uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Song Liu

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Microbiology and Immunology)

GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

The Hepatitis C Virus Persistence: Immunoglobulin Mimicry by E2 Protein

TITRE DE LA THÈSE / TITLE OF THESIS

Earl Brown

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

Yu-Wen Hu

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Karen Copeland

William Cameron

Gary W. Slater

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

THE HEPATITIS C VIRUS PERSISTENCE :

IMMUNOGLOBULIN MIMICRY BY E2 PROTEIN

A thesis submitted to the School of Graduate Studies

University of Ottawa

In partial fulfilment of the requirements for the degree

Master of Science

Department of Biochemistry, Microbiology and Immunology

Faculty of Medicine

By Song Liu

Your file Votre référence
ISBN: 978-0-494-14921-8
Our file Notre référence
ISBN: 978-0-494-14921-8

ABSTRACT

The mechanisms by which hepatitis C virus (HCV) establishes persistent infection in the majority of infected individuals are unknown, but are thought to represent a complex interplay between viral diversity and host immunity. Currently, the most accepted hypothesis is that genetic variation within hypervariable region 1 (HVR1) of glycoprotein E2 may affect recognition by the immune system, leading to chronic infection. However, it can not explain how escape mutants evade subsequent immune recognition. Here, sequences within HCV E2 genes were analyzed during the course of primary HCV infections. We show that E2 possesses sequences that are homologous to sequences in the variable domains of human immunoglobulin (Ig) light chains in particular but also heavy chains and T cell receptors (TCR), and the degree of similarity of HVR1 to Ig types correlated with immune escape and persistence. In addition, recombinant E2 proteins (first 141 aa or full length) were seen to bind to anti-human IgG antibodies, suggesting that N-terminal region of E2 is antigenically and structurally similar to human Ig variable domains. Overall, these data indicate that immunoglobulin mimicry by HCV E2 plays a significant role in viral immune escape and persistence.

This thesis is dedicated to my parents,

Huiming Li and Maokui Liu,

with gratitude for their continuing support and encouragement,

and with love

ACKNOWLEGEMENTS

I am deeply grateful to my supervisors, Dr. Earl. G. Brown and Dr. Yu-Wen Hu for their unlimited patience, advice and encouragement. This work could not be undertaken without their contributions.

I wish to thank Dr. Ken Dimock and Dr. Doug Gray for their guidance and advice throughout this project.

I would like to express my gratitude to my fellow lab members: Heather Perkins, Teye Omlin, Samra Uzicanin, Shabnam Rahimi and Cathy Curley. Thanks for hours of idea-bouncing and good company.

To my parents and my sister, I owe my deepest thanks for believing in me and supporting me throughout my academic journey. I would also like to thank my boyfriend, Haiming Wang - your love and support mean a great deal to me.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

aa, amino acid
bp, base pair
CDR, complementarity determining regions
CHO, Chinese hamster ovarian
eIF 2α, eukaryotic translation initiation factor
ERK, extracellular signal-regulated kinase
ER, endoplasmic reticulum
FR, framework regions
ELISA, enzyme-linked immunosorbent assay
Fab, fragment antigen binding
Fc , fragment crystallizable
hVAP-33, human vesicle-associated protein
HVR1, hypervariable region 1
HVR2, hypervariable region 2
HCMV, human cytomegalovirus
HAV, hepatitis A virus
HBV, hepatitis B virus
HCV, hepatitis C virus
HIV, immunodeficiency virus
HCC, hepatocellular carcinoma
HSV-1, Herpes Simplex Virus-1
IgSF, immunoglobulin superfamily
IMGT, immunogenetics
Ig, immunoglobulin
IFN, interferon
IL, interleukin
IRES, internal ribosomal entry site
IgVL_{κ} , variable region of immunoglobulin kappa light chain
ISDR, interferon sensitivity-determining region
kDa, kilodalton
LDL, low density lipoprotein
MHC, major histocompatibility complex
NK, natural killer cell
NANBH, non-A, non-B hepatitis
PKR, dsRNA dependent protein kinase
PTB, polypyrimidine tract –binding protein
PePHD, PKR-eIF2α phosphorylation homology domain
RdRp, RNA-dependent RNA polymerase
SSc, systemic sclerosis
Sf9, Spodoptera frugiperda
TCR, T cell receptors
TNF, tumor necrosis factor
UTR, untranslated region

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Historical Aspects

For many years, hepatitis A (HAV) and hepatitis B viruses (HBV), so called 'infectious' and 'serum' hepatitis, were viewed as the major agents of viral hepatitis. Diagnostic tests for hepatitis B virus were first developed in 1965 and for hepatitis A virus in 1973 (Feinstone et al, 1973). Shortly after the diagnostic tests for HAV and HBV were developed, they were both applied to sera from patients with hepatitis. Surprisingly, most of the cases were found due to be neither HAV nor HBV infections, and thereby termed non-A, non-B hepatitis (NANBH). (Feinstone et al, 1975; Knodell et al, 1975; Prince et al, 1974). After the introduction of the HBV blood screen, it became apparent that NANBH was responsible for approximately 90% of transfusion-associated hepatitis cases, and the risk of contracting this disease after transfusion of multiple blood units was as high as 20% in some countries (Alter et al, 1989; Dienstag and Alter, 1986). Not surprisingly, much effort was mounted to identify the causative agent or agents of NANBH, but a frustrating period of years followed, during which the methods used successfully to identify HAV and HBV all failed to result in the molecular identification of the aetiological agent of NANBH (Deinstag, 1983; Shih et al, 1986). Nevertheless, many things were known about the etiologic agent. In the late 1970s, the transmission of NANBH from humans to chimpanzees was reported by several groups, indicating that it was caused by a transmissible agent (Alter et al, 1978; Bradley et al, 1979; Shimizu et al, 1979; Tabor et al, 1978; Wyke et al, 1979). The studies on NANBH infected chimpanzees showed that the agent was present in most patients with NANBH at a low titre of 10^2-10^3 chimpanzee infectious doses per ml (CID ml^{-1}). The agent was found to

cause the appearance of distinctive, membranous tubules within the hepatocytes of experimentally infected chimpanzees (Jackson et al, 1979; Pfeifer et al, 1980; Shimizu et al, 1979). Buoyant density studies and filtration experiments with the tubule-forming agent suggested that it was a small enveloped virus because it was sensitive to organic solvents, and could pass through filters of 80 nm pore-size (Bradley et al, 1983; Bradley et al, 1985; Feinstone et at, 1983).

Despite all the conventional virological and immunological approaches to identify this new virus, by the mid 1980s the virus responsible for NANBH remained frustratingly elusive. Only with the advent of recombinant DNA technology was it possible to clone the genome of the virus. In 1989, after six years of work on identifying this new virus, Chiron's group finally succeeded in cloning part of the genome of the virus that was termed hepatitis C virus (HCV). A random-primed lambda phage complementary DNA expression library was constructed from a large chimpanzee plasma pool containing relatively high titre of NANBH virus and screened with serum from a patient clinically diagnosed with NANBH. A complementary DNA clone termed 5-1-1 was isolated that was shown to encode an antigen associated specifically with NANBH infections. This clone was proven not to be derived from host DNA but from a single-stranded RNA molecule present in NANBH infections, which has one continuous long open-reading frame (ORF) of about 10,000 nucleotides (Choo et al, 1989). The antigen (C100-3) used in the first-generation enzyme-linked immunosorbent assay (ELISA) was prepared by expressing clone 5-1-1 as a fusion polypeptide with human superoxide dismutase in yeast (Kuo et al, 1989). Most cases of NANBH were found to have developed antibody against this antigen C100-3 (Bruix et al, 1989; Alter et al, 1989; Colombo et al, 1989). These results indicated that HCV was a major causative agent of parenteral non-A, non-

B hepatitis around the world. Studies of NANBH conducted in 1970 showed the high propensity to develop chronic hepatitis, which could progress to liver cirrhosis (Alter et al, 1989; Berman et al, 1979). Soon after the serological assay for the detection of HCV antibody became available, the incidence of the antibody in patients with hepatocellular carcinoma (HCC) was examined and showed that HCV is closely associated with HCC (Bruix et al 1989; Ohkoshi et al 1990; Saito et al 1990; Kiyosawa et al, 1990).

Since the identification of HCV as the major causative agent of non-A, non-B hepatitis, important aspects of molecular virology, epidemiology, as well as natural course of hepatitis C infection have been elucidated. Specific and sensitive assays for the detection of HCV infections have been developed. However, the lack of an efficient cell culture system or small animal model permissive for HCV infection remain as major hurdles to be overcome in the fight against this virus. Due to these hurdles, many important aspects of HCV infection, such as the mechanisms of HCV infection and replication, pathogenesis of HCV-induced liver diseases, and the mechanisms underlying viral persistence are still undefined. In the past decade, although much effort has been applied to define the mechanisms of viral persistence, it has proven to be especially elusive. The challenges to understanding persistent infection have formed the objectives for conducting the present study described in the thesis.

1.2 Epidemiology

The World Health Organization has estimated that as many as 170 million persons worldwide may be infected with HCV, a global population prevalence of approximately 3% (WHO, 1999), however the prevalence rates of HCV infection in the general population are not available in many parts of the world. Therefore, global prevalence rates are estimated based on HCV infection in the blood donor population. The

prevalence rates in healthy blood donors ranges from 0.01-0.1% in the United Kingdom and Scandinavia, 0.2-0.5% in the Americas, Western Europe, Australia, and South Africa to 1-5% in Brazil, Eastern Europe, the Mediterranean, the Mideast, and the Indian subcontinent. Egypt remains an outlier, with the prevalence rates of HCV infection between 17%-26% (Wasley and Alter, 2000).

There is considerable geographic and temporal variation in the incidence and prevalence of HCV infection. Using age-specific prevalence data, three patterns of HCV transmission have been established. The first pattern, found in the United States and Australia, demonstrates that transmission of most cases of HCV occurred in the age group 30-49 years suggesting the relatively recent incidences of transmission. The second pattern, found in Japan and Italy, shows that the prevalence rates are low in older children and young adults, but high in older adults suggesting that HCV transmission was more distant. In countries with the third pattern high prevalence rates are observed in all age groups indicating both distant and ongoing high level HCV transmission. Egypt demonstrates this pattern. It is postulated that injection drug use is the predominant risk factor in the first pattern, whereas unsafe injections and contaminated medical equipment appears to be the risk factor in the second and third patterns (Wasley and Alter, 2000). In Japan and Italy, the incidence of newly acquired HCV infection has declined as a result of improved sterile techniques. Unfortunately, in Egypt, the ongoing use of contaminated and inadequately sterilized medical equipments continues to cause more new HCV infections.

HCV is transmitted primarily by parenteral routes. Major sources of infection include transfusion of unscreened blood and blood products, injection drug use, needle-stick accidents or reuse of contaminated or inadequately sterilized instruments and needles in

health care, religious or other settings (e.g., tattooing, ear or body piercing), chronic haemodialysis, and organ transplantation from HCV positive donors. (WHO, 1999; Pereira et al, 1991). Non-parenteral modes of HCV include vertical and sexual transmission. Mother to infant transmission of HCV has been observed globally, but the risk is probably less than 5% unless the mother is co-infected with human immunodeficiency virus (HIV) (WHO, 1999)). Sexual transmission of HCV is therefore not as common as for HBV. With the broad introduction of anti-HCV screening of blood and blood products in 1990 transfusion associated HCV infection has been virtually eliminated in developed countries (Schreiber et al; 1996) and injection drug use accounts for the largest risk factor for HCV transmission in many countries. Studies of injection drug users revealed that prevalence rate of antibodies to hepatitis C virus in injection drug users are 76.9% in US (Garfein et al, 1996), 60-70% in Victoria, Australia (Crofts et al, 1997), 74.6% in Saudi Arabia (Njoh et al, 1997) and 95.7% in Switzerland (Diamantis et al, 1997).

1.3 Virologic and Molecular Characteristics

1.3.1 Genomic organization and polyprotein processing

HCV has been classified as the sole member of a separate genus called hepacivirus in the family Flaviviridae which includes the classical flaviviruses, such as yellow fever (YFV), dengue and tick-borne encephalitis (TBE) viruses, and animal pestiviruses, such as bovine viral diarrhea virus (BVDV) (Van Regenmortel et al, 2000). All members of this family are small, enveloped viruses containing a positive-sense single-stranded RNA that, in the case of HCV, has a length of ~9600 nucleotides. The genome contains a single open reading frame (ORF) encoding a polyprotein of about 3,000 amino acids.

The ORF is flanked by two highly conserved untranslated regions (UTRs) at both the 5'and 3' termini.

