: HCV21 (sense; 5'-CGACACTCCACCATGAATCAC-3'; -322 to -302 nts.); HCV22 (anti-sense; 5'-GAGGTTTAGGATTCTGCTCATG-3'; -1 to 22 nts.); EPT1 (sense; 5'-GGCGTTAGTATGAGTGTCGT-3'; -255 to -236 nts.); SF3AS (antisense; 5'-AGGAAGATAGAGAAAGAGCAAC-3'; 512 to 533 nts.); CP11 (antisense; 5'-GAAGATAGAGAAAGAGCAACCA-3'; 510 to 531 nts.); Econ11 (sense; 5'-**AGGAATTCGCCAGGCGT**CCAGGACGACCGGGT-3'; -180 to -151nts.) and Econ12 (antisense; 5'-**GCGAATTCCTCTGGCCTTAAC**GACCACTATGGCTC-TC-3'; -209 to -173nts.), numbering according to Choo *et al* . (1991). The sequences in bold in Econ11 and Econ12 are the unique sequence of bases and those underlined are the 10 complementary bases. Primers were synthesised and purified as described Chapter 3 section 3.1. The concentration of each primer was determined by its optical density at 260nm. Primers were used at a concentration of 40 pmol per PCR reaction.

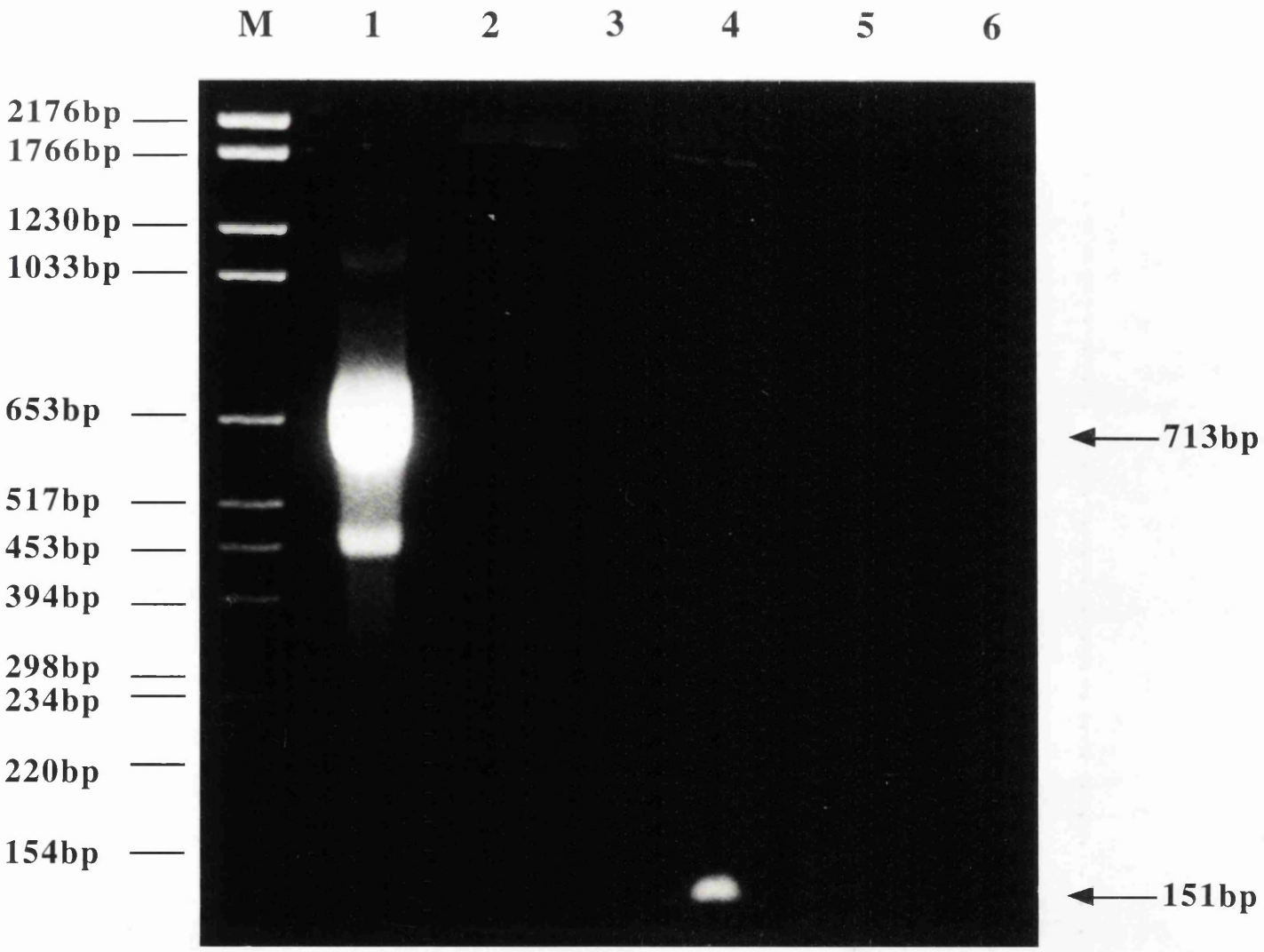
6.2 PCR amplification

To synthesise control HCV RNA, first round PCR products previously amplified with primer pairs HCV21 and HCV22, product one, and EPT1 and SF3AS, product two, were used to introduce the unique 28 nucleotide scramble sequence. Two PCR reactions were performed, product one was used as template for primer pair HCV21 and Econ12 and product two for primer pair Econ11 and CP11. Amplification was performed with an initial denaturation step of 4 minutes at 94°C followed by 2 cycles of denaturation for 30 seconds at 94°C, annealing of primers for 40 seconds at 45°C and extension for 2.5 minutes at 70°C, followed by 28 cycles of amplification with annealing of primers occurring at 55°C for 40 seconds and a final extension step of 7

minutes at 70°C. Two products the correct size were generated (Figure 6.3). The product generated by primer pair Econ11 and CP11 (715 base pairs) produced a strong band of DNA when 10µl of PCR product was run on EtBr stained 2% low melting point agarose gel and visualised under longwave UV light. This product band was excised from the gel. However, the smaller product generated by HCV21 and Econ12 (151 base pairs) was very faint, suggesting that this primer combination did not work well under the PCR conditions used. Since equal amounts of both products was needed, the PCR using primers HCV21 and Econ12 was repeated, only this time 5 cycles of amplification were performed with an annealing temperature of 45°C instead of 2 cycles. This made no difference to the amount of product generated. Problems with plasmid contamination in the laboratory had limited the choice of primers that could be used. Primer HCV21 was the outer sense primer of choice in the laboratory to amplify the 5' non coding region of HCV. In synthesising the control HCV RNA, it was important that it too could be amplified with HCV21, therefore finding an alternative sense primer to use with Econ12 was not feasible. Therefore, to obtain equal amounts of the two PCR products, the technique of 'hot start' PCR was used, with several PCR reactions using HCV21 and Econ12 being performed. The resulting products were run on a 2% low melting point agarose gel, the product band excised and the DNA extracted and re-amplified. To check that equal amounts of each PCR reaction product was being added, 10µl of each PCR product was run out on a 2% low melting point agarose gel and the intensity of each product band was visualised under longwave UV light . The product bands from both PCR reactions were excised from the gel, added to the one eppendorf tube, melted and vortexed to mix to be used as template in a further round of amplification involving the primers HCV21 and CP11.

To amplify the agarose gel template, the appropriate amount of PCR buffer, dNTPs, distilled water and Taq polymerase were added to a 0.5ml eppendorf, to which a wax gem was added and melted by heating the tube for 10 minutes at 70°C. 2µl of melted agarose products, were then added. Denaturation for 4 minutes at 94°C was performed, followed by 3 cycles of denaturation at 94°C for 30 seconds,

Figure 6.3 Agarose gel electrophoresis of the PCR products Econ11/CP11 and HCV21/Econ12. Lane 1 Econ11/CP11 PCR product; Lane 2 and 5, negative serum control; lane 3 and 6 negative control for second round PCR reaction; Lane 4 HCV21/Econ12 PCR product. M, molecular weight marker (Boehringer Mannheim DNA molecular weight marker VI).


M
1
2
3
4
5
6
2176bp
1766bp
1230bp
1033bp
653bp
517bp
453bp
394bp
298bp
234bp
220bp
154bp
713bp
151bp

annealing at 45°C for 40 seconds and extension at 70°C for 2.5 minutes, to allow the two products to anneal together at the region where they were complementary to one another, after which the primers HCV21 and CP11 were added and 25 cycles of amplification were performed using the PCR conditions described above, except annealing occurred at 55°C and there was a final extension step of 7 minutes at 70°C. The resulting PCR product is shown in Figure 6.4. A product the correct size, 854 bases, was produced but smaller products were also generated. The correct size product band was excised and the product reamplified. An EcoR1 restriction digest was then performed on the amplified product to check for the 28 nucleotide scramble sequence (Figure 6.5). A plasmid with known EcoR1 sites was included as a control to confirm complete digestion had taken place. There was no digestion of the amplified product suggesting that the 28 nucleotide scramble sequence had not been incorporated. However, the size of products generated by PCR was all correct suggesting incorporation had taken place. To resolve this issue, the product generated by the primer pair HCV21/CP11 was nucleotide sequenced.

6.3 Single stranded sequencing of PCR product with unique scramble sequence of bases

Single stranded sequencing of the product amplified by primer pair HCV21/CP11 was performed using Dynabead method, as described in Chapter 3 sections 3.9.2 and 3.10. The biotin labelled primers used to generate single stranded DNA were HCV23 (sense; 5'-TCACTCCCCTGTGAGGAACT-3'; -305 to -286 nts.) and NCR4 (antisense; 5'-GCACCCTATCAGGCAGT-3'; -54 to -37 nts.), (numbering according to Choo *et al.* (1991). Figure 6.6 shows the resulting autoradiograph. Up to and after the point where incorporation of the 28bp unique scramble sequence of bases should occur the sequence obtained is correct for the 5'NCR of HCV (Figure 6.7). However, at each nucleotide position (between -194 and -171) where the scramble sequence (5'-