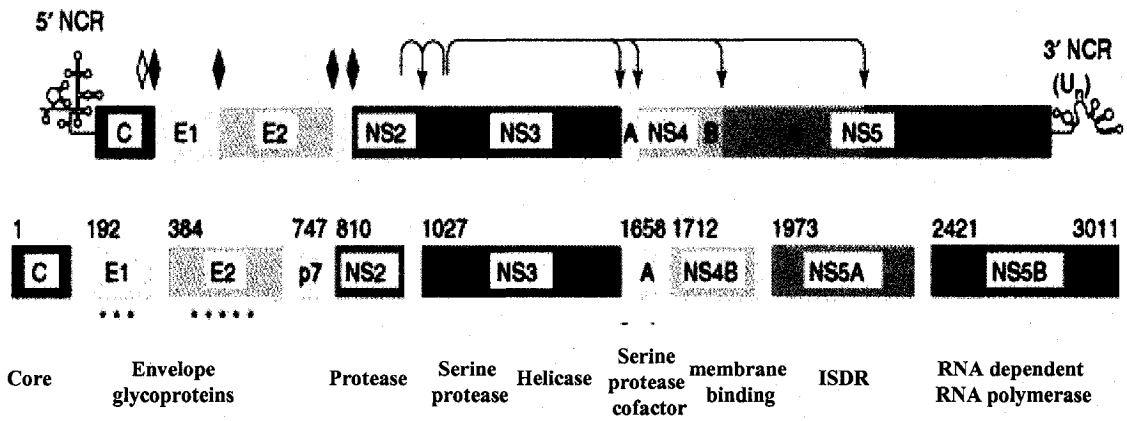
The 5'UTR is 341 nucleotides long (Han et al, 1991) and consists of four highly structured domains numbered I to VI. Domain III contains a pseudoknot, and the ORF translation initiation codon is located in domain IV. It has been found that 5'UTR contains an internal ribosomal entry site (IRES) essential for cap-independent translation of the viral RNA (Fukushi et al, 1994; Reynolds et al, 1996; Rijnbrand et al, 2000). The IRES occupies most of the 5'UTR, including domains II, III, and IV. The specific sequences downstream of the initiator AUG are required for efficient IRES function. In the genome, the first 12 to 30 nucleotides contribute to this IRES function (Honda et al, 1996; Reynolds et al, 1996). A 20-Å resolution 3D model of the mammalian 40S ribosomal subunit bound with either the complete IRES or a fragment lacking domain II was yielded. Intriguingly, binding of the IRES resulted in a significant conformational change in the 40S ribosomal subunit that required domain II, suggesting that the HCV IRES actively manipulates the host translational machinery (Spahn et al, 2001). Sequence heterogeneity between IRES structures has been shown to direct translation with different efficiencies (Laporte et al, 2000). Several cellular proteins, namely, polypyrimidine tract-binding protein and La autoantigen have been found to play functional roles in HCV IRES-mediated translation (Ali et al, 1995; Ali et al, 1997; Ali et al, 2000; Anwar et al, 2000). Because the 5'UTR is one of the most highly conserved regions of the HCV genome, thus most RT-PCR assays for the detection and quantification of HCV RNA use primers specific for the 5' UTR.

The 3'UTR consists of a short variable region of about 40 nucleotides and an internal poly (U)/polypyrimidine tract, followed by a highly conserved 98 nucleotide sequence

which can fold into a very stable stem-and-loop structure (Tanaka et al, 1995; Kolykhalov et al, 1996; Yamada et al, 1996; Blight et al, 1997). This latter tail sequence has been shown to interact with the cellular proteins, including the ubiquitous polypyrimidine tract–binding protein (PTB) and La autoantigen, suggesting that this region is involved in the initiation of replication of the viral genome and probably translation (Tsuchihara et al, 1997; Ito et al, 1999; Spangberg et al, 2001). Studies in chimpanzees using infectious clones of HCV deletion mutants have shown that it is the poly (U/UC) region and the conserved 98 base region but not the variable region that are essential for infectivity. This further confirms that the conserved RNA elements in the 3' UTR play important roles in viral replication (Kolykhalov et al, 2000; Yanagi et al, 1999).

Translation of the HCV ORF results in synthesis of a polyprotein precursor that is co- and post-translationally processed by cellular and viral proteases to yield specific viral gene products outlined in Figure 1. The structural proteins, including the core (C) and the two envelope glycoproteins E1 and E2 are situated in the N-terminal third, with the non- structural proteins in the remaining two thirds of the polyprotein in the order: NS2, NS3, NS4A, NS4B, NS5A and NS5B. The structural proteins are separated from the non-structural proteins by the short membrane peptide p7 of unknown function. The structural proteins are released from the polyprotein precursor by the endoplasmic reticulum signal peptidase. The proteolytic processing of non-structural proteins requires HCV – encoded enzymes: NS2-NS3 protease and NS3 serine protease. NS2-NS3 protease is a zinc-dependent protease and encoded by NS2 and the N-terminal portion of NS3; while NS3 serine protease is located in the N-terminal third of NS3. NS2-NS3 protease appears to be dedicated solely to cleavage at the NS2-NS3 site. The other

Fig.1 HCV genome encoding regions and functions.


5′ NCR
3′ NCR
(Un)
C
E1
E2
NS2
NS3
A
NS4
B
NS5
1
192
384
747
810
1027
1658
1712
1973
2421
3011
C
E1
E2
p7
NS2
NS3
A
NS4B
NS5A
NS5B
Core
Envelope glycoproteins
Protease
Serine protease
Helicase
Serine protease cofactor
membrane binding
ISDR
RNA dependent RNA polymerase

downstream non-structural proteins are cleaved by the NS3 protease associated with its cofactor NS4A (Eckart et al, 1993; Grakoui et al, 1993; Hijikata et al, 1993).

1.3.2 The Viral Proteins

The Core Region:

The core protein, which is thought to form the nucleocapsid in virions, is produced as the most N-terminal component of the polyprotein by host cell signal peptidase. An internal signal sequence, which lies between the core protein and the envelope protein E1, targets the nascent polypeptide to the endoplasmic reticulum (ER) membrane for translocation of E1 into the ER lumen. The immature form of the core protein (191 aa), which contains the E1 signal sequence at its C-terminus, is produced by cleavage of the signal sequence from E1 by the host cell signal peptidase. This signal peptide is subsequently cleaved by signal peptide peptidase, an intramembrane-cleaving protease located in the ER membrane to yield the mature core protein, terminating at or close to amino acid 179 (Mclauchlan et al, 2002). It has been shown that most of the core protein resides in the cytoplasm, where it is bound to ER membranes or located at the surface of lipid droplets, but a small proportion of the core protein is also localized in the nucleus (Yasui et al, 1998; Barba et al, 1997).

The first two-thirds of the core protein constitute a mainly hydrophilic domain that contains several putative nuclear localization signals and a DNA-binding motif. The core protein has highly-conserved sequences among all identified isolates and is thought to bind with viral genomic RNA to form a nucleocapsid because of the basic nature of its amino acid residues, as compared with other flaviviruses (Grakoui et al, 1993; Selby et al, 1993). It has been demonstrated that the core protein binds most efficiently and stably to the 31-nucleotide-long sequence of the loop IIId domain of the 5' UTR, whose

secondary structure is highly conserved among different HCV genotypes (Tanaka et al, 2000).

The core protein has been found to play an important role in the modulation of several cellular events. It has been reported that the core protein of the HCV-1 strain in cooperation with the cellular oncogene H-*ras* can transform primary rat embryo fibroblasts (REFs) to the tumorigenic phenotype (Ray et al, 1996). Another report showed that the core protein in NIH-3T3 cells activated STAT3 through phosphorylation of a critical tyrosine residue, resulting in the tumorigenic transformation of the cells (Yoshida et al, 2002). Several transgenic mouse lines expressing HCV core proteins have been generated in the search for the evidence of involvement of the core protein in the development of hepatocellular carcinoma (HCC). Two independent lines of core transgenic mice developed hepatic steatosis early in life, and HCC at an age of 16 months (Moriya et al, 1998). Recently, another line developed malignant lymphoma at a high frequency (80%) at ages over 20 months, and then developed HCC (Ishikawa et al, 2003). These results indicate a potential role of the core protein in the regulation of cell growth, in the transformation of cells to the tumorigenic phenotype, and in the induction of HCC.

The core protein has been reported to affect apoptosis both negatively and positively. Ray *et al* showed that core protein inhibits cisplatin-mediated apoptosis of human cervical carcinoma (HeLa) cells, c-myc-mediated apoptosis in Chinese hamster ovarian (CHO) cells (Ray et al, 1996), and TNF-α-mediated apoptosis in human breast carcinoma (MCF7) (Ray et al, 1998). Marusawa *et al* also showed that core protein inhibited Fas- and TNF-α-mediated apoptosis via NF-*k*B activation in human hepatoblastoma (HepG2) cells (Marusawa et al, 1999). On the other hand, sensitization

to Fas- and TNF-α-mediated apoptosis by core protein in a mouse fibrosarcoma (BC10ME), HepG2 and HeLa cells was observed by two different groups (Ruggieri et al, 1997; Zhu et al, 1998). The nature of the experiments, cell types and HCV strains used in these studies may contribute to the distinct observations.

Several studies indicate a role for HCV core protein in the transcriptional regulation of cellular and viral promoters. The core protein has been shown to activate the human c-myc promoter, IL-2 promoter, Rous sarcoma virus LTR, simian virus 40 early promoter, and suppress the promoter of HBV, c-fos, p53, p21WafI and human immunodeficiency virus type 1 LTR promoter (Ray et al, 1995; Ray et al, 1997; Ray et al, 1998; Berqvist et al, 2001; Shih et al, 1993). Conflicting results on alteration of activity of transcription factors, however, have also been observed (Shrivastava et al, 1998; Lu et al, 1999; Otsuka et al, 2000).

Cellular proteins associating with HCV core protein have been identified by using the yeast two-hybrid interaction cloning system. The core protein has been found to interact with a variety of cellular proteins, including heterogenous nuclear ribonucleoprotein K, lymphotoxin β, tumor necrosis factor receptor 1, DEAD box RNA helicase, p53, p21, and LZIP protein (Tellinghuisen and Rice, 2002), some of which are involved in transcriptional regulation. As yet, no reports have clarified the mechanism through which the core regulates gene transcription.

The Envelope Region:

The two HCV envelope proteins, E1 and E2, are released from the polyprotein by host signal peptidase cleavages. The cleavage is delayed at the p7/NS2 site and is incomplete at the E2/p7 site, resulting in the production of two forms of E2, fully processed E2 (E2-

A) and E2-p7 (E2-B), differing only in their C-termini (Reed et al, 2000). The function of p7 is currently unknown.

Both E1 and E2 are type I transmembrane (TM) proteins with an N-terminal ectodomains and a short C-terminal hydrophobic transmembrane domain. The ectodomains are targeted to the ER and are modified by N-linked glycosylation at well-conserved sites (5 or 6 in the E1 protein and 9 or 10 in the E2 protein), while the C terminal TM domains are involved in several important protein functions, including membrane anchoring (Cocquerel et al, 2001), endoplasmic reticulum (ER) retention (Cocquerel et al, 1998; Cocquerel et al, 1999), and E1-E2 heterodimer formation (Op De Beeck et al, 2000). E1 and E2 proteins interact either non-covalently to form a heterodimer which is believed to represent the native prebudding complex (Deleersnyder et al, 1997), or covalently through disulfide bonds to produce a complex representing the misfolded aggregates (Dubuisson et al, 1996). Non-covalently heterodimer complexes remain predominantly in the ER where the TM domain of both E1 and E2 possess signals for ER retention (Dubuisson et al, 1996; Duvet et al 1998; Cocquerel et al, 1999; Op De Beeck et al, 2000). Due to their ER retention, the envelope proteins might only be glycosylated by high-mannose type oligosaccharides, (Deleersnyder et al, 1997), but a report suggested complex-type glycans might be involved in the glycosylation (Sato et al, 1993).

The envelope proteins are thought to be the primary mediators of virus attachment and entry. Two molecules have been identified as receptor candidates for HCV: CD81, a member of the tetraspanin family which binds to E2 and the low density lipoprotein (LDL) receptor, which is utilized by virus particles for binding and entry (Agnello et al, 1999; Wunschmann et al, 2000). Recently, with a similar approach, the human

scavenger receptor class B type I (SR-B1), and two closely related membrane-associated C-type mannose-binding lectins, DC-SIGN and L-SIGN have also been proposed as HCV receptors (Scarselli et al, 2002; Lozach et al, 2003).

Comparison of the HCV sequences has shown two hypervariable regions within E2, called HVR1 and HVR2. HVR1 is a 27 amino acid sequence located at the N terminus of E2, and is found in all HCV genotypes. HVR2 is a stretch of 7 amino acids (positions 91-97) and has only been described in the HCV-1b genotype (Hijikata et al, 1991; Weiner et al, 1991; Kato et al, 1992). HVR1 is the target of neutralizing antibodies, and is the only region containing neutralization epitopes so far identified in HCV (Farci et al, 1996). Although HVR1 displays marked sequence variability, some investigators have identified several conserved amino acid residues in HVR1 which define an epitope. Antibodies against this epitope are broadly cross-reactive, as well as having a high capacity to capture HCV variants, suggesting the recognition of a key humoral HCV epitope (Shang et al, 1999; Li et al, 2001).

E2 protein was recently reported to bind to and inhibit double stranded RNA-dependent protein kinase (PKR) through a sequence homologous to the phosphorylation sites of PKR and its target, eukaryotic translation initiation factor eIF 2α, designated PKR-eIF2α phosphorylation homology domain: PePHD. The virally induced antiviral cytokine, type I IFN, acts in part through the double stranded RNA dependent PKR that inhibits protein synthesis by phosphorylation of eIF2α. The interaction of PePHD with PKR inhibited the kinase activity of PKR by preventing its autophosphorylation and blocked its inhibitory effect on protein synthesis through the prevention of the phosphorylation of eIF2α. Therefore, this interaction may provide one of the mechanisms underlying the viral resistance to IFN (Taylor et al, 1999; Taylor et al, 2000).

The NS2 region:

As described in the previous section, the non-structural proteins of the virus, NS2 to NS5B, are located at the C-terminal end of the polyprotein and are released by two virus encoded enzymes: NS2-NS3 zinc-dependent protease and NS3 serine protease. These proteins are not expected to be constituents of the virus particle.

NS2 is released from its polyprotein precursor by two proteolytic cleavages: cellular signal peptidase-mediated cleavage at its N-terminus and NS2-NS3 protease-mediated cleavage at its C-terminus. NS2 is a transmembrane protein with the C-terminus located in the ER lumen and the N-terminus in the cytosol (Santolini et al, 1995). This protein may not be critical for replication of the viral RNA (Lohmann et al, 1999; Blight et al, 2000). The only known function of NS2 is its participation in proteolytic cleavage at the NS2-NS3 junction of the polyprotein, which releases NS2 from the downstream portion of the precursor polyprotein, a process that occurs rapidly after translation and by a conformation-dependent, autocatalytic mechanism (Hijikata et al, 1993). NS2-NS3 protease comprises most of the NS2 region and part of the NS3 domain, appears to be Zinc dependent, and can lose its protease activity by self-cleavage between NS2 and NS3. The biochemical nature of this viral enzyme remains elusive primarily because of the inherent difficulty of expressing and purifying an active protein and the lack of an in *vitro* assay.

The NS3-NS4 Region:

NS3 is a multifunctional protein with a serine protease domain located in the N-terminal one-third and an NTPase/RNA helicase domain in the C-terminal two-thirds. The proteolytic processing by the serine protease involves two sets of independent events: first, cleavage of NS2/NS3 site, in conjunction with NS2, and second, cleavage

at the 3/4A, 4A/4B, 4B/5A, and 5A/5B sites to release the remaining downstream NS proteins (Major et al, 2001). Three highly conserved residues, His-1083, Asp-1107 and Ser-1165 that form the enzyme catalytic triad are found in the NS3 protease domain (Chambers et al, 1990). The cleavage between NS3 and NS4A has been shown only in *cis* and is the first event in the processing cascade, followed by cleavages at the other sites, which can occur in *trans* (Bartenschlager et al, 1993; Bartenschlager et al, 1994).

NS4A, a 54-residue viral protein, functions as a cofactor in NS3 protease activity. It is indispensable for cleavage at the NS3/NS4A and NS4B-NS5A sites and accelerates the rate of cleavage at the NS5A-NS5B (Failla et al, 1994; Tanji et al, 1995). In addition, NS4A operates to stabilize NS3 and targets it to the membranes of the ER (Tanji et al, 1995). The NS3 serine proteinase domain interacts with NS4A, creating a stable complex. A central hydrophobic domain of NS4A and the N-terminal portion of the NS3 serine protease have been found to be essential for both NS3-NS4A complex formation and NS4A-modulated proteolytic activity (Failla et al, 1995; Lin et al, 1995; Tanji et al, 1995). NS4A also associates with and directs hyperphosphorylation of the NS5A protein.

The crystal structures of NS3 serine proteinase, both free and complexed with NS4A, have been determined (Love et al, 1996; Kim et al, 1996). Analysis of the three-dimensional structures of NS3 revealed that HCV NS3 protease domain adopts a chymotrypsin-like fold with two six-stranded β-barrel subdomains of identical topology. The catalytic triad is formed by residues from the same loops of the two β-barrels: His 57 and Asp81 in the N-terminal β-barrel and Ser139 in the C-terminal β-barrel and it is located in a groove between the two barrels. The substrate-binding site that is consistent with the cleavage specificity is also located in the groove. The central part of NS4A

interacts with N-terminal portion of NS3 protease through the formation of a β-strand inserted into the N-terminal β-barrel of NS3. Therefore, NS4A is considered an integral part of the NS3 protease structure. NS3 has a structural tetrahedral zinc-binding site formed by residues Cys-1123, Cys-1125, Cys-1171 and His-1175. This zinc-binding site is more conserved than the catalytic residues and is distal to the active site (Kim et al, 1996).

The RNA helicase and NTPase domain comprises the 465 C-terminal amino acids of NS3 (Kim et al, 1995 Suzich et al, 1993). The helicase unwinds double stranded RNA (dsRNA) and RNA/DNA heteroduplexes in the 3' to 5' direction using any NTP or dNTP as an energy source. Mutations of the conserved residues in the ATPase/helicase motifs severely impair both functions (Kim et al, 1997). The crystal structure of NS3 helicase/NTPase domain has been elucidated, and is representative of superfamily-2 helicases (Kim et al, 1998; Kwong et al, 2000).

In addition to its enzymatic activities, NS3 has been shown to interact with several cellular components. NS3 was reported to transform NIH 3T3 mouse fibroblasts and induce the development of tumours in nude mice (Sakamuro et al, 1995); suppress actinomycin D-induced apoptosis (Fujita et al, 1996), and inhibit phosphorylation mediated by cAMP-dependent protein kinase (Borowski et al, 1997).

NS4B is a very hydrophobic protein that associates with the ER membrane. The function of this protein remains to be established.

The NS5 region:

The NS5 region encodes two viral proteins, NS5A and NS5B. Two phosphoproteins, p56 (56kDa) and p58 (58kDa), are produced from the HCV NS5A region (Kaneko et al, 1994; Koch et al, 1999; Neddermann et al, 1999). P56 is a basally phosphorylated form

of NS5, whereas p58 represented a hyperphosphorylated form of p56. Both proteins are phosphorylated at serine residues, and the kinase(s) responsible for the phosphorylation remains to be determined. Hyperphosphorylation of NS5A requires its expression as part of a continuous polyprotein with NS3, NS4A, and NS4B, and requires the protease activity of NS3. Two regions have been reported to be important for basal phosphorylation of NS5A, located around the center (amino acids 2200 to 2250) and C terminus (amino acids 2350 to 2419) (Tanji et al, 1995). Ser-2231 and Ser-2194 have been identified as the major phosphorylation sites of NS5A in an HCV1a and 1b isolate respectively (Reed et al, 1999; Katze et al, 2000), whereas Ser-2197, Ser-2201, Ser-2204 are shown to be required for hyperphosphorylation of HCV 1b NS5A (Tanji et al, 1995). Recently, an amino-terminal amphipathic alpha helix has been identified in the NS5 protein, and was shown to be both necessary and sufficient to mediated association of NS5A with the ER membrane (Brass et al, 2002).

Although the function of NS5A in the viral replication is unknown, adaptive mutations have been found to cluster in the NS5 region in the replicon system, suggesting NS5A is involved in modulation of the viral replication cycle (Blight et al, 2000). Taken together with the requirement of NS3, NS4A and NS4B for NS5A hyperphosphorylation, it is clear that NS5A is essential for the formation of HCV replication complex.

NS5A has attracted considerable interest because of its potential role in modulation of the response to interferon (IFN) alpha therapy. NS5A was first linked to the IFN response in molecular epidemiology studies by Enomoto and colleagues. By comparing the full-length sequences of IFN-α-sensitive and IFN-α-resistant Japanese HCV1b isolates they found that IFN-α-sensitive isolates contained recurring multiple mutations ($\geq$4) within a discrete region of 40 amino acids in the carboxyl half of NS5 (amino acid

2209-2248), suggesting that this NS5 region, termed the IFN sensitivity-determining region (ISDR), may play a role in conferring IFN resistance (Enomoto et al, 1995; Enomoto et al, 1996). However, studies in Europe and in the North America failed to find a correlation between the ISDR sequence and its ability to inhibit IFN activity (Squadrito et al, 1997; Zeuzem et al, 1997; Duverlie et al, 1998; Nousbaum et al, 2000). The genetic difference between HCV isolates in Japan and in Europe or the United States may account for the discrepant results (Nakano et al, 1999). Nevertheless, new statistical analysis of a database of 675 published individual ISDR sequences strongly supported the correlation of the NS5A ISDR with the IFN response (Witherell et al, 2001).

The HCV NS5A protein binds to and inhibits PKR (Gale et al, 1997), an IFN-induced gene product that is activated by binding to dsRNA. PKR phosphorylates the translation initiation factor eIF-2α, inhibiting protein synthesis. NS5A forms heterodimers with PKR and inhibits PKR-homodimer formation, resulting in the inhibition of both PKR autophosphorylation and phosphorylation of eIF2α. The ability of NS5A to bind PKR requires the ISDR and an additional 26 residues downstream. Mutations in the ISDR can disrupt the NS5A-PKR interaction and possibly render HCV isolates containing such mutations susceptible to IFN therapy (Gale et al, 1998). However, several conflicting results have been presented. Podevin *et al.* (Podevin et al, 2001) and Ezelle *et al.* (Ezelle et al, 2001) failed to observe any effect of NS5A on the activity of PKR in either Huh7 or HeLa cells, although Podevin *et al.* (Podevin et al, 2001) did show that all isolates of NS5A inhibited IFN activity. These studies suggest that NS5A may also counteract the antiviral effects of IFN via other mechanisms independent of its interaction with PKR. It

has been observed that NS5A may be capable of inducing the transcription of interleukin (IL)-8, leading to the inhibition of the antiviral effects of IFN (Polyak et al, 2001).

It has been shown that NS5A protein, with an amino-terminal deletion and fused with the DNA binding domain of yeast transcriptional activator GAL4, activates the transcription in yeast and mammalian cells, suggesting that it is a potential transcriptional regulator, and that amino acids 2135-2248 and amino acids 2207-2325 seem to be crucial for transcriptional transactivation properties of NS5A. These results suggest that NS5A is a potential transcriptional regulator (Tanimoto et al, 1997; Kato et al, 1997). Modulation of p53-dependent p21/Waf-1 gene expression has been attributed to the NS5A protein, although contradictory reports exist as to whether NS5A is an activator or inhibitor of this pathway (Majumder et al, 2001; Ghosh et al, 2000; Arima et al, 2001). In addition, NS5A has been shown to interact with the growth-factor-receptor-bound protein 2 (Grb2) adaptor proteins and block mitogenic signalling (Tan et al, 1999). By contrast, another report showed that NS5A induced anchorage independent growth in murine fibroblasts (NIH3T3) and tumour formation in nude mice. This result also indicated the potential role of NS5A in oncogenesis (Ghosh et al, 1999). NS5A were also found to interact with karyopherin $\beta 3$ (Chung et al, 2000) and hVAP-33 (human vesicle-associated protein) (Tu et al, 1999). The biological significance of the interaction between NS5A and these cellular proteins remains to be elucidated.

NS5B was identified as the RNA-dependent RNA polymerase (RdRp). RdRp activity has been shown for recombinant NS5B protein expressed in insect cells using a baculovirus vector (Behrens et al, 1996; Lohmann et al, 1997), or in *E. coli* (Yuan et al, 1997; Al et al, 1998). Four amino acid sequence motifs have been found to be crucial for RdRp, designated A (amino acids 2640 to 2645), B (amino acids 2702 to 2711), C

(amino acids 2737 to 2739), D (amino acids 2762 to 2766) (Lohmann et al, 1997). These four motifs may be involved in nucleotide binding and catalysis (A), template and /or primer positioning (B), and NTP binding and catalysis (C and D). The recombinant NS5B protein prefers a primer-dependent initiation of RNA synthesis, either by elongation of a primer hybridized to an RNA homopolymer or via a copy-back mechanism when using heteropolymeric templates (Behrens et al, 1996; Lohmann et al, 1997). A full length of NS5B expressed in E. coli was found to be capable of synthesizing the full-length viral RNA in a primer-independent manner (Oh et al, 1999). As mentioned in the last section, hVAP-33 has been shown to interact with NS5A. Interestingly, this interaction has also been observed for NS5B. NS5A and NS5B bind to different domains of hVAP-33: NS5A binds to the C-terminus, whereas NS5B binds to the N-terminus of hVAP-33 (Tu et al, 1999).

The crystal structure of NS5B has been determined and reveals a catalytic domain followed by a C-terminal extension that connects to the TM region via the active-site groove.

1.3.3 Virion Morphology

The morphology of the virion remained unclear until 1994. Kaito's group described HCV particles in plasma of blood donors with high HCV-RNA titre by indirect immunogold electron microscopy using anti-HCV envelope 1 (E1) antibody, and demonstrated that HCV virions are 55-65 nm spherical particles with fine surface spike-like projections (Kaito et al, 1994). The nucleocapsid of the particle was found to be 50 nm in diameter, icosahedral in structure, and surrounded by an envelope covered with surface projections (Li et al, 1995). The morphological features of the HCV particle

were shown to be consistent with the characteristics of flaviviruses and pestiviruses of the *Flaviviridae* family.

1.3.4 Genetic Diversity

The first complete genome sequence was determined in 1991 (Choo et al, 1991). As additional genome sequences from isolates from different geographical regions were elucidated and compared, it was evident that HCV exhibited substantial nucleotide sequence heterogeneity throughout the viral genome (Bukh et al, 1995). Like many RNA viruses, the existence of significant genetic diversity among HCV genomes is assumed to be the result of the accumulation of mutations during viral replication. This high mutation rate, which is characteristic of RNA viruses, can be attributed to an error-prone RNA-dependent RNA polymerase that lacks proofreading activity. Because HCV exists as quasispecies in vivo, it is difficult to determine the actual mutation rate of HCV genome. However, it is possible to know the rate of fixation of mutations that is determined by the accumulated number of mutations in the viral genome per unit of time during the infection. The average rate of fixation of mutations in the HCV genome has been estimated to be 1.44 to 1.92×10^{-3} substitutions per genomic site per year, by comparing the consensus sequences of the isolates obtained from individuals at intervals of 8 to 13 years, in both humans and chimpanzees (Okamoto et al, 1992; Ogata et al, 1991). The rate of fixation of mutations is not evenly distributed throughout the genome, with the highest rate in HVR-1 of the E2 gene (Bukh et al, 1995).