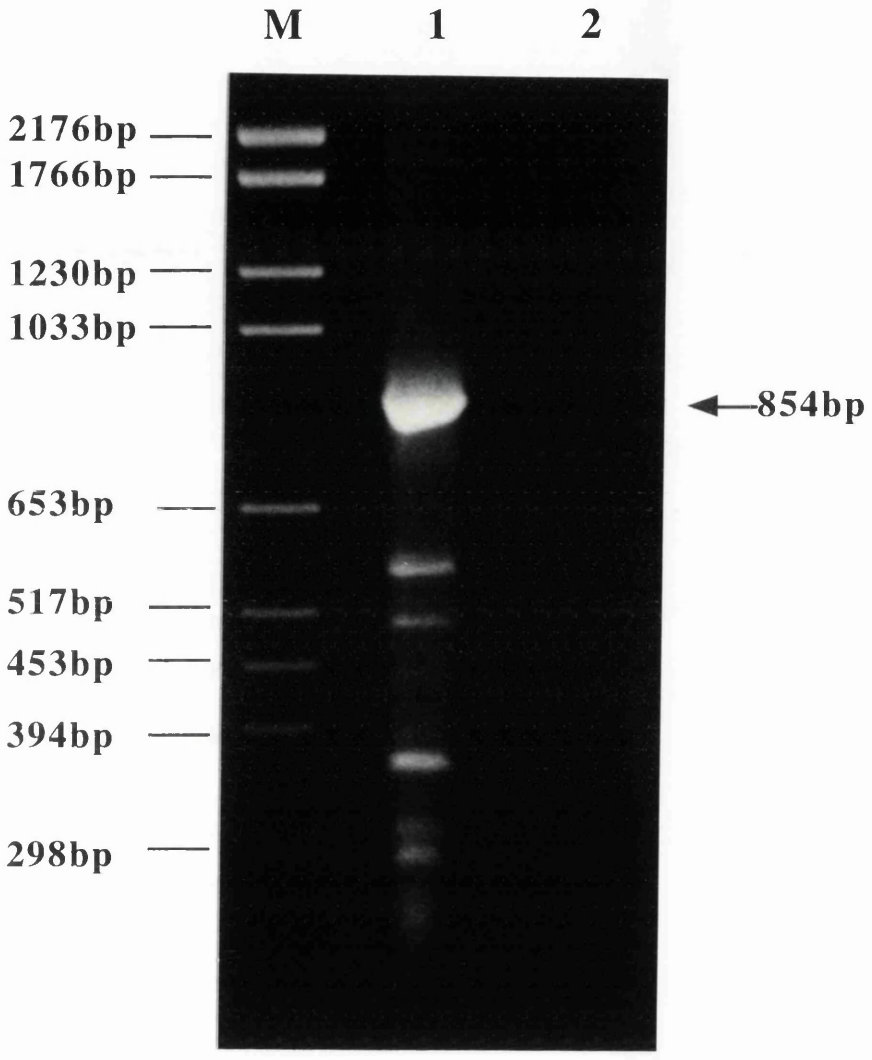


Figure 6.4 Agarose gel electrophoresis of the PCR product generated by amplifying the Econ11/CP11 and HCV21/Econ12 generated PCR fragments with the primers Econ11 and Econ12. Lane 1, the resultant Econ11/Econ12 amplified fragment; lane 2, negative control for the PCR reaction. M, molecular weight marker (Boehringer Mannheim DNA molecular weight marker VI).

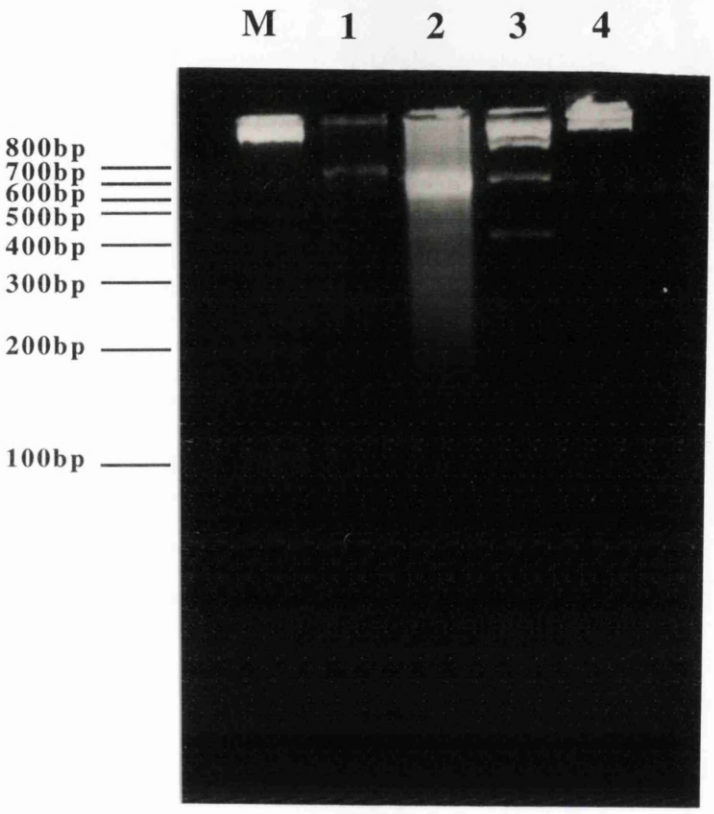
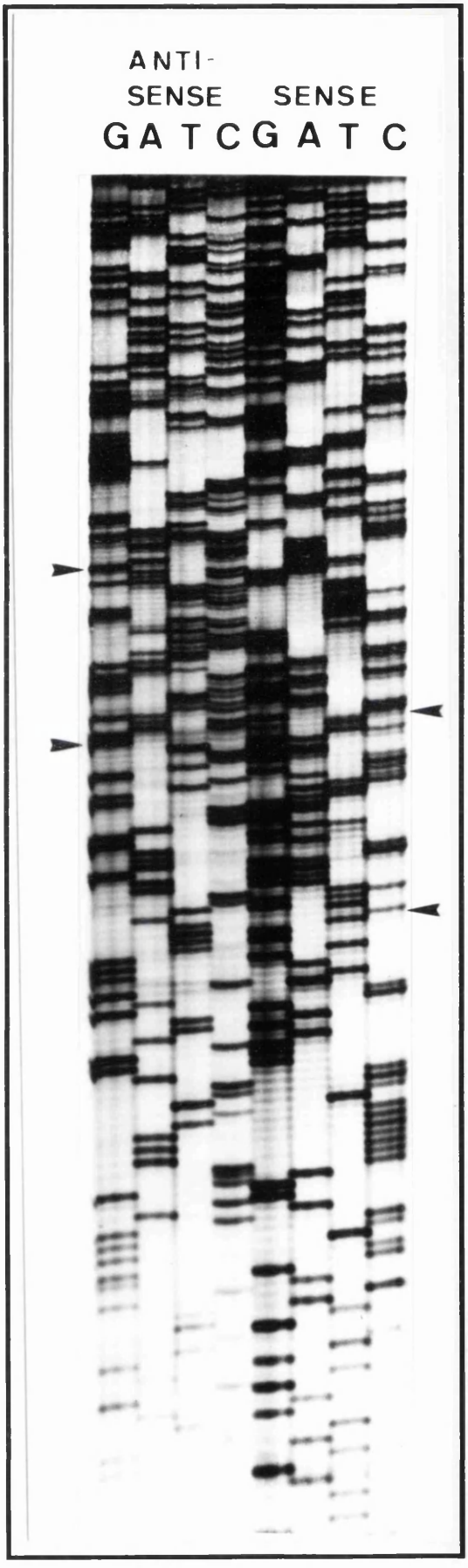


Figure 6.5 Agarose gel electrophoresis of Econ11/Econ12 PCR product EcoR1 restriction digest. Lane 1, digested Econ11/Econ12 fragment; lane2, undigested Econ11/Econ12 fragment; lane3, digested control plasmid; lane4, undigested control plasmid. M, molecular weight marker (Gibco, BRL 100bp DNA ladder).

Figure 6.6 Direct DNA sequencing of the Econ11/Econ12 amplified PCR product. Both the sense and antisense sequencing direction are shown. The arrows indicate the incorporation site of the 28bp unique scramble sequence.


ANTI-
SENSE
SENSE
G A T C G A T C

GTTAAGGCCAGAGGAATTCGCCAGGCGT-3') should have been incorporated, two bases are present (Figure 6.7). The sequence of the mixed base positions appears to indicate the presence of both control and native HCV DNA in the final PCR product. The sequence of native HCV between the primers Econ12 and Econ11 is 5'-TGCGCAACCGGTGAGTACACCGGAATTG-3', suggesting that at some stage during the synthesis of the PCR product "contamination" with native HCV DNA has occurred. If time had permitted, one could resolve this by cloning the PCR product and selecting clones with the desired insert.

1 46

TAGCCATGGCGTTAGTATGAGTGTCGTACAGCCTCCAGGACCCCCC

47 90

CTCCCGGGAGAGCCATAGTGGTCNNNNNNNNNNNNNNNNNNNN

91 136

NNNNNNNCCAGGACGACCGGGTCCTTTCTTGGATAAAACCCGCTCA

137 177

ATGCCTGGAGATTTGGGCGTGCCCCCGCAAGACTGCTRGCC

Figure 6.7 Sequence of PCR product (primers HCV21/CP11). The sequence in blue and red type face denotes the HCV sequence of primers Econ12 (antisense) and Econ11 (sense), respectively.

6.4 Discussion

The work presented in this chapter was undertaken to provide control HCV-RNA transcripts for quantitative competitive RNA PCR analysis. Several authors have used either internal control cDNA or HCV RNA to quantitiate HCV RNA levels (Kaneko *et al.*, 1992; Hagiwara *et al.*, 1993; Kato *et al.*, 1993). In the study by Kaneko *et al.* (1992) competitive PCR was used to analyse serial changes in serum HCV RNA levels from patients with chronic hepatitis C who had received α-interferon treatment. A deletion mutant HCV cDNA plasmid was used as the internal control for competitive PCR. The control plasmid was constructed by cloning a fragment of HCV cDNA into a pGEM vector and performing restriction digests on the cloned HCV cDNA. The resultant subclone contained the same primer sites as target cDNA but it was 106 base pairs shorter. Target HCV cDNA, reverse transcribed from HCV RNA extracted from patient sera, was co-amplified with known amounts of the deletion mutant HCV-cDNA. The PCR products were visualised by UV fluorescence after electrophoresis. Two product bands were seen, the smaller corresponding to the internal control HCV. At the point where the amount of target HCV and internal control HCV concentrations are equivalent (i.e. 1:1 ratio), the PCR product bands were of equal intensity.

Performing competitive PCR using HCV cDNA as an internal control does not take into account the efficiency of the RNA extraction method or reverse transcription reaction. To overcome this, control HCV RNA transcripts are produced which can be added in known amounts to patient serum or plasma and hence be co- extracted, reverse transcribed and amplified. Studies by Hagiwara *et al.* (1993) and Kato *et al.* (1993b) used control HCV RNA transcripts to quantitate HCV RNA levels. Both groups used site-directed mutagenesis by RT-PCR and *in vitro* transcription to produce control HCV RNA transcripts. The resultant product was cloned into a vector with a promoter to generate the RNA transcripts.