In order to impose some structure on the global genetic diversity of HCV, a classification scheme based on molecular phylogenetic analysis of viral sequences has been developed. In this scheme, all current HCV complete genome sequences are classified into six distinct genotypes (labelled 1-6) and more than 50 subtypes (labelled

alphabetically in their order of discovery). Complete genomes from different types differ at approximately 30-35% of nucleotide sites (Simmonds, 2000), with more variability concentrated in regions such as the envelope proteins, whereas the 5' UTR, core, and NS3 are highly conserved. Each of the six major genetic groups of HCV contains a series of more closely related subtypes that typically differ from each other by 20-25% in nucleotide sequences (Simmonds, 2000).

HCV genotypes and subtypes display clear differences in their geographic distribution. HCV subtypes 1a, 1b, 2a, 2b and 3a are distributed very widely and account for the vast majority of infections in North and South America, Europe, and Japan. Genotype 4 is prevalent in North Africa and the Middle East, and genotypes 5 and 6 are found in South Africa and South East Asia respectively (Simmonds et al, 1993, Abdulkarim et al, 1998; Chamberlain et al, 1997; Cha et al, 1992). Infections in Egypt are caused predominantly by HCV Subtype 4a.

Repeated studies have shown that patients with HCV subtypes 1a, 1b and 4a infection have lower response to anti-viral therapy than those infected with type 2 and 3 strains, suggesting that genetic variation among subtypes can generate significant differences in clinical outcome (Bell et al, 1997; Mondelli et al, 1999). In addition, HCV genotype has been found to be an important factor in the progression of the liver disease. Subtype 1b infections are more likely to cause liver disease and have been associated with more advanced liver diseases, cirrhosis, and hepatocellular carcinoma than other subtype infections (Brechot et al, 1997; Pozzato et al, 1994; Zein et al, 1996)

Significant genetic heterogeneity was found among isolates from an HCV infected individual, indicating the existence of HCV genomes as a quasispecies. Analogous to other RNA viruses, HCV circulates in an infected individual as a population of closely

related, yet heterogeneous, sequences: the quasispecies (Martell et al, 1992; Eigen et al, 1996). The viral population is composed of a dominant sequence, termed the master or consensus sequence, and a large spectrum of mutants. The consensus sequence is the sequence resulting from combining the most common nucleotides at each position of the genome and is obtained by direct sequencing of PCR products or by alignment of cloned sequences. It is assumed that the master sequence is the most fit or best adapted sequence with respect to viral replication in the particular host at that given time point.

Developement of quasispecies implies a significant adaptation advantage for a virus because the simultaneous existence of a large and diverse population of genomes allows for the rapid selection of the mutants with better fitness for any new environmental condition. On the other hand, a quasispecies will remain in stable equilibrium with little evolution of its consensus or master sequences as conditions are unchanged (Martell et al, 1992). Development of quasispecies has many important biological implications, including the development of escape mutants to cellular or humoral immunity, the generation of "defective" viral particles, variable cell tropism, and the development of drug resistance (Marrone and Sallie, 1996).

The relation of HCV quasispecies evolution with the progression of liver disease is more clear-cut. A study by Farci et al showed that the outcome of acute hepatitis C could be predicted by the evolution of the viral quasispecies. Acute resolving hepatitis was associated with relative evolutionary stasis of the quasispecies, whereas progressing hepatitis correlated with genetic evolution of HCV. In all cases, the genetic diversity was significantly higher in HVR1 than in other regions of the E1 and E2 genes, suggesting that this region is under selective pressure by the host immune system (Farci et al, 2000).

Because of lack of a vaccine, effective treatment for hepatitis C is especially important. Interferon (IFN), alone or in combination with ribivirin, is the only drug with proven efficacy against HCV, but only 40% of treated patients have a sustained response (Farci et al, 2002). Previous work has shown a relationship between the complexity of the quasispecies within HVR1 and the response to IFN therapy (Farci et al, 1997). There is evidence that a greater degree of quasispecies complexity is associated with a lack of response to IFN. A recent study found a significant decrease in viral quasispecies diversity in sustained responders compared with nonresponders during the first two weeks of the treatment, although the genetic diversity before treatment did not correlate with treatment outcome (Farci et al, 2002). Therefore, low genetic diversity of HCV quasispecies during the early stages of IFN treatment is usually found to be predictive of therapy success.

Genetic diversity is of course also affected by host factors, most importantly the strength of the humoral and cellular immune responses. Over the course of an infection, immune pressure could increase diversity by repeatedly selecting for escape mutants, and individuals that are immuno-suppressed have almost homogenenous viral populations (Lawal et al, 1997; Ni et al, 1999).

1.4 Mechanisms of persistent infections

More than 85% of individuals infected by hepatitis C virus become chronically infected, but the mechanism(s) underlying persistent HCV infection remains elusive. The high rate of chronic HCV infections suggests that the virus, which is non-cytopathic, has evolved one or more mechanisms aimed at evading host immunity.

Evasion of the immune response by quasispecies variation

HCV is a rapidly replicating virus, with approximately 10^{10}-10^{12} new virions produced daily. This high replication rate, in combination with the absence of any HCV polymerase proofreading ability, leads to the emergence of viral quasispecies, which may allow the virus to circumvent the immune response, leading to persistent infection.

Variation of the neutralizing epitopes

Several studies have shown that during HCV infection, the HVR1 sequence of E2 becomes progressively heterogeneous, and HVR1 mutated less frequently in immunocompromised patients than in immunocompetent patients, suggesting that HVR1 is a target of anti-HCV antibody attack and that mutations in the HVR1 are the result of selective pressure (Farci et al, 1997; Kumar et al, 1994; Booth et al, 1998). Kato *et al* reported that B-cell epitope, found in HVR1 shifted during HCV infection, and amino acid substitutions in B epitopes led to the escape from recognition by anti-HVR1 antibody (Kato et al, 1993; Kato et al, 1994). Using an experimental chimpanzee model, HVR1 has been shown to be a critical neutralization domain of HCV (Farci et al, 1996; Esumi et al, 1999). Farci *et al* (Farci et al, 2000) reported that the pattern of quasispecies (in HVR1) during acute phase predicts the outcome of infection. If the quasispecies pattern is limited, infection is circumvented and the virus is eliminated. However, if the quasispecies pattern continues to evolve, persistent infection results. These observations suggest that selection of variants that are not neutralized by anti-HVR1 antibody may contribute to the mechanism responsible for persistent HCV infection.

Variation of CTL epitopes

As an alternative suggestion, it has been hypothesized that immune selection of CTL escape variants may be responsible for HCV persistence (Chang et al, 1997; Weiner et

al, 1995). In keeping with this notion, a HCV variant that escaped CTL recognition has been reported in an experimentally infected chimpanzee who subsequently developed chronic HCV infection (Weiner et al, 1995). Mutations of key epitopes targeted by CD^{8+} CTL have been observed in the HCV quasispecies from chimpanzees and humans (Weiner et al, 1995; Chang et al, 1997; Tsai et al, 1998; Erickson et al, 2001) with chronic hepatitis C, and many of these mutations inhibited CD^{8+} CTL recognition. Some variants with these mutations have been reported to function as T cell receptor antagonists and inhibit CTL activity (Tsai et al, 1998). In these cases, selection of CTL escape variants which have multiple escape mutations in CTL epitopes during early stages of infection may lead to chronic infection, whereas a narrow spectrum of CTL epitopes correlates with virus clearance.

Implication of HCV proteins in the modulation of cellular immune functions

Given the fact that there are multiple antigenic sites to be recognized by T and B cells, the loss of some epitopes is unlikely to confer a significant survival advantage to the variant virus, and the high rate of HCV persistent infection can not be fully attributed to the selection of antibodies and HCV-specific CTL escape mutants. One possibility is that viral proteins may facilitate immune escape and the establishment of persistent infection through the interaction with host proteins.

The tumor necrosis factor (TNF) receptor family is involved in the immune system and particularly in the control of apoptosis. The core protein has been reported to suppress the host immune responses through interaction between intracellular core and TNFR or Fas, and this interaction has been shown to induce apoptosis in both hepatocytes and lymphocytes (Matsumoto et al, 1997). C1q, the natural ligand of complement receptor gC1qR, is a part of the C1 complex and plays a critical role in early immune responses.

Kittlesen *et al* has shown that the extracellular core binds to a C1qR on T cells, suppressing T cells activity (Kittlesen et al, 2000).

Natural killer (NK) cells are a critical component of innate immunity and play a pivotal role in the elimination of many viruses. The HCV E2 protein binds CD81 expressed on the surface of host cells, and this was recently linked to the suppression of NK activity. It has been shown that the ligation of CD81 on NK cells by immobilized E2 inhibits non-MHC-restricted cytotoxicity mediated by NK cells, IL-2-induced proliferation, as well as IL-2-, IL-12-, or IL-15-induced IFN-γ production (Tseng et al, 2002). Additionally, cross-linking of CD81 specifically inhibits CD16-mediated activation of extracellular signal-regulated kinase (ERK) and mitogen-activated protein kinase (MAPK) in NK cells.

Interferon is an antiviral cytokine induced by viral infection. As previously described, at least two HCV proteins, E2 and NS5A, have been shown to inhibit the IFN response. The role and the mechanisms of IFN resistance in persistent HCV infections are not fully understood.

From the above review on HCV, we notice that the underlying mechanism for HCV persistence is still unknown. In this study, we focus our discussion on how molecular mimicry of immunoglobulin by E2 can result in persistent HCV infection. Therefore, it is necessary to give a short review on the structure of immunoglobulin.

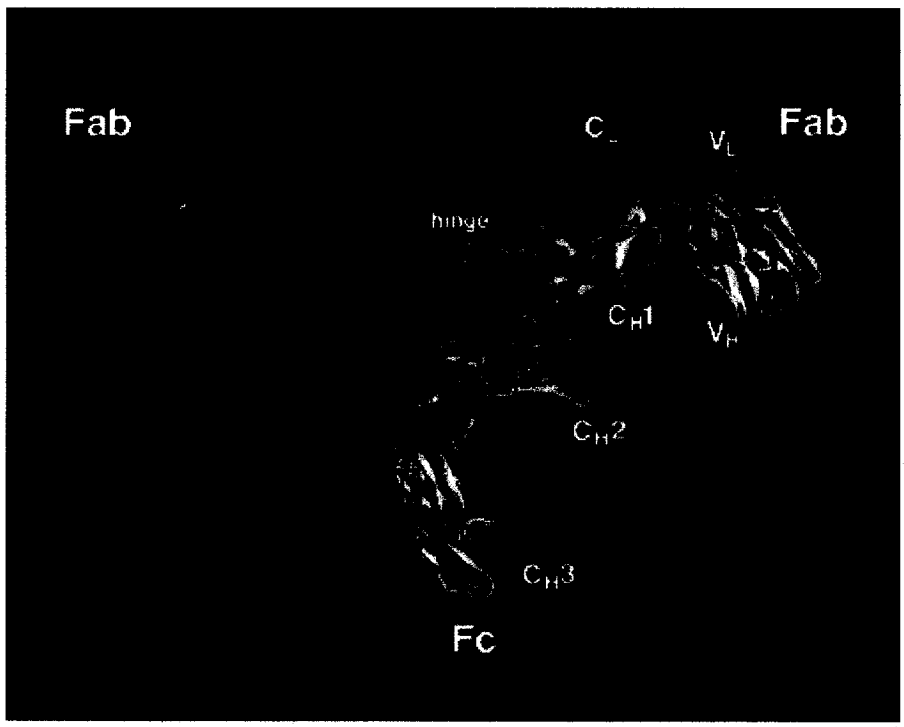
1.5 Structure of immunoglobulin

The analysis of the structure of antibody molecules really began in 1959 with two discoveries which, for the first time, revealed that the molecule could be separated into analyzable parts suitable for further study. In England, Porter found that proteolytic treatment with the enzyme papain split the immunoglobulin molecule (molecular weight

150,000 daltons) into three fragments of about equal size. Two of these were found to retain the antibody's ability to bind antigen specifically, designated Fab (fragment antigen binding) fragment; the third fragment could be crystallized out of solution, called Fc (fragment crystallizable).

At about the same time, Edelman in the United States discovered that treatment with mercaptoethanol (a reagent that breaks S-S bonds) split the immunoglobulin molecule into four chains: two identical chains with a molecular weight of about 53,000 daltons each, designated heavy (*H*) chain, and two others of about 22,000 daltons each, called light (*L*) chain. Based on these results, the structure of immunoglobulin molecules, as depicted in Figure 2, was proposed. Thus, the general Y shaped configuration of an immunoglobulin corresponds to two identical Fab arms and Fc region. Each of the Fab arms consists of the light chain and the N-terminal half of the heavy chain, while the Fc region comprises the C-terminal half of the two heavy chains. Human immunoglobulins have been shown to consist of five different classes (isotypes), IgM, IgD, IgG, IgA and IgE, which differ in the structure of their H chains (termed mu, delta, gamma, alpha and epsilon, respectively). The two light chain types, kappa and lambda, are common to all five classes. The immunoglobulin molecules are assemblies of separate domains, each centered on a disulfide bond. The N-terminal domain on L and H chains is highly variable, in terms of amino acid sequence, from one immunoglobulin to the next, and it is therefore designated the variable domain (V_L or V_H). The other domains on both chains are constant in amino acid sequence and are designated constant domains. There is one constant (C) domain in the kappa and lambda light chain, three (CH1, CH2 and CH3) in the delta, gamma, or alpha heavy chains, and four (CH1-CH4) in the mu and epsilon heavy chains. The immunoglobulin domain (variable or constant) folds into two

Fig. 2 Structure of an immunoglobulin molecule. Modified from *Nature* 1992, *360*, 369-372


Fab
C_L
V_L
Fab
hinge
C_H1
V_H
C_H2
C_H3
Fc

layers of antiparallel β-sheet arranged in a β-barrel, which consists of nine and seven β-strands for variable and constant domain respectively. Each variable domain is composed of four "framework region" (FR) supporting three hypervariable loops or complementary determining regions (CDRs), which are spatially close to each other and constitute the antigen binding site. In this study we have observed a structural relationship between variable domains of immunoglobulins to the E2 region of HCV.

1.6 Research rationales, hypotheses and objectives

It is noted from a thorough review of the relevant literature that the mechanism by which HCV persists in the majority of infected individuals has been studied for some time, but it remains elusive. Although the currently accepted hypothesis is that amino acid substitutions in HVR1 leads to the emergence of quasispecies, some of which are antibody escape mutants that are not recognized by the immune response and persist after seroconversion (Kato, 2001; Pavio et al, 2003), the important question as to why the virus epitopes within HVR1 cannot be subsequently recognized by the immune system is unknown. To fully elucidate the mechanism underlying the persistent HCV infection, a study aiming at identifying the mechanism by which the virus epitopes within HVR1 cannot be recognized by the immune system becomes necessary. This has formed the rational for the present study.

Molecular mimicry is defined as similar structures shared by molecules from dissimilar genes or by their protein products. Either the molecules' linear amino acid sequences or their conformational structure may be shared, even though their origins are as separate as, for example, a virus and a normal host-self determinant. Molecular mimicry, where viruses encode proteins that are homologs of host defense proteins and immunomodulators, has been shown to be an important strategy used by viruses

(especially DNA viruses) to subvert the host immune responses and become persistent (Ploegh, 1998; Seet et al, 2003; Vossen et al, 2002). HCV has been found to employ molecular mimicry to resist type I IFN, as previously described. We reason that other instances of molecular mimicry could also be contributing to HCV persistent infection and may function to avoid immune detection. To address this question, we searched for homologous sequences to the HCV polyprotein among all proteins in GenBank. We found that HCV encodes a sequence in envelope region 2 (E2) that is highly homologous to human imunoglobulin (Ig). We hypothesized that molecular mimicry of immunoglobulin by HCV E2 was an important strategy for immune evasion and the establishment of persistent infection. To test this hypothesis, we compared V-genes of various Ig types and E2 sequences of HCV variants that arise in the course of primary infection by using bioinformatic and evolutionary approaches. We found that a high level of sequence similarity existed between E2 and Ig, and the degree of similarity of HVR1 or its epitopes with Ig types was directly related to viral escape and persistence in a host specific manner. To further confirm our hypothesis, recombinant E2 proteins were expressed and reacted with anti-human-IgG Fab fragment. This approach provided evidence of structural similarity of E2 to IgG. Taken together, these data indicated HCV employ a strategy of mimicry of immunoglobulin to evade the host immune attack and become persistent.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Patient Samples

Two patients with primary HCV infections transmitted from known virus sources were retrospectively included in the study. Patient A is a 43-year-old male who had acquired an acute nosocomial HCV 1a infection from chronically infected patient S. (Larke et al, 2002). Serum samples were collected before and soon after seroconversion as well as four weeks after seroconversion (A1.2m, A1.6m and A2.4m at days 36, 48 and 72 post infection (PI), respectively). From day 74 pi, patient A received a six month of IFN and ribavirin therapy. Serum samples were also obtained 1, 2 and 7 months after 6 months of IFN therapy. Patient B was a lung transplant patient who had cleared a previous HCV infection and subsequently acquired HCV 3a from an organ donor (patient D). Serial samples were taken from patient B on days 8, 50, 78, 109 and 171 after transplant/infection. All samples were stored at -70° C until used. Anti-HCV antibody was detected by using an EIA (ortho HCV version 3.0 ELISA, Ortho Diagnostic Systems Inc., Raritan, NJ or Abbott AxSym EIA) and confirmed by using RIBA (Chiron RIBA HCV 3.0 SIA, CHIRON, Emeryville, CA). Sequential HVR1 sequences from several follow up studies published were also used for this study.

2.2 RNA Extraction

Viral RNA was extracted from 140 μl patient serum using the QIAamp viral RNA purification kit (Qiagen, Germany). The procedure and conditions were the same as those described in the manufacturer's instructions. The RNA pellet thus obtained was resuspended in 50 μl of preheated (80°C) RNase-free water.

2.3 Reverse Transcription and PCR

cDNA synthesis was performed by using random hexamer primer (pd(N)6) and the first-strand cDNA synthesis kit as per manufacturer's instructions (Amersham Pharmacia Biotech, Piscataway, USA). Briefly, reverse transcription was carried out in 15 μl of the reaction containing 8 μl of RNA, 0.2 μg of $(N)_6$ primer (random hexadeoxynucleotides), 25 U volume of FPLC*pure*^{TM} murine reverse transcriptase, 45mM Tris (pH8.3), 68 mM KCl, 15 mM DTT, 9mM $MgCl_2$, 0.08mg/ml BSA and 1.8 mM each dNTP. The RNA was heated at 65°C for 10 min, and quenched briefly on ice before the reaction mixture was added for 1 hr of incubation at 37℃. The E2 region, including HVR1 was amplified by nested PCR using 15 μl of the entire cDNA reaction in 50 μl of a reaction mixture containing 2.5 U Taq DNA polymerase (Qiagen, Chatsworth, CA) and 0.5 μg of each primer. Second PCR was performed by using Taq PCR Core Kit (Qiagen, Chatsworth, CA) as per manufacturer's instructions. Briefly, one microliter of first PCR product was amplified in a second PCR step in a volume of 100 μl containing 2.5 U Taq DNA polymerase, 1x PCR buffer, 1x Q solution, 200 μM of each dNTP, 2.0 mM MgCl and 0.5 μg of each forward and reverse primer. The primers used in HVR1/E2 amplification are as follows: Patient A (HCV 1a): outer sense: (1290-1313) 5'-CGCATGGCATGGGA TATGATGATG-3', outer anti-sense: (1901-1882) 5'-CGCGCCCGACCTGTCGGTCG -3', inner sense: (1293-1316) 5'-ATGATGATGAAC TGGTCCCCTACG-3', inner anti-sense: (1891-1872) 5'-CCCACCACCACGGGGCTG GG-3'. Patient B (HCV 3a): outer sense: 5'-TGGCGGGCCTAGCCTATT-3', outer anti-sense: 5'-ACGATGTTTTGGTGGAGGTG- 3', inner sense: 5'-ATGTTTTCAGG GGTCGATGCC-3', inner anti-sense: 5'-GGTGGAGGTGTATTAGACCTGT-3'. The same thermocycling conditions were used for both PCR rounds, which included 40

cycles of 94°C for 30 sec, 45°C for 30 sec, and 72°C for 1 min, and one cycle at 72°C for 7 min.

The PCR products obtained were analyzed on a 1.5 % agarose gel and visualized in the presence of ethidium bromide under ultraviolet transillumination.

PCR products were purified by using a QIAquick PCR purification kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. In this study, one hundred microliter of PCR product was used to be purified and fifty microliter of ddH_2O were added to elute DNA.

2.4 Cloning of the PCR Products into pCR2.1 Vector

The purified PCR products were ligated into the pCR2.1 vector using the TA cloning kit (Invitrogen, Carlsbad, CA). Ligations were performed with 1 μl (50ng) of PCR insert, 1 μl (10ng) of pCR2.1 vector, 1.2M NaCl and 0.06 M MgCl, and the mixture was incubated at room temperature for 10 min.

For the generation of recombinant E. coli clones, two microliters of ligation products were used to transform 50 μl of competent E. coli cells using TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA). Briefly, two microliters of ligation products were added into 50 μl of competent cells and mixed gently by tapping. The cell suspension was incubated on ice for 30 min, heat-shocked for 30 sec in the 42° C water bath without shaking and immediately incubated on ice for 2 min. Two hundred and fifty microlitres of S.O.C medium was added to the cells and incubated at 37°C with constant shaking at 225 rpm for 1 hour. A volume of 40 μl of the transformed cells was then plated on LB agar plates containing 100 μg/ml ampicillin and 1.6 mg of X-Gal, and incubated overnight at 37°C.

2.5 PCR Amplification of Cloned Samples

Colonies were screened for E2-containing plasmids by PCR amplification of colonies (Colony PCR) using the Taq Core Kit (QIAGEN). Twenty to forty positive recombinant clones (20 to 40 white colonies) per sample were randomly picked from the LB agar plates containing 100 μg/ml ampicillin, and each was added to 20 μl of PCR reaction mixture containing 1x PCR buffer, 125 μM of each dNTP, 0.04 μg of M13 Forward primer (5'-GTAAAACGACGGCCAG-3'), 0.04 μg of M13 Reverse primer (5'-CAGGAAACAGCTATGAC-3'), and 1 U of Taq DNA polymerase.Thermal cycling was performed as follows: 94°C for 2 min; 40 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min; and 72°C for 10min. The PCR products thus obtained were analyzed in 1.5% agarose gel containing ethidium bromide.

2.6 Sequencing of the Positive Clones

For each patient, both strands of twenty to forty E2-containing plasmids were sequenced with Cy5 M13 (universal) forward (5'-GTAAAACGACGGCCAG-3') and Cy5.5 M13 (universal) reverse primers (5'-CAGGAAACAGCTATGAC-3') by the dideoxy chain termination method with the Visible Genetics Sequencing Kit (Visible Genetics, Toronto, ON) on a Long-Read Tower™ OpenGene™ automated DNA sequencer (Visible genetics, Toronto, ON). Briefly, a master mix containing 2.2 μl of sequencing buffer (260 mM Tris-Hcl, pH 8.3, 39 mM $MgCl_2$), 1.1 μl of each forward and reverse primers (4 μM), 2.75 μl (3.5 U) of thermo sequence™ enzyme, 2.0 μl of colony PCR product and 5.15 μl of ddH_2O was prepared and 3 μl of the above master mix was aliqotted into each of four PCR tubes containing 3 μl of each ddATP, ddCTP, ddGTP, ddTTP termination mixes. The reaction mixture was overlaid with a small

amount of mineral oil to minimize evaporation, and incubated at 94° C for 2 min, 18 cycles of 94° C for 15 sec, 55° C for 15 sec, 70° C of 15 sec; 15 cycles of 94° C for 15 sec, 70° C for 15 sec. Once the reaction was complete, three microlitres of Stop Loading Dye was added into each of termination tubes, and the samples were heated at 90° C and chilled on ice. The Next step was running the sequencing gel plate. Six percent sequencing gel was filled into the plate and toasted in the toaster. When complete, the gel was inserted into the sequencer and 1x TBE buffer was added to top and bottom reservoirs. Two microlitres of each A, C, G, T termination mix was loaded in four different lanes using the multi-channel pipette.

2.7 Genetic Analysis

Viral sequences were derived from patient samples, Genbank or the LANL HCV database (hcv.lanl.gov). Ig genes were obtained from the NREF database. Sequences were locally aligned using ClustalW v1.82 and BLAST v2.26 followed by manual alignment according to the IMGT numbering scheme (imgy.cines.fr, (Lefranc et al., 2003)) which is approved by the Human Genome Organization (HUGO) Nomenclature Committee (HGNC). HVR1 sequences were then tested for Ig similarity (relative to IMGT alignment files of the variable domains of human heavey (n=984); light kappa (n=843); light lambda (n=1212); TCRα (n=111); and TCRβ (n=346) chains) using a scoring system based on a position-specific scoring matrix that quantifies the physical and chemical properties of aas. Specially, for a given region of comparison, the similarities of each constituent aa to the corresponding position of reference sequences, was multiplied by a value found in the updated Dayhoff aa similarity matrix (Feng et al., 1984; Jones et al,, 1992), derived form the observed frequencies of aa substitutions in

related proteins which is a function of the physical (size, charge, and hydropathy) and chemical similarities between aas. Using this algorithm, similarity scores range up to 1500/aa for identity (maximum score = 1500 L; where L = sequence length). Other similarity scoring matrices were used including an identity matrix, PAM250 and the BLOSUM62 matrices. All the matrices produced results supporting similar trends of change for HVR1 evolution (not shown). The three-dimensional model structure of the N-terminal portion of E2 was made using the SWISS-MODEL server (Schwede et al., 2003) for HCV genotype 2a (AAF59944) relative to the tertiary structure of the variable domain of humanized antibody 4D5 (PDB:1FVD). The model image was generated with both VMD v1.8.2 and Raster3D v2.6.