The method described in this chapter to produce control HCV RNA transcripts, with a unique 28 base sequence and EcoR1 restriction site (Figure. 6.2), was designed

to avoid cloning the PCR product into a vector containing a T7 promoter to produce RNA transcripts, thus avoiding the risk of PCR contamination in the laboratory. Unfortunately, "contamination" appears to have occurred at some stage during the synthesis of the PCR product which was to be used to produce the control HCV RNA transcripts, which resulted in a heterogeneous product. The sequencing result suggests "contamination" with native HCV cDNA rather than mis-annealing of the two products. If mis-annealing had occurred, the two strands would not overlap by 10 bases every time resulting in a product of fixed length but one would expect mixed base positions outwith nucleotide positions -194 to -171 and products of varying length. This is not what is observed. The contamination with native HCV cDNA probably occurred during amplification of template DNA with the primer combinations HCV21/Econ12 and Econ11/CP11. Although during amplification the shorter cDNA copies, with both ends defined by the primer sequences, are amplified, the original full-length template DNA will remain, resulting in copies of native HCV cDNA being present. In addition, there may have been mispriming upstream. Since the primer combination Econ11/CP11 produced a more diffuse PCR product band than the primer combination HCV21/Econ12, it is more likely that the Econ11/CP11 PCR reaction was the source of contamination. The size of the "short product" DNA and original full-length template DNA produced by the primer combination Econ11/CP11 are 715 and 787 base pairs, respectively. Because the original full-length template DNA differs in size from the "short product" DNA by only 72 base pairs, it is possible it was excised along with the "short product" from the low melting point agarose gel. If both native and control HCV cDNA are present, why did the restriction digest with EcoR1 not cut the control HCV cDNA present ? Perhaps it did but the amount of resultant product was present in insufficient quantities to be detected by gel electrophoresis. To prove the presence of native HCV DNA in the PCR product, the product would need to be cloned into bacteria and the resultant clones sequenced. This could also allow selection of clones with the desired insert to act as template for control RNA production.

In a paper by Gretch *et al.* (1994) reporting the relationship between HCV RNA levels and HCV pathogenesis, internal control HCV RNA was synthesised using a similar technique to the one described in this chapter. A 60 base pair *Salmonella typhimurium* DNA insert was introduced into a PCR product amplified from the HCV 5' NCR. Instead of 10 base pairs of overlapping sequence between the two internal primers used to introduce the DNA insert into the PCR product, Gretch *et al.* (1994) synthesised two internal primers JHC160 (sense) and JHC161 (antisense), with 60 base pairs of overlapping *S.typhimurium* sequence. However, after performing two separate PCR reactions with the primers JHC160 and JHC51 (HCV specific antisense primer) and JHC161 and JHC93 (HCV specific sense primer), the resultant double stranded DNA products were digested with lambda exonuclease to produce single stranded DNA. The single stranded DNA products were then purified before being reamplified with the primers JHC93 and JHC51. A further round of PCR was then performed with primers containing bacteriophage T7 promoter sequences to allow *in vitro* transcription. If the method described by Gretch *et al.* (1994) had been followed, the problem of contamination with native HCV DNA may not have occurred. Also, after each round of PCR, the amplified DNA should have been purified, thereby removing any PCR reaction primers and non-specific products which may interfere with the subsequent round of amplification. Another factor to be considered if repeating the work presented here, would be to quantitiate the HCV21/Econ12 and Econ11/CP11 PCR product DNA by measuring the optical density of the DNA at 260nm, quantitating the products using this method would perhaps be more accurate. Thus, the actual amount of each product present before they are co-amplified together is known.

6.4.1 Future Studies

The ability to measure the level of HCV RNA present in serum or plasma from infected persons is important in understanding the pathogenesis of HCV infection, deciding when to administer anti-viral therapy and assessing treatment efficacy. The findings of Gretch *et al.* (1994) and Naito *et al.* (1994) indicate that high levels of HCV RNA may correlate with the progression of liver disease. It would therefore, be of interest to measure the HCV RNA levels present in the three renal dialysis patients, studied in Chapter 4, to assess the relationship between the level of HCV RNA and the emergence of new genetic variants.

Chapter 7 General Discussion

Hepatocyte damage in HCV infected individuals may result from direct cytopathic action of the virus or may be the result of the host's immune response. RNA viruses are noted for their variability in genome sequence, as has been shown *in vivo* for HCV (Martell *et al.*, 1992). Aspects of genetic variation of HCV which could influence pathogenesis are (i) antigenic variation and (ii) variants that differ in virulence. Antigenic variation can lead to escape from the host's humoral and cell-mediated immune responses. If damage is due to a direct cytopathic action of the virus, a new genetic variant would be expected to emerge at any time. Alternatively, if damage was as a result of the destruction of infected hepatocytes by the host's immune system, a new genetic variant would appear after an elevation in ALT. It is widely believed that variation in the HVR1 domain in E2 is driven by immune selection pressure. The findings of the study described in Chapter 4 showed sequence variation of the E2 HVR1 domain in the predominant virus population to occur during the natural course of hepatitis C infection in renal dialysis patients. Unfortunately, the results of this study did not show an association between the emergence of new sequence variants and either mechanism of liver damage, based on ALT levels as a measurement of liver damage. At present, liver biopsies are the only way to evaluate accurately the severity of liver disease but performing biopsies frequently to monitor liver dysfunction is not feasible and since there is no alternative biochemical marker, ALT levels are monitored.

It is possible that sequence variants play another role which is selectively advantageous for the virus. Genetic mutation of the virus may result in the generation of variants that differ in virulence. This may affect the severity and progression of disease. The HVR1 may be important for cell tropism, the formation of virus/antibody complexes or viral entry. It is also possible that the HVR1 domain itself has no role in viral persistence. Before understanding what role, if any, HVR1 variants play in

persistence of HCV within the host, several issues will have to be addressed. These are: (1) the biological function of HVR1; (2) whether antibodies to epitopes in the HVR1 are neutralising; (3) whether cytotoxic T-cells against HCV provide a protective response against the virus; (4) the structure of the envelope glycoproteins. Rosa *et al.* (1996) have recently provided evidence that although antibodies to the HVR1 domain are neutralising, the protection induced by vaccination with recombinant envelope proteins is not dependent on them.

The structure of the HCV envelope proteins may be similar to those of alpha- and flaviviruses. Like the two envelope glycoproteins (E1 and E2) of Semliki Forest virus (SFV), an alphavirus, the E1 and E2 proteins of HCV appear to form a complex (Grakoui *et al.*, 1993a; Ralston *et al.*, 1993). Indeed chimpanzees vaccinated with recombinant E1/E2 protein were shown to be protected upon challenge with the virus (Choo *et al.*, 1994). The E1 and E2 proteins of SFV form a heterohexameric complex that projects out from the surface. Although, structurally different from the influenza HA protein, a solid trimer, SFV also undergoes a pH-induced conformational change to allow the fusion protein to associate with the plasma membrane of the cell. The fusion protein (E1) is brought to the surface in the form of a homotrimer (Helenius, 1995). Alternatively, the structure adopted by the HCV envelope proteins may be similar to the major envelope protein of the flavivirus tick-borne encephalitis (TBE) virus. Unlike SFV, influenza and HIV membrane proteins, the major envelope protein of TBE is a flat, elongated homodimer extending laterally along and parallel to the membrane surface. TBE, like influenza and alphaviruses, enters the cell via an endosome and a low pH-conformational change occurs. Rey *et al.* (1995) has suggested that this conformational change allows the protein to protrude from the surface i.e. stand up and interact with the cell to allow fusion. It is possible that the structure adopted by the HCV envelope proteins is such that HVR1 domain is not exposed on the surface of the virion and therefore, not exposed to immune selection pressure. It has been suggested that the gp120 in HIV exists as a tetrameric protein with the V3 loop being in the middle of it unexposed to the host immune system and

under no structural constraints, therefore it can mutate at will (Howard Marsden, personal communication). One could speculate that the sequence variation displayed by HVR1 arises because this region is under no structural constraint to remain conserved.

A modified, non-pathogenic strain of an infectious virus, able to stimulate the host's immune system but whose replication and spread is impaired compared to the wild type virus, may be considered an ideal vaccine. The favoured type of vaccine against viruses in the past has been live attenuated virus. Vaccines of this nature are able to stimulate both arms of the immune system. The live attenuated virus will encounter the same cells of the host's immune system as the wild type virus and so stimulate them to respond in a similar manner to that seen in a natural infection. Antibodies and T-cells are produced to a large number of epitopes that give rise to a large population of memory cells and a high level of neutralising antibody. However, the disadvantage of attenuated virus vaccines is the possibility of reversion to a virulent form. For many viral vaccines, surface antigens of the virus are among the most important to elicit neutralising antibodies. In the case of HCV, the envelope glycoproteins E1 and E2 are likely to be the targets of the host's humoral immune response. Neutralising antibodies against the envelope proteins of HCV would prevent the virus from infecting hepatocytes. The findings of Choo *et al.* (1994) demonstrated that chimpanzees vaccinated with recombinant E1/E2 protein were protected from infection. However, by mutating the genes encoding the E1 and E2 proteins the virus could evade the host's humoral immune response. The E2 HVR1 domain in particular is highly variable and appears to contain neutralising epitope(s) (Zibert *et al.*, 1995). If the HVR1 domain were to contain the only neutralising epitope(s) of the virus, neutralising antibodies induced by vaccination may not afford protection from infection because of the sequence variation displayed by this region between different HCV variants. In terms of developing an effective vaccine against HCV it will be important to identify a region of the genome which contains conserved epitopes, so that the response elicited will protect against infection from all strains. Recent evidence suggests that at least two neutralising epitopes are present on the E2 protein, one of

which may be conserved in some infected individuals (Rosa *et al.*, 1996). Yet, despite this finding it may prove difficult to develop a multivalent vaccine against HCV which elicits a neutralising antibody titre which is high enough to protect against infection, since HCV infection elicits low or no neutralising antibody titres and chimpanzees can be reinfected with same homologous strain (Farci *et al.*, 1992; Rosa *et al.*, 1996). Thus, in the case of HCV a vaccine capable of stimulating a good cell-mediated immune response may be of more importance, especially it's ability to stimulate a CTL response in the vaccinee.

A new approach to immunisation is the use of DNA vaccines. Such vaccines are plasmids that contain the gene(s) for the antigenic portion of the target virus. By injecting naked DNA into a person both humoral and cell-mediated responses are induced but naked DNA appears, in animals models at least, to stimulate cell-mediated immunity more efficiently than a live attenuated vaccine (McDonnell & Askari, 1996). The naked DNA enters the host cell, where it is expressed and the corresponding protein is synthesised inside the cell, allowing the viral protein to enter the MHC class I pathway and stimulate a CTL response. One advantage a DNA vaccine has over a standard live attenuated vaccine is that genes from several different strains could be included on the same plasmid and a multivalent response could be elicited. This could prove useful if no one conserved epitope(s) exists between all HCV variants.

The second part of the work presented in this thesis was aimed at determining whether serum storage conditions could affect the detection of HCV RNA by RT-PCR. As a result of difficulties experienced in amplifying the HVR1 region from stored patient sera, it was important to consider what effect, if any, storage conditions had on the stability of HCV RNA. The study described in Chapter 4 of this thesis like many other published studies was retrospective. The results of the work described in Chapter 5 showed that HCV RNA was relatively stable despite prolonged storage at 4°C or successive freeze-thaw cycles. The results of this work are also important from a diagnostic viewpoint. When a person becomes infected with HCV it may be several weeks or months before seroconversion occurs. If so, then it is important to be able to

go back and test sera which may have been taken from the patient before seroconversion for the presence of HCV RNA so that the time of infection may be known.