2.8 Statistical Analysis

Population prevalence of differences in mutant populations was determined by Chi-squared analysis and differences in Ig similarity scores were determined using the Student's *t* test.

2.9 Expression of N-terminally 6xHis Tagged HCV E2 Proteins

2.9.1 Preparation of N-terminally 6x His tagged E2 fragments

2.9.1.1 PCR mutagenesis to insert 6x His tag and restriction enzyme sites to E2 fragments

From colony PCR products obtained above, based on the amino acid sequence alignment, four clones from patient A and two clones from patient B were chosen to be expressed in E.coli or in Spodoptera frugiperda (Sf9) insect cells. To isolate target E2 genes, and to add restriction sites, an initiator ATG and stop codon TAG as well as

sequences coding for 6 His residues into each construct, the E2 sequences containing HVR1 were amplified by PCR using the selected colony PCR products as templates. Amplification of the E2 fragment was performed with the following primers. Each forward primer contained a NheI site (shown in italics), followed by an initiator ATG , sequences encoding 6x His residues, another initiator ATG, and nucleotides representing the 5' end of E2 (in boldface type). Forward primers for cloning into the pET 17a vector: HCV1aE2-A13-NheI-F: 5'-ATTA*GCTAGC*ATGCATCACCATCACCATCACATG**GAAACCTACGTCACCGG**-3' (HCV 1a A13 variant). HCV1aE2-A19-NheI- F: 5'-ATTA*GCTAGC*ATGCATCACCATCACCATCACATG**ACAACCTACGTCACCGGG**-3' (HCV1aA19variant), HCV1aE2-A5/9-NheI-F: 5'-ATTA*GCTAGC*ATGCATCACCATCACCATCACATG**ACAACCGACATCTCCGGG**-3' (HCV1a A5 and A9 variants). The forward primers for Baculovirus expression vector were similar to those above except they carried a BgLII restriction site (*AGATCT)* and were called HCV1aE2-A13-BgLII-F, HCV1aE2-A19-BgLII-F, HCV1aE2-A5-BgLII-F, and HCV1aE2-A9-BgLII-F respectively. Reverse primers: HCV1aE2-A13/9-R: 5'-CATG *GAATTC*CTA**CGCGCCCGACCTGCGGTGGTCCC**-3' (HCV 1a A13 and A9 variants), HCV1aE2-A5/19-R: 5'-CATG*GAATTC*CTA**CGCGCCCGACCTGTCG GTCGTCCC**-3' (HCV 1a A19 and A5 variant). Each reverse primer carried 24 bases complementary to the E2 (in boldface type) followed by TAG stop codon, and an EcoR1 site (shown in italitics). PCR was performed by using Taq PCR Core Kit (Qiagen, Chatsworth, CA) as indicated by the manufacturer. One microlitre of PCR product was amplified in a volume of 100 μl containing 2.5 U Taq DNA polymerase, 1x PCR buffer, 1x Q solution, 200 μM of each dNTP, 2.0 mM MgCl and 0.5 μg of each forward and

reverse primer. The thermal cycle files were programmed as follows: 40 cycle of 94 °C for 30 sec, 45 °C for 30 sec, and 72° C for 1 min, and one cycle at 72 °C for 7 min.

The PCR products were analyzed by electrophoresis through a 1.5 % agarose gel and staining with ethidium bromide.

PCR products were purified by using a QIAquick PCR purification kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions where 100 μl of PCR product was purified and eluted with 50 μl of ddH_2O.

2.9.1.2 Subcloning of modified E2 fragments into pCR2.1 vector

The purified PCR products were subcloned into pCR2.1 vector using the TOPO TA cloning kit (Invitrogen, Carlsbad, CA) as described above.

Colony PCR was performed to verify the presence of the insert by using Taq Core Kit from Qiagen (Qiagen, Chatsworth, CA) as described above. The PCR products were analyzed by electrophoresis through a 1.5 % agarose gel and staining with ethidium bromide. One positive recombinant clone per variant was selected to be sequenced to ensure that the insert DNA encoded the authentic HCV E2 sequence, using Cy5 M13 (universal) forward (5'-GTAAAACGACGGCCAG-3') and Cy5.5 M13 (universal) reverse primers (5'-CAGGAAACAGCTATGAC-3') by the dideoxy chain termination method with the Visible Genetics Sequencing Kit (Visible Genetics, Toronto, ON) on a Long-Read Tower™ OpenGene™ automated DNA sequencer (Visible Genetics, Toronto, ON) as described above. The selected clones were propagated in competent

E.coli (TOPO TA Cloning, Invitrogen) in LB broth containing ampicillin (100μg/ml) at 37°C overnight with shaking.

2.9.1.3 Purification of pCR2.1-E2 plasmids by miniprep

The plasmid DNA was extracted from 3 ml of overnight culture using a QIAprep Spin Miniprep Kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. The extracted DNA was eluted with 50 μl of ddH_2O and stored at -20°C.

2.9.1.4 Double digestion of pCR2.1-E2 plasmids

The plasmid DNA thus obtained was double digested with NheI (BgLII) and EcoRI. In brief, one hundred microlitres of digestion mixture containing 50 μl of plasmid DNA, 10 μl of 10x NheI (BgLII) buffer (New England Biolab), 1 μl of 100x BSA (100 μg/ml) (New England Biolab), 10 μl of NheI (BgLII) (10 U/μl) (New England Biolab), 5μl of EcoRI (20 U/μl) (New England Biolab) and 24 μl of ddH_2O were mixed gently and incubated at 37°C water bath for one hour. The digested DNA fragments were separated by gel electrophoresis in 1 % agarose gels containing 0.5 μg/ml ethidium bromide. DNA fragments corresponding to the sizes of 470 bp (E2 fragments) were exercised from agarose gels.

2.9.1.5 Gel -purification of E2 fragments

The E2 fragments were purified from agarose gels by using a QIAquick gel extration kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. E2 fragments contained in approximately 300 mg agarose gel were eluted from the column with 30 μl of ddH_2O.

2.9.2 Expression of 6xHis tagged E2 proteins in E.coli

2.9.2.1 Preparation of the expression vector pET17b

2.9.2.1.1 Mini-preparation of the pET17b vector

The pET17b vector (Novagen) was propagated in competent E.coli (TOPO TA Cloning, Invitrogen) in LB broth containing ampicillin (100μg/ml) at 37° C overnight with shaking. The plasmid DNA was extracted from 3 ml of overnight culture using a QIAprep Spin Miniprep Kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. The extracted DNA was eluted with 50 μl of ddH_2O and stored at -20°C.

2.9.2.1.2 Double digestion of pET17b vector

The pET17b vector was double digested with NheI and EcoRI in a volume of 100 μl containing 50 μl of pET17b (50 μg), 10 μl of 10x NheI buffer (New England Biolab), 1 μl of 100x BSA (100 μg/ml) (New England Biolab),10 μl of NheI (10 U/μl) (New England Biolab), 5μl of EcoRI (20 U/μl) (New England Biolab) and 24 μl of ddH_2O. The digestion mixture was incubated at 37° C water bath for one hour. The digested pET17b plasmids were separated by gel electrophoresis in 1 % agarose gels containing 0.5 μg/ml ethidium bromide and purified from agarose gels by using the QIAquick gel extraction kit (Qiagen, Chatsworth, CA) as described above.

2.9.2.2 Cloning of 6x His tagged E2 fragments into pET17b vector

The NheI/EcoRI digested E2 fragments were ligated into NheI/EcoRI digested pET17b plasmid with T_4 DNA ligase (Biolab). The ligation mixture containing 5 μl of digested E2 fragment (250ng), 2 μl of digested pET17b plasmid (200ng), 2 μl of 10x T_4 DNA

ligase buffer (containing 10mM ATP) (New England Biolab), 1 μl of T_4 DNA ligase (400U/ml) (New England Biolab) and 10 μl of ddH_2O was incubated in a 37°C water bath for 10 minutes.

For the generation of recombinant E. coli clones, five microlitres of ligation products (designated pET17b-E2) were used to transform 50 μl of competent E. coli cells using TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA). Briefly, five microlitres of ligation products were mixed gently with 50 μl of competent cells. The cell suspension was incubated on ice for 30 min, heat-shocked for exactly 30 sec in the 42°C water bath without shaking and immediately incubated on ice for 2 min. Two hundred and fifty microlitres of S.O.C medium was added to the cells and incubated at 37°C with constant shaking at 225 rpm for 1 hour. Forty microlitres of the transformed cells were then plated on LB agar plates containing 100 μg/ml ampicillin, and incubated overnight at 37°C.

Colony PCR was performed to see if E2 fragment was ligated into pET17b vector, using Taq Core Kit (QIAGEN). Five positive recombinant colonies (shown in white) per sample were randomly picked from the LB agar plates containing 100 μg/ml ampicillin, and each was added to 20 μl of PCR reaction mixture containing 1x PCR buffer, 125 μM of each dNTP, 0.04 μg of forward primer, 0.04 μg of reverse primer, and 1 U of Taq DNA polymerase. The forward and reverse primers used here were the E2 primers containing NheI and EcoRI restriction enzyme sites as we used for PCR mutagenesis to insert 6-His tag and restriction enzyme sites. As described above, they were HCV1aE2-A13-NheI-F, HCV1aE2-A13-EcoRI-R; HCV1aE2-A19-NheI-F, HCV1aE2-A19-EcoRI-R; HCV1aE2-A5-NheI-F, HCV1aE2-A5-EcoRI-R; HCV1aE2-A9-NheI-F, HCV1aE2-A9-EcoRI-R. Thermal cycling was performed as follows: 94°C for 2 min; 40 cycles of

94° C for 1 min, 55° C for 1 min, and 72° C for 1 min; and 72° C for 10 min. The PCR products thus obtained were analyzed in a 1.5% agarose gel containing ethidium bromide.

Two positive clones per sample were propagated in competent E.coli (TOPO TA Cloning, Invitrogen) in 3 ml of LB broth containing ampicillin (100μg/ml) at 37° C overnight with shaking.

2.9.2.3 Purification of pET17b-E2 plasmids by miniprep

The plasmid DNA was extracted from 3 ml of overnight culture using a QIAprep Spin Miniprep Kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. The extracted DNA was eluted with 50 μl of ddH_2O and stored at -20°C.

2.9.2.4 Mini-digestion of pET17b-E2 plasmids

The plasmid DNA thus obtained was double digested with NheI and EcoRI to check for the correct E2 insert that should be the same size as the corresponding PCR. In brief, twenty microliters of digestion mixture containing 10 μl of plasmid DNA, 2 μl of 10x NheI buffer (New England Biolab), 0.2 μl of 100x BSA (100 μg/ml) (New England Biolab), 2 μl of NheI (10 U/μl) (New England Biolab), 1μl of EcoRI (20 U/μl) (New England Biolab) and 4.8 μl of ddH_2O were mixed gently and incubated in a 37°C water bath for one hour. The digested DNA fragments were separated by gel electrophoresis in 1.5 % agarose gels. DNA fragments corresponding to the sizes of 470 bp (E2 fragments) were E2 inserts.

2.9.2.5 Preparation of BL21 competent cells:

Forty microliters of BL21 cells (Novagen) were plated on LB agar plates containing 100μg/ml ampicillin, and incubated overnight at 37°C. A single colony was picked and incubated in 100 ml of LB broth at 37°C with vigorous agitation. The growth of the culture was monitored by measuring the OD every 15-20 min. When the OD_{600} value reached about 0.4, the culture was incubated on ice for 5 min, and then the cells were recovered by centrifugation at 2700g for 10 min at 4°C. The cell pellets were resuspended in 30 ml of ice-cold $MgCl_2$-$CaCl_2$ solution (80mM $MgCl_2$ and 20mM $CaCl_2$). The cells were recovered by centrifugation at 2700g for 10 min at 4°C. Four millilitres of ice-cold $CaCl_2$ (0.1M) was used to resuspend the pellets and 140 μl of DMSO was added to the resuspended cells. After incubation on ice for 15 min, the resuspended cells were mixed with an additional 140 μl of DMSO, dispensed into several chilled, sterile microfuge tubes and immediately snap-freezed by immersing the tightly closed tubes in a bath of liquid nitrogen. The competent cells thus obtained were stored at -80°C until used.

2.9.2.6 Transformation of competent E.coli BL21 cells with pET17b-E2 plasmids

For expression of recombinant E2 proteins, the pET-E2 plasmids were used to transform the competent E.coli BL21 cells. Briefly, two microlitres of pET17b-E2 plasmids were mixed gently with 100 μl of competent cells. The cell suspension was incubated on ice for 10 min, heat-shocked for 2 min in the 42°C water bath without shaking and immediately incubated on ice for 10 min. One millilitre of LB medium was added to the cells and incubation at 37°C with constant shaking at 250 rpm for 1 hour was performed. One hundred microlitres of the transformed cells were then plated on LB agar plates containing 100μg/ml ampicillin, and incubated overnight at 37°C.