The need to measure the level of HCV RNA is important for studies on the relationship between HCV RNA levels in serum or liver tissue and the evaluation of response of HCV infected individuals to anti-HCV therapies. It is also important in studying the natural course of acute and chronic infection. For these reasons I attempted to construct control HCV RNA transcripts for quantitative analysis of the HCV RNA levels in patient serum, notably that of the three renal dialysis patients studied in Chapter 4, to assess the relationship between the level of HCV RNA in patient serum and the emergence of HVR1 genetic variants. Unfortunately, a problem with contamination prevented the construction of control PCR products. Cloning of the resultant product would allow the control cDNA to be selected, from which RNA transcripts could be synthesised.

REFERENCES

Abe, K., Inchauspe, G. & Fujisawa, K. (1992a). Genomic characterization and mutation rate of hepatitis C virus isolated from a patient who contracted hepatitis during an epidemic of non-A, non-B hepatitis in Japan. *Journal of General Virology* **73**, 2725-2729.

Abe, K., Inchauspe, G., Shikata, T. & Prince, A. M. (1992b). Three different patterns of hepatitis C virus infection in chimpanzees. *Hepatology* **15**, 690-695.

Alexander, S. & Elder, J. H. (1984). Carbohydrate dramatically influences immune reactivity of antisera to viral glycoproteins. *Science* **226**, 1328-1330.

Allander, T., Medin, C., Jacobson, S. H., Grillner, L. & Persson, M. A. (1994). Hepatitis C transmission in a haemodialysis unit: Molecular evidence for spread of virus among patients not sharing equipment. *Journal of Medical Virology* **43**, 415-419.

Alter, H. J. (1991). Descartes before the horse: I clone, therefore I am: The hepatitis C virus in current perspective. *Annals of Internal Medicine* **115**, 644-649.

Alter, H. J., Purcell, R. H., Holland, P. V., Feinstone, S. M., Morrow, A. G. & Moritsugu, Y. (1975). Clinical and serological analysis of transfusion-associated hepatitis. *The Lancet* **1**, 838-841.

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

Alter, H. J., Purcell, R. H., Shih, J. W., Melpolder, J. C., Houghton, M., Choo, Q.-L. & Kuo, G. (1989). Detection of antibody to hepatitis C virus in prospectively followed transfusion recipients with acute and chronic non-A, non-B heaptitis. *The New England Journal of Medicine* **321**, 1494-1500.

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., Samliner, R. E., Meeks, E. L. & Beach, M. J. (1992). The natural history of community-acquired hepatitis C virus in the United States. *The New England Journal of Medicine* **327**, 1899-1905.

Bartenschlager, R., Ahlborn-Laake, L., Mous, J. & Jacobsen, H. (1993). Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *Journal of Virology* **67**, 3835-3844.

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

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

Bartenschlager, R., Lohmann, V., Wilkinson, T., & Koch, J. O. (1995b). Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and the NS4A and its importance for polyprotein maturation. *Journal of Virology* **69**, 7519-7528.

Battegay, M., Fikes, J., Bisceglie, A. M. D., Wentworth, P. A., Sette, A., Celis, E., Ching, W.-M., Grakoui, A., Rice, C. M., Kurokohchi, K., Berzofsy, J. A. & Akatsuka, T. (1995). Patients with Chronic Hepatitis C Have Circulating Cytotoxic T Cells Which Recognize Hepatitis C Virus-Encoded Peptides Binding to HLA-A2.1 Molecules. *Journal of Virology* **69**, 2462-2470.

Bertoletti, A., Sette, A., Chisari, F. V., Penna, A., Levrero, M., Carli, M. D., Fiaccadori, F. & Ferrari, C. (1994). Natural variants of cytotoxic epitopes are T-cell receptor antagonists for antiviral cytotoxic T cells. *Nature* **369**, 407-410.

Borrego, B., Novella, I. S., Giralt, E., Andreu, D. & Domingo, E. (1993). Distinct Repertorie of Antigenic Variants of Foot-and-Mouth Disease Virus in the Presence or Absence Immune Selection. *Journal of Virology* **67**, 6071-6079.

Bradley, D. W. (1990). Hepatitis non-A, non-B viruses become identified as hepatitis C and E viruses. *Progress in Medical Virology* **37**, 101-135.

Bradley, D. W., Cook, E. H., Maynard, J. E., McCaustland, K. A., Cook, E. H., Ebert, J. W., Dolana, G. H., Petzel, R. A., Kantor, R. J., Heilbrunn, A., Fields, H. A. & Murphy, B. L. (1979). Experimental infection of chimpanzees with antihaemophilic (factor VIII) materials. Recovery of virus-like particles associated with non-A, non-B hepatitis. *Journal of Medical Virology* **3**, 253-269.

Bradley, D. W., Maynard, J. E., Cook, E. H., Ebert, J. W., Gravelle, C. R., Tsiquaye, K. N., Kessler, H., Zuckerman, A. J., Miller, M. F., Ling, C.-M. & Overby, L. R. (1980). Non-A, non-B hepatitis in experimentally infected chimpanzees: Cross-challenge and electron microscopic studies. *Journal of Medical Virology* **6**, 185-201.

Bradley, D. W., Maynard, J. E., Popper, H., Cook, E. H., Ebert, J. W., McCaustland, K. A., Schable, C. A. & Fields, H. A. (1983). Posttransfusion non-A, non-B hepatitis: Physicochemical properties of two distinct agents. *The Journal of Infectious Diseases* **148**, 254-265.

Bradley, D. W., McCaustland, K. A., Cook, E. H., Schable, C. A., Ebert, J. W. & Maynard, J. E. (1985). Posttransfusion non-A, non-B hepatitis in chimpanzees. *Gastroenterology* **88**, 773-779.

Bradley, D., McCaustland, K., Krawczynski, K., Spelbring, J., Humphrey, C. & Cook, H. E. (1991). Hepatitis C virus: Buoyant density of the factor VIII-derived isolate in sucrose. *Journal of Medical Virology* **34**, 206-208.

Bresters, D., Mauser-Bunschoten, E. P., Reesink, H. W., Roosendaal, G., Poel, C., Chamuleau, R. A., Jansen, P. L., Weegink, C. J., Cuypers, H. T. M., Lelie, P. N. & Berg, H. M. (1993). Sexual ransmission of hepatitis C virus. *The Lancet* **342**, 210-211.

Brillanti, S., Garson, J., Foli, M., Whitby, K., Deaville, R., Masci, C., Miglioli, M., & Barbara, L. (1994). A pilot study of combination therapy with ribavirin plus interferon alfa for interferon alfa-resistant chronic hepatitis C. *Gastroenterology* **107**, 812-817.

Brillanti, S., Miglioli, M., & Barbara, L. (1995). Combination antiviral therapy with ribavirin and interferon alfa in interferon alfa relapsers and non-responders: Italian experience. *Journal of Hepatology* **23** (Suppl. 2), 13-16.

Broder, C. C., Earl, P. L., Long, D., Abedon, S. T., Moss, B. & Doms, R. W. (1994). Antigenic implications of human immunodeficiency virus type 1 envelope quaternary structure: Oligomer-specific and -sensitive monoclonal antibodies. *Proc. Natl. Acad. Sci USA* **91**, 11699-11703.

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

Bruno, S., Rossi, S., Petroni, M. L., Villa, E., Zain, M. & Podda, M. (1994). Normal aminotransferase concentrations in patients with antibodies to hepatitis C virus. *British Medical Journal* **308**, 697-698.

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

Burns, D. P. W. & Desrosiers, R. C. (1994). Envelope sequence variation and primate lentivirus persistence. *Current Topics in Microbiology and Immunology* **188**, 185-220.

Busch, M. P., C, W. J., Johnson, P., Tobler, L. & Evans, C. S. (1992). Impact of specimen handling and storage on detection of hepatitis C virus RNA. *Transfusion* **32**, 420-425.

Carpenter, S., Evans, L. H., Sevoian, M. & Chesebro, B. (1987). Role of the host immune response in selection of equine infectious anaemia virus variants. *Journal of Virology* **61**, 3783-3789.

Cha, T.-A., Beall, E., Irvine, B., Kolberg, J., Chien, D., Kuo, G. & Urdea, M. S. (1992). At least five related but distinct hepatitis C viral genotypes exist. *Proc. Natl. Acad. Sci. USA* **89**, 7144-7148.

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

Choo, Q.-L., Kuo, G., Ralston, R., Weiner, A., Chien, D., Nest, G. v., Han, J., Berger, K., Thudium, K., Kuo, C., Kansopon, J., McFarland, J., Tabrizi, A., Ching, K., Moss, B., Cummins, L. B., Houghton, M. & Muchmore, E. (1994). Vaccination of chimpanzees against infection by the hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **91**, 1294-1298.

Choo, Q.-L., Kuo, G., Weiner, A. J., OVerby, L. R., Bradley, D. W. & Houghton, M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **244**, 359-362.

Choo, Q.-L., Richman, K. H., Han, J. H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, A., Barr, P. J., Weiner, A. J., Bradley, D. W., Kuo, G. & Houghton, M. (1991). Genetic organization and diversity of the hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **88**, 2451-2455.

Clements, J. E., Gdovin, S. L., Montelaro, R. C. & Narayan, O. (1988). Antigenic variation in lentiviral diseases. *Annual Reviews of Immunology.* **6**, 139-159.

Collett, M. S., Anderson, D. K. & Retzel, E. (1988). Comparisons of the pestivirus bovine viral diarrhoea virus with members of the flavivirdae. *Journal of General Virology* **69**, 2637-2643.

Cuypers, H. T. M., Bresters, D., Winkel, I. N., Reesink, H. W., Weiner, A. J., Houghton, M., van der Poel, C. L. & Lelie, P. N. (1992). Storage conditions of blood samples and primer selection affect the yield of cDNA polymerase chain reaction products of hepatitis C virus. *Journal of Clinical Microbiology* **30**, 3220-3224.

Cuypers, H. T. M., Winkel, I. N., Poel, C. L. v. d., Reesink, H. W., Lelie, P. N., Houghton, M. & Weiner, A. (1991). Analysis of genomic variability of hepatitis C virus. *Journal of Hepatology* **13** (Suppl. 4), S15-S19.

Davis, G. L., Lau, J. Y.-N., Urdea, M. S., Neuwald, P. D., Wilber, J. C., Lindsay, K., Perrillo, R. P. & Albrecht, J. (1994). Quantitative detection of hepatitis C virus RNA with a solid-phase signal amplification method: definition of optimal conditions for specimen collection and clinical application in interferon-treated patients. *Hepatology* **19**, 1337-1341.

Deka, N., Sharma, M. D. & Mukerjee, R. (1994). Isolation of the novel agent from human stool samples that is associated with sporadic non-A, non-B hepatitis. *Journal of Virology* **68**, 7810-7815.

Di Bisceglie, A., Goodman, Z., Ishak, K., Hoofnagle, J., Melpolder, J. & Alter, H. (1991). Long-term clinical and histological follow-up of chronic posttransfusion hepatitis. *Hepatology* **14**, 969-974.