2.9.2.7 Mini-Protein induction:

Three colonies from each LB agar plate were randomly picked, and each was used to inoculate 3 ml of LB broth containing 100μl/ml ampicillin and incubated overnight at 37°C with shaking. Next morning, one hundred microlitres of the cell culture was added to 3 ml of LB broth containing ampicillin and incubated at 37°C with shaking until the density of the cells reached 0.3-0.6 OD_{600}. When the desired OD was reached, one millilitre of aliquot of the cell culture (0.3-0.6 OD_{600}) was taken and spun at 13,000 rpm for 1 min, and the pellet was resuspended in 2x SDS sample buffer with the volume defined by the formula: V_{SDS} (μl) = OD_{600} x 100. This was the T_0 sample. Another one millilitre of the cell culture (0.3-0.6 OD_{600}) was used for glycerol stock. One millimolar of IPTG was added to the remaining 1 ml of cell culture (0.3-0.6 OD_{600}), and the cells were further grown for 3 hours at 37°C with vigorous shaking. The recombinant E2 proteins were expressed in BL21 cells following induction with IPTG. The cells were harvested by centrifugation and the pellet was resuspended in 2x SDS sample buffer with the volume defined by the same formula as described before. This was the T_3 sample. All the samples were boiled for 10 min and stored at -20 °C until used.

2.9.2.8 SDS-PAGE

SDS-Polyacrylamide gel electrophoresis was performed by the method described by Laemmli (Laemmli et al, 1970), with 12 % resolving gel, 4 % stacking gel and a Bio-Rad mini-gel apparatus. Protein samples (5 μl) were loaded into the wells and run at 180 V for approximately 1 hour. To assess the expression level of the recombinant proteins,

the separated proteins were either stained with Coomassie blue or analyzed by western blots.

2.9.2.9 Western Blot

The separated proteins from SDS-PAGE were equilibrated with a transfer buffer on a rocker for 10 min and electrotransferred onto nitrocellulose membranes at 100 V for 1 hour by using a Mini Trans-Blot Electrophoretic Transfer Cell (Biorad, US) according to the manufacturer's instruction. The membranes were blocked with 1% BSA in PBS containing 0.1 % Tween 20 (blocking buffer) for at least 1 hour, then incubated at room temperature for 1 hour with 6x His monoclonal antibody (BD Biosciences, US) at a dilution of 1:5000 in blocking buffer. After 1 h of incubation, the membranes were washed three times for 10 min each with 10 ml of PBS containing 0.1% Tween 20 (washing buffer). Then, the membranes were incubated for 1 hour with alkaline phosphatase conjugated anti-mouse polyvalent immunoglobulin (Sigma, US) at a dilution of 1:6000 in blocking buffer. The membranes were washed 3x10 min with washing buffer and 10 min with alkaline phosphatase buffer (0.1 M Tris-HCl, 0.1 M NaCl and 5 mM MgCl). Antibody was visualized by development with NBT and BCIP using Alkaline Phosphatase Conjugate Substrate Kit according to the manufacturer's instructions. In brief, the membranes were immunostained with 10 ml of substrate solution containing 100 μl of NBT (2.5 mg/ml), 100 μl of BCIP (2.5 mg/ml) and 400 μl of 25x AP colour developing buffer (2.5 M Tris-HCl, 2.5 M NaCl and 125 mM MgCl).

2.9.2.10 Mega-Protein Induction:

Five microliter of cell culture glycerol stock from mini-induction was used to inoculate 25 ml of LB broth containing 100μl/ml ampicillin and incubated overnight at 37°C with shaking. Next morning, all the 25 ml of the cell culture was added to 500 ml of pre-warmed LB broth containing ampicillin and incubated at 37° C with shaking until the density of the cells reached 0.3-0.4 OD_{600}. When desired OD was reached, a 1 ml aliquot of the cell culture (0.3-0.6 OD_{600}) was taken and spun at 13,000 rpm for 1 min, and the pellet was resuspended in 2x SDS sample buffer with the volume defined by the formula: V_{SDS} (μl) = OD_{600} x 100. This was the T_0 sample. To induce the recombinant protein expression, IPTG was added to the remaining cell culture (0.3-0.4 OD_{600}) to a final concentration of 1 mM, and the cells were further grown for 3 hours at 37°C with vigorous shaking. The recombinant E2 proteins were expressed in BL21 cells following induction with IPTG. One millilitre aliquot of the cell culture was taken and spun, and the pellet was resuspended with 2x SDS sample buffer with the volume defined by the same formula as described before. This was the T_3 sample. The cells were harvested at 4° C by centrifugation at 4000 rpm for 20 min, and the pellet was resuspended in 20 ml of cell lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 1 protease inhibitor tablet, 1 mM PMSF and 10 mM β-mercaptoehanol). Then the suspension was incubated on ice for at least 30 min, sonicated by ultrasound for 1 min and immediately put on ice for 30 sec. After centrifugation at 10000 rpm and 4°C for 20 min, the supernatant (soluble fraction of the recombinant protein) was saved and stored at 4°C until used, and the pellet was resuspended in 10 ml of pH 8 insoluble fraction protein buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM β-mercaptoehanol, 10mM immidazole and 8 M urea), which

generated the insoluble fraction of recombinant protein. Fifty microliter aliquots from each fraction weres taken and mixed with 50 μl of 2x SDS sample buffer. All the samples were boiled for 10 min and stored at -20° C until used.

Five microliters of each sample (T_0, T_3, soluble fraction, and insoluble fraction) was run on SDS-PAGE as described before. The separated proteins were stained with coomassie blue to see where the recombinant E2 protein was located (soluble fraction or insoluble fraction).

2.9.3 Expression of recombinant E2 proteins in baculovirus-infected Sf9 cells

Recombinant E2 proteins were expressed in baculovirus-infected Sf9 cells using the BaculoGold Kit (BD Biosciences, Mississauga, ON) according to manufacturer's instructions.

2.9.3.1 Preparation of baculovirus transfer vector PVL1392

The PVL1392 transfer vector was double digested with BgLII and EcoRI in a volume of 100 μl containing 50 μl of PVL1392 (50μg),10 μl of 10x BgLII buffer (New England Biolab), 10 μl of BgLII (10 U/μl) (New England Biolab), 5μl of EcoRI (20 U/μl) (New England Biolab) and 25 μl of ddH_2O. The digestion mixture was incubated in a 37° C water bath for one hour. The digested PVL1392 plasmids were separated by gel electrophoresis in 1% agarose gels containing 0.5 μg/ml ethidium bromide and purified from agarose gels by using the QIAquick gel extraction kit (Qiagen, Chatsworth, CA) as described above.

2.9.3.2 Cloning of 6x His tagged E2 fragments into baculovirus transfer vector PVL1392

The BglII/EcoRI digested E2 fragments were ligated into the baculovirus transfer vector PVL1392 (BD Biosciences, Mississauga, Canada) with T_4 DNA ligase (New England Biolab). The ligation mixture containing 5 μl of digested E2 fragment (250ng), 2 μl of baculovirus transfer vector PVL1392 (200ng), 2 μl of 10x T_4 DNA ligase buffer (containing 10mM ATP) (New England Biolab) , 1 μl of T_4 DNA ligase (400U/ml) and 10 μl of ddH_2O was incubated at 37°C water bath for 10 minutes.

For the generation of recombinant E. coli clones, 5 μl of ligation products (designated PVL1392-E2) was used to transform 50 μl of competent E. coli cells using TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) as described before.

Colony PCR was performed to see if E2 fragment was ligated into PVL1392 vector using Taq Core Kit (QIAGEN). Five positive recombinant clones (shown in white color) per sample were randomly picked from the LB agar plates containing 100 μg/ml ampicillin, and each was added to a 20 μl of PCR reaction mixture containing 1x PCR buffer, 125 μM of each dNTP, 0.04 μg of forward primer, 0.04 μg of reverse primer, and 1 U of Taq DNA polymerase. The forward and reverse primers used here were the E2 primers containing BglII and EcoRI restriction enzyme sites as we used for PCR mutagenesis to insert 6-His tag and restriction enzyme sites. As described above, they were HCV1aE2-A13-BgLII-F, HCV1aE2-A13-EcoRI-R; HCV1aE2-A19-BgLII-F, HCV1aE2-A19-EcoRI-A19-R; HCV1aE2-A5-BgLII-F, HCV1aE2-A5-EcoRI-R; HCV1aE2-A9-BgLII-F, HCV1aE2-A9-EcoRI-R. Thermal cycling was performed as follows: 94°C for 2 min; 40 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1

min; and 72° C for 10min. The PCR products thus obtained were analyzed in 1.5% agarose gels containing ethidium bromide.

Two positive clones per sample were propagated in competent E.Coli (TOPO TA Cloning, Invitrogen) in 3 ml of LB broth containing ampicillin (100μg/ml) at 37° C overnight with shaking.

2.9.3.3 Purification of PVL1392-E2 plasmids by miniprep

The plasmid DNA was extracted from 3 ml of overnight culture using a QIAprep Spin Miniprep Kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. The extracted DNA was eluted with 50 μl of ddH_2O and stored at -20°C.

2.9.3.4 Mini-digestion of PVL1392-E2 plasmids

The plasmid DNA thus obtained was double digested with BgLII and EcoRI to check for the correct E2 insert that should be the same size as the corresponding PCR. In brief, twenty microlitres of digestion mixture containing 10 μl of plasmid DNA, 2 μl of 10x BgLII buffer (New England Biolab), 2 μl of BgLII (10 U/μl) (New England Biolab), 1μl of EcoRI (20 U/μl) (New England Biolab) and 5 μl of ddH_2O was mixed gently and incubated in a 37° C water bath for one hour. The digested DNA fragments were separated by gel electrophoresis in 1.5 % agarose gels. DNA fragments corresponding to the size of 470 bp (E2 fragments) were E2 inserts.

2.9.3.5 Co-transfection of linearized baculovirus DNA and recombinant transfer vector into Sf9 insect cells

Cotransfection of Sf9 insect cells with linearized baculovirus DNA and recombinant transfer vector PVL1392 was performed using BaculoGold Kit (BD Biosciences, Mississauga, Canada) according to manufacturer's instructions. The transfected Sf9 cells were incubated in a humid atmosphere at 27°C for 5 days.

2.9.3.6 Amplify the recombinant baculoviruses

After five days of incubation, the recombinant baculovirus supernatant was collected and used to infect the Sf9 cells. The infected Sf9 cells were incubated at 27°C for 72 hr. This procedure was repeated one more time.

2.9.3.7 Expression of recombinant E2 protein in Sf9 cells

Sf9 cells infected with recombinant baculoviruses were harvested 72 hr after infection. Briefly, the infected cells were harvested by centrifugation at 10,000 rpm for 2 min, and the pellet was resuspended in 300 μl of ddH_2O.

Lysates of infected cells were electrophoresed in 12% polyacrylamide Laemmli gels and proteins were identified by Coomassie blue staining and Western blot analyses using 6xHis monoclone antibody and alkaline phosphatase conjugated anti-mouse polyvalent immunoglobulin as described before.

2.10 Purification of Recombinant E2 Proteins

Purification of 6xHis-tagged proteins was performed by Ni-NTA (Ni^{2+}-nitrilotriacetate-agarose) affinity chromatography (Qiagen, San Diego, US).

2.10.1 Binding proteins to beads:

To pre-equilibrate the Ni-NTA beads with pH 8 insoluble protein buffer (the recombinant E2 proteins are insoluble proteins), four microlitres of Ni-NTA beads were washed three times with 5 ml of insoluble protein buffer (PH 8) and spun at 1000 rpm for 3 min each prior to adding 10 ml of the insoluble protein fraction to resuspend the pelleted beads. The protein was allowed to bind to resin by incubation with end-over-end rocking at room temperature for at least 1 h.

2.10.2 Column elution of protein fractions:

The unbound material was washed with 20 ml of protein buffers (pH8 and pH6.3)(50 mM NaH_2PO_4, 300 mM NaCl, 10 mM β-mercaptoehanol, 20mM immidazole and 8 M urea) (25 ml each) by passing through the column (Qiagen, US), and the gel-bound proteins were eluted with 18ml of the protein buffer pH 4.5 (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM β-mercaptoehanol, 100mM immidazole and 8 M urea) by going through the column. The eluted protein was collected into 6 different tubes with 3 ml for each fraction. Fifty microlitre aliquots from each wash and each pH4.5 fraction were mixed with 50 μl of 2x SDS sample buffer. All the samples were boiled for 10 min and stored at -20° C until used.

An aliquot of each above sample was run on SDS-PAGE as described before. The separated proteins were stained with Coomassie blue to assess the effectiveness of protein purification. The eluted proteins in different pH4.5 fractions were combined together in one tube.

2.11 Determination of the Concentration of Recombinant Proteins

The concentration of recombinant E2 proteins was determined by using the Bio-Rad Protein Assay (Bio-Rad, US) according to the manufacturer instructions. Briefly, Dye reagent was prepared by diluting 1 part Dye Reagent Concentrate with 4 parts distilled deionized water. Five dilutions of BSA (protein standard), ranging from 0.2 to 0.9 mg/ml, was prepared. One hundred microlitres of each standard and sample solution was mixed with 5 ml of diluted dye reagent and incubated at room temperature for at least 5 min. Absorbance for each standard and sample solution was measured with a spectrophotometer at 595 nm. A standard curve was made and comparison to a standard curve provides a relative measurement of protein concentration.

2.12 SDS-PAGE and Western Blot

SDS-PAGE was carried out using equal amounts of crude or purified protein samples as previously described. Recombinant E2 proteins were reacted with a commercial preparation of alkaline phosphatase conjugated goat antibodies raised against the Fab fragment of human IgG at a dilution of 1:15,000 (Catalogue no. A8542, Sigma-Aldrich Canada Ltd). Antibody was visualized by development with NBT and BCIP using Alkaline Phosphatase Conjugate Substrate Kit as previously described.