Di Bisceglie, A., Shindo, M., Fong, T., Fried, M., Swain, M., Bergasa, N., Axiotis, C., Waggoner, J., Park, Y. & Hoofnagle, J. (1992). A pilot study of ribavirin therapy for chronic hepatitis C. *Hepatology* **16**, 649-654.

Diez, J., Mateu, M. G. & Domingo, E. (1989). Selection of Antigenic Variants of Foot-and-Mouth Disease Virus in the Absence of Antibodies, as Revealed by an in situ Assay. *Journal of General Virology* **70**, 3281-3289.

Domingo, E., Diez, J., Martinez, M. A., Hernandez, J., Holguin, A., Borrego, B. & Mateu, M. G. (1993). New observations on antigenic diversification of RNA viruses. Antigenic variation is not dependent on immune selection. *Journal of General Virology* **74**, 2039-2045.

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

Dusheiko, G. & Simmonds, P. (1994a). Sequence variability of hepatitis C virus and its clinical relevance. *Journal of Viral Hepatitis* **1**, 3-15.

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

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

Enomoto, N., Takada, A., Nakao, T. & Date, T. (1990). There are two major types of hepatitis C virus in Japan. *Biochem Biophys Res Commun.* **170**, 1021-1025.

Esteban, R. (1993). Epidemiology of hepatitis C virus infection. *Journal of Hepatology* **17** (Suppl. 3), S67-S71.

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

Farci, P. & Purcell, R. H. (1993). Hepatitis C virus: Natural history and experimental models. In *Viral Hepatitis: Scientific basis and clinical mamagement*, pp. 241-267. Edited by A. J. Zuckerman & H. C. Thomas. Churchill Livingstone.

Farci, P., Alter, H. J, Govindarajan, S., Wong, D., Engle, R., Lesniewski, R., Mushahwar, I., Desai, S., Miller, R., Ogata, N. & Purcell, R. (1992). Lack of protective immunity against reinfection with hepatitis C virus. *Science* **258**, 135-140.

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

Farci, P., Alter, H. J., Wong, D. C., Miller, R. H., Govindarajan, S., Engle, R., Shapiro, M. & 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.

Feinstone, S. M., Kapikian, A. Z., Purcell, R. H., Alter, H. J. & Holland, P. V. (1975). Transfusion-associated hepatitis not due to viral hepatitis type A or B. *The New England Journal of Medicine* **292**, 767-770.

Felsentein, J. (1993). Phylip Inference Version 3.5. Seattle: Department of Genetics University of Washington.

Fong, T.-L., Charboneau, F., Valinluck, B. & Govindarajan, S. (1993). The stability of serum hepatitis C viral RNA in various handling and storage conditions. *Archives Pathology Laboratory Methods* **117**, 150-151.

Garcia-Barreno, B., Delgado, T. & Melero, J. A. (1994). Oligo(A) Sequences of Human Respiratory Syncytical Virus G Protein Gene: Assessment of Their Genetic Stability in Frameshift Mutants. *Journal of Virology* **68**, 5460-5468.

Gerlich, W. (1993). Hepatitis B virus: Structure and molecular virology. In *Viral Hepatitis: Scientific basis and clinical mamagement.*, pp. 83-113. Edited by A. J. Zuckerman & H. C. Thomas. Churchill Livingstone.

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

Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M. & Rice, C. M. (1993c). A second hepatitis C virus-encoded proteinase. *Proc. Natl. Acad. Sci. USA* **90**, 10583-10587.

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

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

Hagiwara, H., Hayashi, N., Mita, E., Naito, M., Kasahara, A., Fusamoto, H. & Kamada, T. (1993). Quantitation of hepatitis C virus RNA in serum of asymptomatic blood donors and patients with type C chronic liver disease. *Hepatology* **17**, 545-550.

Hahm, B., Han, D. S., Back, S. H., Song, O.-K., Cho, M.-J., Kim, C.-J., Shimotohno, K. & Jang, S. K. (1995). NS3-4A of hepatitis C virus is a chymotrypsin-like protease. *Journal of Virology* **69**, 2534-2539.

Haldor, S. C. & Margolis, H. S. (1991). Viral hepatitis. In *Viral infections of humans. Epidemiology and control.*, pp. 351. Edited by A. S. Evans. Plenum Medical Book Company.

Han, D. S., Hahm, B., Rho, H.-M. & Jang, S. K. (1995). Identification of the protease domain in NS3 of hepatitis C virus. *Journal of General Virology* **76**, 985-993.

Han, J. H., Shyamala, V., Richman, K. H., Brauer, M. J., Irvine, B., Urdea, M. S., Tekamp-Olson, P., Kuo, G., Choo, Q.-L. & Houghton, M. (1991). Characterization of the terminal regions of the hepatitis C viral RNA: Identification of conserved sequences in the 5' untranslated region and poly(A) tails at the 3' end. *Proc.Natl. Acad. sci. USA* **88**, 1711-1715.

Harada, S., Watanabe, Y., Takeuchi, K., Suzuki, T., Katayama, T., Yutaka, Takebe, Saito, I. & Miyamura, T. (1991). Expression of processed core protein of hepatitis C virus in mammalian cells. *Journal of Virology* **65**, 3015-3021.

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

Helenius, A. (1995). Alphavirus and Flavivirus Glycoproteins: Structures and Functions. *Cell* **81**, 651-653.

Higashi, Y., Kakumu, S., Yoshioka, K., Wakita, T., Mizokami, M., Ohba, K., Ito, Y., Ishikawa, T., Takayanagi, M. & Nagai, Y. (1993). Dynamics of genome change in the E2/NS1 region of hepatitis C virus in vivo. *Virology* **197**, 659-668.

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

Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S. & Shimotohno, K. (1991b). Hypervariable regions in the putative glycoprotein of hepatitis C virus. *Biochem. Biophy. Res. Commun.* **175**, 220-228.

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

Hijikata, M., Mizushima, H., Tanji, Y., Komoda, Y., Hirowatari, Y., Akagi, T., Kato, N., Kimura, K. & Shimotohno, K. (1993b). Proteolytic processing and membrane association of putative nonstructural proteins of hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **90**, 10773-10777.

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

Hino, K., Sainokami, S., Shimoda, K., Iino, S., Wang, Y., Okamoto, H., Miyakawa, Y. & Mayumi, M. (1994). Genotypes and Titers of Hepatitis C Virus for Predicting Response to Interferon in Patients With Chronic Hepatitis C. *Journal of Medical Virology* **42**, 299-305.

Holland, J. J., De La Torre, J. C., & Steinhauer, D. A. (1992). RNA virus populations as quasispecies. *Current Topics in Microbiology and Immunology* **176**, 1-20.

Hollinger, F. B., Mosley, J. W., Szmuness, W., Aach, R. D., Peters, R. L. & Stevens, C. (1980). Transfusion-transmitted viruses study: Experimental evidence for two non-A, non-B hepatitis agents. *The Journal of Infectious Diseases* **142**, 400-407.

Houghton, M., Weiner, A., Han, J., Kuo, G. & Choo, Q.-L. (1991). Molecular biology of the hepatitis C viruses: Implications for diagnosis, development and control of viral disease. *Hepatology* **14**, 381-388.

Inchauspe, G., Zebedee, S., Lee, D.-H., Sugitiani, M., Nasoff, M. & Prince, A. M. (1991). Genomic structure of the human prototype strain H of hepatitis C virus: Comparison with the American and Japanese isolates. *Proc. Natl. Acad. Sci. USA* **88**, 10292-10296.

Inoue, Y., Takeuchi, K., Chou, W.-H., Unayama, T., Takahasi, K., Saito, I. & Miyamura, T. (1992). Silent mother-to-child transmission of hepatitis C virus through two generations determined by comparative nucleotide sequence analysis of the viral cDNA. *Journal of Infectious Disease* **166**, 142-1428.

Jin, H. & Elliott, R. M. (1993). Characterization of Bunyamwera virus S RNA that is transcribed and replicated by the L protein expressed from recombinant vaccinia virus. *Journal of Virology* **67**, 1396-1404.

Kaito, M., Watanabe, S., Tsukiyama-Kohara, K., Yamaguchi, K., Kobayashi, Y., Konishi, M., Yokoi, M., Ishida, S., Suzuki, S. & Kohara, M. (1994). Hepatitis C virus particle detected by immunoelectron microscopic study. *Journal of General Virology* **75**, 1755-1760.

Kaneko, S., Murakami, S., Unoura, M. & Kobayashi, K. (1992). Quantitiation of hepatitis C virus RNA by competitive polymerase chain reaction. *Journal of Medical Virology* **37**, 278-282.

Kanto, T., Hayashi, N., Takehara, T., Hagiwara, H., Mita, E., Naito, M., Kasahara, A., Fusamoto, H. & Kamada, T. (1995). Density analysis of hepatitis C virus particle population in the circulation of infected hosts: implications for virus neutralisation or persistence. *Journal of Hepatology* **22**, 440-448.

Kato, N., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T. & Shimotohno, K. (1990). Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. *Proc. Natl. Acad. Sci. USA* **87**, 9524-9528.

Kato, N., Ootsuyama, Y., Ohkoshi, S., Nakazawa, T., Sekiya, H., Hijikata, M. & Shimotohno, K. (1992b). Characterisation of hypervariable regions in the putative envelope protein of hepatitis C virus. *Biochem. Biophy. Res. Commun.* **189**, 119-127.

Kato, N., Ootsuyama, Y., Sekiya, H., Ohkoshi, S., Nakazawa, T., Hijikata, M. & Shimotohno, K. (1994). Genetic drift in hypervariable region 1 of the viral genome in persistent hepatitis C virus infection. *Journal of Virology* **68**, 4776-4784.

Kato, N., Ootsuyama, Y., Tanaka, T., Nakagawa, M., Nakazawa, T., Muraiso, K., Ohkoshi, S., Hijikata, M. & Shimotohno, K. (1992a). Marked sequence diversity in the putative envelope proteins of hepatitis C viruses. *Virus Research* **22**, 107-123.

Kato, N., Sekiya, H., Ootsuyama, Y., Nakazawa, T., Hijikata, M., Ohkoshi, S. & Shimotohno, K. (1993a). Humoral immune response to hypervariable region 1 of the putative envelope glycoprotein (gp70) of hepatitis C virus. *Journal of Virology* **67**, 3923-3930.

Kato, N., Yokosuka, O., Hosoda, K., Ito, Y., Ohto, M. & Omata, M. (1993b). Quantification of hepatitis C virus by competitive reverse transcription-polymerase chain reaction: increase of the virus in advanced liver disease. *Hepatology* **18**, 16-20.

Kiyosawa, K., Tanaka, E., Sodeyama, T. & Furuta, S. (1994). Natural history of hepatitis C. *Intervirology* **37**, 101-107.

Klenerman, P., Rowland-Jones, S., McAdam, S., Edwards, J., Daenke, S., Lalloo, D., Koppe, B., Rosenberg, W., Boyd, D., Edwards, A., Giangrande, P., Phillips, R. E. & McMichael, A. (1994). Cytotoxic T-Cell actvity anatagonized by naturally occurring HIV-1 Gag variants. *Nature* **369**, 403-407.

Kobayashi, Y., Watanabe, S., Konishi, M., Yokoi, M., Kakehashi, R., Kaito, M., Kondo, M., Hayashi, Y., Jomori, T. & Suzuki, S. (1993). Quantitiation and typing of serum hepatitis C virus RNA in patients with chronic hepatitis C treated with interferon-b. *Hepatology.* **18**, 1319-1325.

Kojima, M., Osuga, T., Tsuda, F., Tanaka, T. & Okamoto, H. (1994). Influence of antibodies to the hypervariable region of E2/NS1 glycoprotein on the selective replication of hepatitis C virus in chimpanzees. *Virology* **204**, 665-672.

Kolykhalov, A., Feinstone, S.M., & Rice, C. M. (1996). Identification of a highly conserved sequence element at the 3' terminus of hepatitis C virus genome RNA. *Journal of Virology* **70**, 3363-3371.

Koonin, E. (1991). The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J Gen Virol* **72**, 2197-2206.

Kumar, S., Tamura, K. & Nei, M. (1993). MEGA: Molecular Evolutionary Genetics Analysis Version 1.01. The Pennsylvania State University, University Park, PA 16802.

Kumar, U., Brown, J., Monjardino, J. & Thomas, H. C. (1993). Sequence variation in the large envelope glycoprotein (E2/NS1) of hepatitis C virus during chronic infection. *The Journal of Infectious Diseases* **167**, 726-730.

Kumar, U., Monjardino, J. & Thomas, H. C. (1994). Hypervariable region of hepatitis C virus Envelope glycoprotein (E2/NS1) in an agammaglobulinemic patient. *Gastroenterology* **106**, 1072-1075.

Kuo, G. & Houghton, M. (1992). Characterization of the hepatitis C virus E2/NS1 gene product expressed in mammalian cells. *Virology* **188**, 819-830.

Kuo, G., Choo, Q.-L., Alter, H. J., Gitnick, G. L., Redeker, A. G., Purcell, R. H., Miyamura, T., Dienstag, J. L., Alter, M. J., Stevens, C. E., Tegtmeier, G. E., Bonino, F., Colombo, M., Lee, W.-S., Kuo, C., Berger, K., Shuster, J. R., Overby, L. R., Bradley, D. W. & Houghton, M. (1989). An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science* **244**, 362-364.

Kurosaki, M., Enomoto, N., Marumo, F. & Sato, C. (1993). Rapid sequence variation of the hypervariable region of hepatitis C virus during the course of chronic infection. *Hepatology* **18**, 1293-1299.

Kurosaki, M., Enomoto, N., Marumo, F. & Sato, C. (1994). Evolution and selection of hepatitis C virus variants in patients with chronic hepatitis C. *Virology* **205**, 161-169.

Lanford, R. E., Notvall, L., Chavez, D., White, R., Frenzel, G., Simonsen, C. & Kim, J. (1993). Analysis of hepatitis C virus capsid, E1 and E2/NS1 proteins expressed in insect cells. *Virology* **197**, 225-235.

Lemon, S. M. (1994). Hepatitis A virus. In *Encyclopedia of Virology*, pp. 546. Edited by R. G. Webster & A. Granoff: Academic Press.

Lemon, S. M. & Brown, E. A. (1995). Hepatitis C virus. In *Principles and Practice of Infectious Diseases.*, 4th edn, pp. 1474-1486. Edited by G. L. Mandell, J. E. Bennett & R. Dolin: Churchill Livingstone.

Lesniewski, R. R., Boardway, K. M., Casey, J. M., Desai, S. M., Devare, S. G., Leung, T. K. & Mushahwar, I. K. (1993). Hypervariable 5'terminus of hepatitis C virus E2/NS1 encodes antigenically distinct variants. *Journal of Medical Virology* **40**, 150-156.

Liang, T. J., Jeffers, L. J., Reddy, K. R., Medina, M. D., Parker, I. T., Cheinquer, H., Idrovo, V., Rabassa, A. & Schiff, E. R. (1993). Viral pathogenesis of hepatocellular carcinoma in the United States. *Hepatology* **18**, 1326-1333.

Lin, C., Lindenbach, B. D., Pragai, B. M., McCourt, D. W. & Rice, C. M. (1994). Processing in the hepatitis C virus E2-NS2 region: Identification of p7 and two distinct E2-specific products with different C termini. *Journal of Virology* **68**, 5063-5073.

Linnen, J., Wages Jr, J., Zhang-Keck, Z.- Y., Fry, K. E., Krawczynski, K. Z., Alter, H., Koonin, E., Gallagher, M., Alter, M., Hadziyannis, S., Karayiannis, P., Fung, K., Nakatsuji, Y., Shih, J. W.- K., Young, L., Piatak Jr, M., Hoover, C., Fernandez, J., Chen, S., Zou, J.- C., Morris, T., Hyams, K. C., Ismay, S., Lifson, J. D., Hess, G., Foung, S. K. H., Thomas, H., Bradley, D., Margolis, H. & Kim, J. P. (1996). Molecular cloning and disease association of hepatitis G virus: A transfusion-transmissible agent. *Science* **271**, 505-508.

Manniatis, T., Fritsch, E. F. & Sambrook, J. (1982). "Molecular Cloning; A Laboratory Manual". N.Y: Cold Spring Harbour Laboratory.

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

Matsuura, Y., Harada, S., Suzuki, R., Watanabe, Y., Inoue, Y., Saito, I. & Miyamura, T. (1992). Expression of processed envelope protein of hepatitis C virus in mammalian and insect cells. *Journal of Virology* **66**, 1425-1431.

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

McDonnell, W. M. & Askari, F. K. (1996). Molecular Medicine: DNA Vaccines. *The New England Journal Of Medicine* **334**, 42-45.

McFarlane, I. G., Smith, H. M., Johnson, P. J., Bray, G. P., Vergani, D. & Williams, R. (1990). Hepatitis C virus antibodies in chronic active hepatitis: pathogenic factor or false-positive result? *The Lancet* **335**, 754-757.

McIntyre, P., McCruden, E. A. B., Dow, B. C., Cameron, S. O., McMillan, M. A., Allison, M. E. M. & Briggs, J. D. (1994). Hepatitis C virus infection in renal dialysis patients in Glasgow. *Nephrology Dialysis Transplanation.* **9**, 291-295.

Mehdi, H., Kaplan, M. J., Anlar, F. Y., Yang, X., Bayer R., Sutherland, K., & Peeples, M. E. (1994). Hepatitis B virus surface antigen binds to apolipoprotein H. *Journal of Virology* **68**, 2415-2424.

Mehdi, H., Yang, X., & Peeples, M. E. (1996). An altered form of apolipoprotein H binds virus surface antifen most efficiently. *Virology* **217**, 58-66.

Meier, U.-C., Klenerman, P., Griffin, P., James, W., Koppe, B., Larder, B., McMichael, A. & Phillips, R. (1995). Cytotoxic T Lymphocytes Lysis Inhibited by Viable HIV Mutants. *Science* **270**, 1360-1362.

Mellor, J., Holmes, E. C., Jarvis, L. M., Yap, P. L., Simmonds, P. And The International HCV Collaborative Study Group. (1995). Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification. *Journal of General Virology* **76**, 2493-2507.

Miller, R. H. & Purcell, R. H. (1990). Hepatits C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. *Proc. Natl. Acad. Sci. USA* **87**, 2057-2061.

Miyamoto, H., Okamoto, H., Sato, K., Tanaka, T. & Mishiro, S. (1992). Extraordinarily low density of hepatitis C virus estimated by sucrose density gradient centrifugation and the polymerase chain reaction. *Journal of General Virology* **73**, 715-718.

Mizokami, M., Gojobori, T., Ohba, K.-I., Ikeo, K., Ge, X.-M., Ohno, T., Orito, E., & Lau, J. Y. N. (1996). Hepatitis C virus types 7, 8 and 9 should be classified as type 6 subtypes. *Journal of Hepatology* **24**, 622-624.

Mizushima, H., Hijikata, M., Asabe, S.-I., Hirota, M., Kimura, K. & Shimotohno, K. (1994b). Two hepatitis C virus glycoprotein E2 products with different C termini. *Journal of Virology* **68**, 6215-6222.

Mizushima, H., Hijikata, M., Tanji, Y., Kimura, K. & Shimotohno, K. (1994a). Analysis of N-terminal processing of hepatitis C virus nonstructural protein 2. *Journal of Virology* **68**, 2731-2734.

Montefiori, D. C., Zhou, J., Barnes, B., Lake, D., Hersh, E. M., Masuho, Y. & Jr., L. B. L. (1991). Homotypic antibody responses to fresh clinical isolates of human immunodeficiency virus. *Virology* **182**, 635-643.

Moore, J. P., Cao, Y., Conley, A. J., Wyatt, R., Robinson, J., Gorny, M. K., Zolla-Pazner, S., Ho, D. D. & Koup, R. A. (1994). Studies with Monoclonal Antibodies to the V3 Region of HIV-1 gp120 Reveal Limitations to the Utility of Solid-Phase Peptide Binding Assays. *Journal of Acquired Immune Deficiency Syndromes* **7**, 332-339.

Mullis, K. & Faloona, F. (1987). Specific synthesis of DNA in vitro via a polymerase-catalysed chain reaction. *Methods Enzymology* **155**, 335-350.

Muerhoff, A. S., Leary, T. P., Simons, J. N., Pilot-Matias, T. P., Dawson, G. J., Erker, J. C., Chalmers, M. L., Schlauder, G. G., Desai, S. M. & Mushahwar, I. K. (1995). Genomic Organization of GB viruses A and B: Two new members of the *Flaviviridae* associated woth GB agent hepatitis. *Journal of Virology* **69**, 5621-5630.

Nagayama, R., Tsuda, F., Okamoto, H., Wang, y., Mitsui, T., Tanaka, T., Miyakawa, Y. & Mayumi, M. (1994). Diagnostic and clinical implications of the different genotypes of hepatitis C virus. *Hepatology* **20**, 256-259.

Naito, M., Hayashi, N., Hagiwara, H., Katayama, K., Kasahara, A., Fusamoto, H., Kato, M., Masuzawa, M. & Kamada, T. (1994). Serial quantitative analysis of serum hepatitis C virus RNA level in patients with acute and chronic hepatitis C. *Journal of Hepatology* **20**, 755-759.

Nara, P. L., Smit, L., Dunlop, N., Hatch, W., Merges, M., Waters, D., Kelliher, J., Gallo, R. C., Fischinger, P. J. & Goudsmit, J. (1990). Emergence of viruses resistent to neutralisation by V3-specific antibodies in experimental human immunodeficiency virus type 1 IIIB infection of chimpanzees. *Journal of Virology* **64**, 3779-3791.

Negro, F., Pacchioni, D., Shimizu, Y., Miller, R. H., Bussolati, G., Purcell, R. H., & Bonino, F. (1992). Detection of intrahepatic replication of hepatitis C virus RNA by in situ hybridization and comparison with histopathology. *Proc. Natl. Acad. Sci. USA* **89**, 2247-2251.

Nowak, T., Faber, P. M., Wengler, G. & Wengler, G. (1989). Analyses of the terminal sequences of West Nile structural proteins and of the in vitro translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. *Virology* **169**, 365-376.

Ogata, N., Alter, H. J., Miller, R. H. & Purcell, R. H. (1991). Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **88**, 3392-3396.

Ohto, H., Terazawa, S., Sasaki, N., Sasaki, N., Hino, K., Ishiwata, C., Kako, M., Ujiie, N., Endo, C., Matsui, A., Okamoto, H. & Mishiro, S. (1994). Transmission of hepatitis C virus from mothers to infants. *The New England Journal of Medicine* **330**, 744-756.

Okada, S.-I., Akahane, Y., Suzuki, H., Okamoto, H. & Mishiro, S. (1992). The degree of variability in the amino terminal region of the E2/NS1 protein of hepatitis C virus correlates with responsiveness to interferon therapy in viraemic patients. *Hepatology* **16**, 619-624.

Okamoto, H. & Mishiro, S. (1994a). Genetic heterogeneity of hepatitis C virus. *Intervirology* **37**, 68-76.

Okamoto, H., Kojima, M., Okada, S.-I., Yoshizawa, H., Iizuka, H., Tanaka, T., Mucnmore, E. E., Peterson, D. A., Ito, Y. & Mishiro, S. (1992b). Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability. *Virology* **190**, 894-899.

Okamoto, H., Kojima, M., Sakamoto, M., Iizuka, H., Hadiwandowo, S., Suwignyo, S., Miyakawa, Y. & Mayumi, M. (1994b). The entire nucleotide sequence and classification of a hepatitis C virus isolate of a novel genotype from an Indonesian patient with chronic liver disease. *Journal of General Virology* **75**, 629-635.

Okamoto, H., Kurai, K., Okada, S.-I., Yamamoto, K., Lizuka, H., Tanaka, T., Fukuda, S., Tsuda, F. & Mishiro, S. (1992a). 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.

Okamoto, H., Okada, S., Sugiyama, Y., Kurai, K., Iizuka, H., Machida, A., Miyakawa, Y. & Mayumi, M. (1991). Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: Comparison with reported isolates for conserved and divergent regions. *Journal of General Virology* **72**, 2697-2704.

Oldstone, M. B. A. (1989). Viral Persistence. *Cell* **56**, 517-520.

Oshima, M., Tsuchiya, M., Yagasaki, M., Orita, T., Hasegawa, M., Tomonoh, K., Kojima, T., Hirata, Y., Yamamoto, O., Sho, Y., Maeda, E. & Arima, T. (1991). cDNA clones of Japanese hepatitis C virus genomes derived from a single patient show sequence heterogeneity. *Journal of General Virology* **72**, 2805-2809.

Phillips, R. E., Rowland-Jones, S., Nixon, D. F., Gotch, F. M., Edwards, J. P., Ogunlesi, A. O., Elvin, J. G., Rothbard, J. A., Bangham, C. R. M., Rizza, C. R. & McMichael, A. J. (1991). Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature* **354**, 453-459.

Poole, T., Wang, C., Popp, R. A., Potgieter, L. N. D., Siddiqui, A. & Collett, M. S. (1995). Pestivirus Translation Initiation Occurs by Internal Ribosome Entry. *Virology* **206**, 750-754.

Prieto, M., Olaso, V., Verdu, C., Cordoba, J., Gisbert, C., Rayon, M., Carrasco, D., Berenguer, M., Higon, M. D. & Berenguer, J. (1995). Does the Healthy Hepatitis C Virus Carrier State Really Exist? An Analysis Using Polymerase Chain Reaction. *Hepatology* **22**, 413-417.

Prince, A. M., Brotman, B., Grady, G. F., Kuhns, W. J., Hazzi, C., Levine, R. W. & Millian, S. J. (1974). Long-incubation post-transfusion hepatitis without serological evidence of exposure to hepatitis B virus. *The Lancet* **ii**, 241-246.

Purcell, R. H. (1994). Hepatitis viruses: Changing patterns of human disease. *Proc. Natl. Acad. Sci. USA* **91**, 2401-2406.

Ralston, R., Thudium, K., Berger, K., Kuo, C., Gervase, B., Hall, J., Selby, M., Kuo, G., Houghton, M. & Choo, Q.-L. (1993). Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. *Journal of Virology* **67**, 6753-6761.

Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C. & Harrison, S. C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2A resolution. *Nature* **375**, 291-298.

Reynolds, J., Grace, K., Clarke, B., Rowlands, D., Kaminski, A & Jackson, R. (1994). Features of HCV RNA involved in the initiation of protein translation. Abstract. 2bd International Meeting on Hepatitis C and Related Viruses. University of California, San Diego.

Rice, C. M. & Strauss, J. H. (1990). Production of flavivirus polypeptides by proteolytic processing. *Seminars in Virology* **1**, 357-367.

Rocha, E., Cox, N. J., Black, R. A., Harmon, M. W., Harrison, C. J. & Kendal, A. P. (1991). Antigenic and Genetic Variation in Influenza A (H1N1) Virus Isolates Recovered from a Persistently Infected Immunodeficient Child. *Journal of Virology* **65**, 2340-2350.

Rosa, D., Campagnoli, S., Moretto, C., Guenzi, E., Cousens, L., Chin, M., Dong, C., Weiner, A. J., Lau, J. Y. N., Choo, Q.-L., Chien, D., Pileri, P., Houghton, M. & Abrignani, S. (1996). A quantitative test to estimate neutralizing antibodies to the hepatitis C virus: Cytofluorimetric assassment of envelope glycoprotein 2 binding to target cells. *Proc. Nalt. Acad. Sci. USA* **93**, 1759-1763.

Saiki, R., Scarf, S., Faloona, F., Mullis, K., Horh, G., Erlich, H. & Arnheim, N. (1985). Enzymatic amplification of b-globin genmic sequence and restriction site analysis for diagnosis of sickle cell anaemia. *Science* **230**, 1350-1354.

Saito, H., Miyamura, T., Ohbayashi, A., Harada, H., Katayama, T., Kikuchi, S., Watanabe, Y., Koi, S., Onji, M., Ohta, Y., Choo, Q.-L., Houghton, M. & Kuo, G. (1990). Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. *Proc. Nalt. Acad. Sci. USA* **87**, 6547-6549.

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

Sakamoto, N., Enomoto, N., Kurosaki, M., Marumo, F. & Sato, C. (1994). Sequential change of the hypervariable region of the hepatitis C virus genome in acute infection. *Journal of Medical Virology* **42**, 103-108.

Salinovich, O., Payne, S. L., Montelaro, R. C., Hussain, K. A., Issel, C. J. & Schnorr, K. L. (1986). Rapid Emergence of Novel Antigenic and Genetic Variants of Equine Infectious Anaemia Virus during Persistent Infection. *Journal of Virology* **57**, 71-80.

Sanchez-Palomino, S., Rojas, J. M., Martinez, M. A., Fenyo, E. M., Najera, R., Domingo, E. & Lopez-Galindez, C. (1993). Dilute passage promotes expression of genetic and phenotypic variants of human immunodeficiency virus type 1 in cell culture. *Journal of Virology* **67**, 2938-2943.

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Nalt. Acad. Sci. USA.* **74**, 5463-5467.

Santolini, E., Migliaccio, G. & Monica, N. L. (1994). Biosynthesis and biochemical properties of the hepatitis C virus core protein. *Journal of Virology* **68**, 3631-3641.

Sato, K., Okamoto, H., Aihara, S., Hoshi, Y., Tanaka, T. & Mishiro, S. (1993). Demonstration of a sugar moiety on the surface of hepatitis C virions recovered from the circulation of infected humans. *Virology* **196**, 354-357.

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

Scarselli, E., Cerino, A., Esposito, G., Silini, E., Mondelli, M. U. & Traboni, C. (1995). Occurrence of Antibodies Reactive with More then One Variant of the Putative Envelope Glycoprotein (gp70) Hypervariable Region 1 in Viremic Hepatitis C Virus-Induced Patients. *Journal of Virology* **69**, 4407-4412.

Schlauder, G. G., Pilot-Matias, T. J., Gabriel, G. S., Simons, J. N., Muerhoff, A. S., Dawson, G. J. & Mushahwar, I. K. (1995). Origin of GB-hepatitis. *The Lancet* **346**, 447-448.

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

Selby, M. J., Glazer, E., Masiarz, F. & Houghton, M. (1994). Complex processing and protein:protein interactions in the E2:NS2 region of HCV. *Virology* **204**, 114-122.

Shimizu, Y. K., Feinstone, S. M., Kohara, M., Purcell, R. H., & Yoshikara, H. (1996). Hepatitis C virus: Detection of intracellular virus particles by electron microscopy. *Hepatology* **23**, 205-209.

Shimizu, Y. K., Feinstone, S. M., Purcell, R. H., Alter, H. J. & London, W. T. (1979). Non-A, non-B hepatitis: ultrastructural evidence for two agents in experimentally infected chimpanzees. *Science* **205**, 197-200.

Shimizu, Y. K., Hijikata, M., Iwamoto, A., Alter, H. J., Purcell, R. H. & Yoshikura, H. (1994). Neutralizing Antibodies against Hepatitis C Virus and the Emergence of Neutralization Escape Mutant Viruses. *Journal of Virology* **68**, 1494-1500.

Shimizu, Y. K., Iwamoto, A., Hijikata, M., Purcell, R. H. & Yoshikura, H. (1992). Evidence for in vitro replication of hepatitis C virus genome in a human T-cell line. *Proc. Natl. Acad. Sci. USA* **89**, 5477-5481.

Shimizu, Y. K., Weiner, A. J., Rosenblatt, J., Wong, D. C., Shapiro, M., Popkin, T., Houghton, M., Alter, H. J. & Purcell, R. H. (1990). Early events in hepatitis C virus infection of chimpanzees. *Proc. Nalt. Acad. Sci. USA* **87**, 6441-6444.

Shindo, M., Arai, K., Sokawa, Y. & Okuno, T. (1995). The Virological and Histological States of Anti-Hepatitis C Virus-Positive Subjects With Normal Liver Biochemical Values. *Hepatology* **22**, 418-425.

Shirai, M., Okada, H., Nishioka, M., Akatsuka, T., Wychowski, C., Houghten, R., Pendleton, C. D., Feinstone, S. M. & Berzofsky, J. A. (1995). An Epitope in Hepatitis C Virus Core Region Recognized by Cytotoxic T Cells in Mice and Humans. *Journal of Virology* **68**, 3334-3342.

Simmonds, P., Balfe, P., Ludlam, C. A., Bishop, J. O. & Brown, A. J. L. (1990b). Analysis of Sequence Diversity in Hypervariable Regions of the External Glycoprotein of Human Immunodeficiency Virus Type 1. *Journal of Virology* **64**, 5840-5850.

Simmonds, P., Holmes, E. C., Cha, T.-A., Chan, S.-W., McOmish, F., Irvine, B., Beall, E., Yap, P. L., Kolberg, J. & Urdea, M. S. (1993b). Classfication of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS5 region. *Journal of General Virology* **74**, 2391-2399.

Simmonds, P., McOmish, F., Yap, P. L., Chan, S.-W., Lin, C. K., Dusheiko, G., Saeed, A. A. & Holmes, E. C. (1993a). Sequence variability in the 5' non-coding region of hepatitis C virus: identification of a new virus type and restrictions on sequence diversity. *Journal of General Virology* **74**, 661-668.

Simmonds, P., Zhang, L. Q., Watson, H. G., Rebus, S., Ferguson, E. D., Balfe, P., Leadbetter, G. H., Yap, P. L., Peutherer, J. F. & Ludlam, C. A. (1990). Hepatitis C quantification and sequencing in blood products, haemophiliacs, and drug users. *The Lancet* **336**, 1469-1472.

Simons, J. N., Pilot-Matias, T. J., Leary, T. P., Dawson, G. J., Desai, S. M., Schlauder, G. G., Muerhoff, A. S., Erker, J. C., Buijk, S. L., Chalmers, M. L., Sant, C. L. V. & Mushahwar, I. K. (1995). Identification of two flavivirus-like genomes in the GB hepatitis agent. *Proc. Nalt. Acad. Sci. USA* **92**, 3401-3405.

Smith, D. B., Mellor, J., Jarvis, L. M., Davidson, F., Kolberg, J., Urdea, M., Yap, P.-L., Simmonds, P., Conradie, J. D., Neill, A. G. S., Dusheiko, G. M., Kew, M. C., Crookes, R., Koshy, A., Lin, C. K., Lai, C., Murray Lyon, I. M., El Guneid, A., Gunaid, A. A., *et al.* (1995). Variation of the hepatitis C virus 5' non-coding region: Implications for secondary structure, virus detection and typing. *Journal of General Virology* **76**, 1749-1761.

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

Spaete, R. R., Alexander, D. A., Rugroden, M. E., Choo, Q.-L., Berger, K., Crawford, K., Kuo, C., Leng, S., Lee, C., Ralston, R., Thudium, K., Tung, J. W., Kuo, G. & Houghton, M. (1992). Characterization of the hepatitis C virus E2/NS1 gene product expressed in mammalian cells. *Virology* **188**, 819-830.

Steinhauer, D. A. & Holland, J. J. (1987). Rapid evolution of RNA viruses. *Annual Review of Microbiology* **41**, 409-33.

Stuyver, L., Rossau, R., Wyseur, A., Duhamel, M., Vanderborght, B., Heuverswyn, H. v. & Maertens, G. (1993). Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay. *Journal of General Virology* **74**, 1093-1102.

Tabor, E., Gerety, R. J., Drucker, J. A., Seeff, L. B., Hoofnagle, J. H., Jackson, D. R., April, M., Barker, L. F. & Pineda-Tamondong, G. (1978). Transmission of non-A, non-B hepatitis from man to chimpanzee. *The Lancet* **i**, 463-466.

Tabor, S. & Richardson, C. C. (1989). Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by *in vitro* mutagenesis. *J.ournal Biological Chemistry* **264**, 6447-6458.

Takada, N., Takase, S., Enomoto, N., Takada, A. & Date, T. (1992). Clinical backfrounds of the patients having different types of hepatitis C virus genomes. *Journal of Hepatology* **14**, 35-40.

Takahashi, K., Kishimoto, S., Yoshizawa, H., Okamoto, H., Yoshikawa, A. & Mishiro, S. (1992a). p26 protein and 33-nm particle associated with nucleocapsid of hepatitis C virus recovered from the circulation of infected hosts. *Virology* **191**, 431-434.

Takahashi, K., Okamoto, H., Kishimoto, S., Munekata, E., Tachibana, K., Akahane, Y., Yoshizawa, H. & Mishiro, S. (1992b). Demonstration of a hepatitis C virus-specific antigen predicted from the putative core gene in the circulation of infected hosts. *Journal of General Virology* **73**, 667-672.

Takamizawa, A., Mori, C., Fuke, I., Manabe, S., Murakami, S., Fujita, J., Onishi, E., Andoh, T., Yoshida, I. & Okayama, H. (1991). Structure and organisation of the hepatitis C virus genome isolated from human carriers. *Journal of Virology* **65**, 1105-1113.

Takeuchi, K., Kubo, Y., Boonmar, S., Watanabe, Y., Katayama, T., Choo, Q.-L., Kuo, G., Houghton, M., Saito, I. & Miyamura, T. (1990b). The putative nucleocapsid and envelope protein genes of hepatitis C virus determined by comparison of the nucleotide sequences of two isolates derived from an experimentally infected chimpanzee and healthly human carriers. *Journal of General Virology* **71**, 3027-3033.

Tanaka, T., Kato, N., Cho, M.-J., Sugiyama, K., & Shimotohno, K. (1996). Structure of the 3' terminus of the hepatitis C virus genome. *Journal of Virology* **70**, 3307-3312.

Tanaka, T., Kato, N., Hijikata, M., & Shimotohno, K. (1992). Base transitions and base transversions seen in mutations among various types of the hepatitis C viral genome. *FEBS Letters* **315**, 201-203.

Tanaka, T., Kato, N., Nakagawa, M., Ootsuyama, Y., Cho, M.-J., Nakazawa, T., Hijikata, M., Ishimura, Y. & Shimotohno, K. (1992). Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals. *Virus Research* **23**, 39-53.

Taniguchi, S., Okamoto, H., Sakamoto, M., Kojima, M., Tsuda, F., Tanaka, T., Munekata, E., Muchmore, E. E., Peterson, D. A. & Mishiro, S. (1993). A structurally flexible and antigenically variable N-terminal domain of the hepatitis C virus E2/NS1 protein: Implication for an escape from antibody. *Virology* **195**, 297-301.

Thomssen, R., Bonk, S. & Thiele, A. (1993). Density heterogeneities of hepatitis C virus in human sera due to the binding of b-lipoproteins and immunoglobulins. *Medical Microbiology and Immunology* **182**, 329-334.

Thomssen, R., Bonk, S., Propfe, C., Heermann, K.-H., Kochel, H. G. & Uy, A. (1992). Association of hepatitis C virus in human sera with b-lipoprotein. *Medical Microbiology and Immunology* **181**, 293-300.

Tokita, H., Okamoto, H., Luengrojanakul, P., Vareesangthip, K., Chainuvati, T., Iizuka, H., Tsuda, F., Miyakawa, Y., & Mayumi, M. (1995). Hepatitis C virus variants from Thailand classifiable into five novel genotypes in the sixth (6b), seventh (7c, 7d) and ninth (9b, 9c) major genetic groups. *Journal of General Virology* **76**, 2329-2335.

Tokita, H., Okamoto, H., Tsuda, F., Song, P., Nakata, S., Chosa, T., Iizuka, H., Mishiro, S., Miyakawa, Y., & Mayumi, M. (1994). Hepatitis C virus variants from Vietnam are classifiable into the seventh, eighth and ninth major genetic groups. *Proc. Natl. Acad. Sci. USA* **91**, 11022-11026.

Tomei, L., Failla, C., Santolini, E., Francesco, R. D. & Monica, N. L. (1993). NS3 is a serine proteinase required for processing of the hepatitis C virus polyprotein. *Journal of Virology* **67**, 4017-4026.

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. & Nomoto, A. (1992). Internal ribosome entry site within hepatitis C virus RNA. *Journal of Virology* 66, 1476-1483.

Ulrich, P. P., Romeo, J. M., Lane, P. K., Kelly, I., Daniel, L. J. & Vyas, G. N. (1990). Detection, semiquantitation, and genetic variation in hepatitis C virus sequences amplified from plasma of blood donors with elevated alanine aminotransferase. *Journal Clinical Investigation* **86**, 1609-1614.

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

Walker, C. M. (1996). Cytotoxic T-lymphocyte responses to the hepatitis C virus in humans and chimpanzees. *Seminars in Virology* **7**, 13-21.

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

Wang, C., Sarnow, P. & Siddiqui, A. (1994). A conserved helical element is essential for internal initiation of translation of hepatitis C virus RNA. *Journal of Virology* **68**, 7301-7307.

Wang, J.-T., Wang, T.-H., Sheu, J.-C., Lin, S.-Y., Lin, J.-T. & Chen, D.-S. (1992). Effects of Anticoagulants and Storage of Blood Samples on Efficacy of the Polymerase Chain Reaction Assay for Hepatitis C Virus. *Journal of Clinical Investigation* **30**, 750-753.

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M. & Kawaoka, Y. (1992). Evolution and Ecology of Influenza A viruses. *Microbiological Reviews* **56**, 152-179.

Weiner, A. J., Brauer, M. J., Rosenblatt, J., Richman, K. H., Tung, J., Crawford, K., Bonino, F., Saracco, G., Choo, Q.-L., Houghton, M. & Han, J. H. (1991). Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins. *Virology* **180**, 842-848.

Weiner, A. J., Geysen, H. M., Christopherson, C., Hall, J. E., Mason, T. J., Sarcco, G., Bonino, F., Crawford, K., Marion, C. D., Crawford, K. A., Brunetto, M., Barr, P. J., Miyamura, T., McHutchinson, J. & Houghton, M. (1992). Evidence of immune selection of hepatitis C virus (HCV) putative envelope glycoprotein variants: Potential role in chronic HCV infections. *Proc. Natl. Acad. Sci. USA* **89**, 3468-3472.

Weiner, A., Erickson, A. L., Kansopon, J., Crawford, K., Muchmore, E., Hughes, A. L., Houghton, M. & Walker, C. M. (1995). Persistent hepatitis C virus infection in a chimpanzee is associated with emergence of a cytotoxic T lymphocyte escape variant. *Proc.Natl. Acad. Sci USA* **92**, 2755-2759.

Wiley, D. C. & Skehel, J. J. (1987). The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annual reviews of Biochemistry* **56**, 365-394.

Yuasa, T., Ishikawa, G., Manabe, S.-I., Sekiguchi, S., Takeuchi, K. & Miyamura, T. (1991). The particle size of non-A, non-B hepatitis virus by filtration through microporous regenerated cellulose fibre. *Journal of General Virology* **72**, 2021-2024.

Zhang, L.-Q., MacKenzie, P., Cleland, A., Holmes, E. C., Brown, A. J. L. & Simmonds, P. (1993). Selection for Specific Sequences in the External Envelope Protein of HIV-Type 1 Upon Primary Infection. *Journal of Virology* **67**, 3345-3356.

Zibert, A., Schreier, E. & Roggendorf, M. (1995). Antibodies in Human Sera to Hypervariable Region 1 of Hepatitis C Virus Can Block Viral Attachment. *Virology* **208**, 653-661.