CHAPTER THREE

RESULTS

3.1 E2 Contains a Sequence Homologous to Immunoglobulin

Analysis of genomic data sets for sequence homology to a query protein using SSEARCH and BLAST alignment algorithms yields E value scores, for pairs of proteins, which indicate the probability that two proteins are related. E value between 1 and 10 are usually found to be related and values less than 0.01 almost always represent homologous proteins (Thompson et al, 2001; Sauder et al, 2000). On amino acid sequence comparison of representative HCV polyproteins against all currently known proteins using the SSEARCH alignment program we found that all nine common genotypes of HCV E2 were homologous to immunoglobulin kappa light chain (IgL_k) (E values from 4.5 to 0.11 and amino acid identities from 23.2 to 38.7%), (Table 1). In particular the amino terminal region of E2 was homologous to the variable region of immunoglobulin kappa light chain ($IgVL_\kappa$), while the E2 carboxyl-terminal region exhibited similarity to the constant region (see genotypes 3a and 5a), suggesting that the entire E2 gene may share structural similarity to immunoglobulin light chains. To assess the significance of these findings, we searched all known virus sequences in GenBank for the presence of similar levels of homology to two germline antibodies using NCBI BLAST. Only HCV and to a lesser extent Herpes Simplex Virus-1 (HSV-1), and some retroviruses were found to possess proteins that share homology with human antibodies (Table 2).

Table 1. Amino acid sequence homology between human immunoglobulin kappa light chain (Ig_k) proteins and regions of E2 from 9 genotypes of HCV

Genotype (Acc. #)*	Protein homologue (PDB #)**	E2 region of polyprotein (aa)	E value	Homology
1a (S1 strain)	IgL_k (1cdob)	87 (406-490)	5.2	24.1
1b (AF333324)	IgL_k (1c5d)	130 (397-526)	4.2	25.4
2a (AF238485)	IgL_k (1fvd)	81 (443-523)	0.11	38.7
2b (AF238486)	IgL_k (2rcs)	86 (438-523)	5.2	30.2
3a	IgL_k (1mcoH)	79 (501-578)	2.8	26.6
(BAA06044)	IgL_k (1dgd)	187 (395-570)	0.65	23.2
3b (D49374)	IgL_k (1c5d)	110 (416-526)	0.41	30.0
4a (Y11604)	IgL_k (2fbj)	163 (506-669)	4.5	23.3
5a (Y13184)	IgL_k (1fveC)	76 (501-576)	1.3!	31.6
6a (CAA72801)				

Data base access numbers are indicated for protein sequences analyzed using NCBI BLAST (* GeneBank accession number), and SSEARCH (** Protein Data Bank, (PDB) accession number). A 4 amino acids insertion was deleted in 6a E2 before alignment with other protein sequences (!).

Table 2 Amino acid homology between germline antibodies and all known virus proteins in GenBank detected by NCBI Blast

Antibody	Viral protein (No. of sequences found)	Degree of homology	
		Identity (overlap aa)	E value
2rcs	HCV E2 polyprotein (58)	28-34% (436-514)	0.50-0.044
	Herpes virus UL6 proteins (3)	21%	5.2-0.85
1gaf	HCV E2 polyrotein (67)	28-34% (436-514)	0.49-0.018
	Herpes virus UL6 proteins (3)	21%	4.9-0.87

3.2 Sequence Alignment of E2 with Immunoglobulin and T cell Receptors

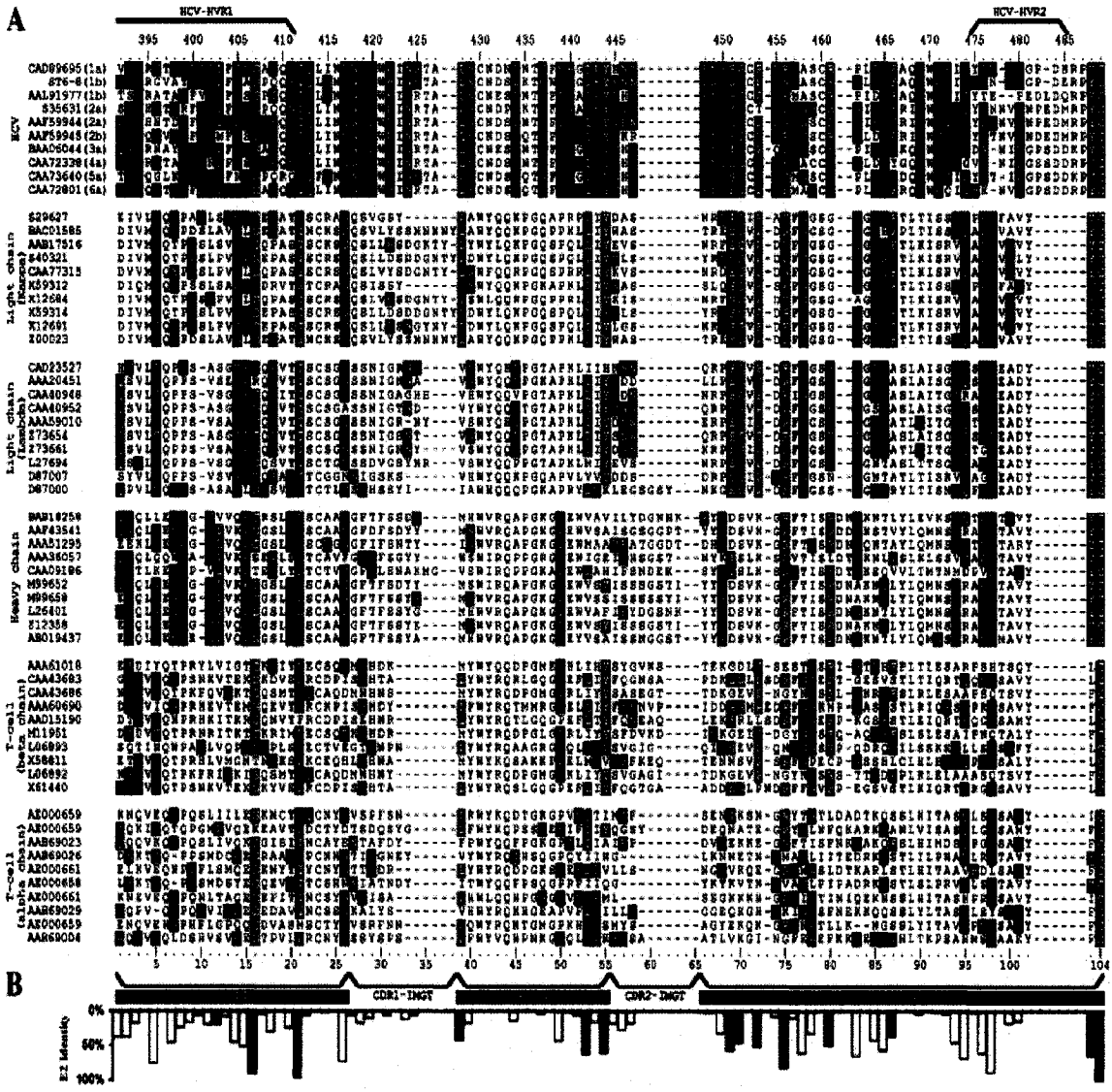
To further assess the sequence similarity between E2 and Igs, we aligned N-terminal portions of a variety of E2 sequences representing the six major HCV genotypes with sequences of various Ig types. By using computer-generated alignments (i.e., BLAST and Clustal W), we found that the regions of highest sequence similarity were almost exclusively restricted to the variable region (v-gene) of Ig molecules and the first 104 aa of E2. Then, the immunogenetics (IMGT) unique numbering system (Lefrane et al, 2003) was used for the comparison among members of the Ig superfamily (Barclay, 2003; Harpaz and Chothia, 1994). This numbering is derived from the sequence alignment and 3-dimensional structural comparisons of Ig superfamily members and provides a definition of the highly conserved framework regions (FR) that support the antibody binding site formed by complementarity determining regions (CDR). Therefore in the IMGT numbering system, conserved amino acids always have the same number and position for immunoglobulins and Ig superfamily v-like genes. The numbering system accommodates the longest forms of CDR which are placed in justified alignments between the framework regions as shown for heavy, light chains and TCR in Fig3.

From this alignment, amino acids in each human Ig framework region that were completely conserved among E2 and Ig types (i.e., kappa light chain) were identified, and the corresponding E2 positions (IMGT numbering: G_{16}, L_{53}, Y_{55}, G_{70}, R_{75}, Y_{103}, and C_{104}) were used as references to manually complete the IMGT formatted alignment with the FR regions of immunoglobulins and TCR. Therefore, FR1, FR2, and FR3 could be aligned with corresponding regions in E2, which allowed the alignment of intervening

sequence regions that were the same or comparable lengths to the CDR1 and CDR2 of heavy and light immunoglobulin chains. In addition, a single gap was inserted before aa position 103 to achieve the most parsimonious alignment because the E2 sequences were 6 aas longer than the corresponding Ig region. Significantly, 14 aa positions were found to be highly maintained (>70% identity), among E2 sequences and all or individual Ig protein types including the heavy and the light chains (kappa and lambda) as well as T cell receptor (TCR) α and β chains (see red bars in Fig. 3B).

For most aa positions in E2, the identical aas could be found at the corresponding positions in Ig members (masked in Fig. 3A), and most of the highly conserved Ig aas are identical to the amino acids at the corresponding positions in E2 (red aa in fig. 3A). There were 14 amino acid sites that had ≥70 shared identities between both E2 and any or all Ig groups and were thus were highly maintained (red bars in Fig. 3B). In the highly conserved sites, such as at aa20 and 80, amino acid replacements were found to be frequently restricted to those with similar biochemistry (masked in blue Fig.3). The sequence homology between E2 and Ig types is high. Twenty out of 104 had >50% common identity and several other sites had conserved chemistry (i.e., aa26 (small hydroxyl); 50, 54, 76, 94 (hydrophobic); and 74 (acidic)). E2 sequences were more homologous to kappa light chains (for example 24% identity of HCV 2a AAF59944.1 with kappa X59312, fig. 2) than to the heavy and TCRα and TCRβ genes (p<0.0001) suggesting specific rather than random relationships of E2 to Ig gene types. The sequence identity of E2 to kappa light chains was significantly higher than calculated for random sequences ($P \leq 4 \times 10^{-72}$); and more importantly randomized Ig consensus sequences ($P \leq 1 \times 10^{-9}$). There were common amino acids (usually P and G) that control

Fig. 3. IMGT sequence alignment of the N-terminal domain of HCV E2 and the variable region of human immunoglobulins and T cell receptor genes. (A) Groups of 10 sequences representing the major genotypes of E2, as well as groups of 5 expressed and 5 germ line variable gene sequences are shown for each set of light kappa, light lambda, heavy, T cell receptor β, and T cell receptor α v-genes; (genes are named with accession numbers). The numbering of the HCV polyprotein is shown at the top along with the location of HVR1 and HVR2. The location of the framework (FR) and complementarity determining regions (CDRs) are shown within the IMGT numbered region at the bottom of the alignment. Amino acid sites with $\geq$70% shared identity between both E2 and any or all Ig groups are masked in red with alternative common substitutions at these sites (of similar chemistry) shown in blue. All other instances of shared aa identity with E2 are masked in black. Average v-gene region identities of E2 relative to all sequences shown (i.e., 100 pair wise comparisons for each group) were significantly higher for the light chain genes (kappa 17.1 ± 2.3%; lambda 16.3 ± 2.1%) than for the heavy (11.8 ± 1.8%) and TCRα (9.9 ± 1.3%) and TCRβ chains (12.1 ± 0.8%) genes ($P < 0.0001$) indicating specific relationships of E2 Ig and TCR gene types rather than a random similarity that would not be expected to differ among groups. (B) The percent identity to E2 is shown in the bar graph with those aa sites that share >70% identity within E2 and any Ig group shown in red.


A
HCV-HVR1
HCV-HVR2
HCV
Light chain (kappa)
Light chain (lambda)
Heavy chain
T-cell (beta chain)
T-cell (alpha chain)
B
CDR1-IMGT
CDR2-IMGT
EI Identity
0%
50%
100%

CDR molecular shape in immunoglobulin heavy and light chains (Kappa and Lambda), such as those at positions 15P and 16G in FR1 of Ig heavy and light chains that were identical in HVR1 of HCV E2. These amino acids allowed bending through adoption of unusual torsion angles (Chothia and Lesk, 1987). Other common features shared by immunoglobulin framework and the corresponding regions of E2 were amino acid content, which was rich in small flexible residues S, A, T, G, as well as specific motifs such as ^{14}SPG^{17}, ^{53}LFYRNN^{58}, ^{85}TDF^{87}, ^{103}YC^{104}. Taken together these data indicated a conserved core structure among E2 and Ig-like genes rather than a chance association of amino acids in the variable CDR regions.

Such a high level of sequence similarity beween E2 and Ig suggests that this region of E2 has the ability to form the basic Ig fold structure (Barclay, 2003; Harpaz and Chothia, 1994) thus providing further support for an Ig-like nature of the N-terminal domain of E2 (Fig. 4). Indeed, three-dimensional modeling shows an overlap of FR regions of the kappa variable domain (in yellow in Fig. 4A) with E2 (superimposed in blue and red), including HVR1 (in red), where all the framework regions align within 2 Å of the kappa peptide chain. Although the E2 structure does not maintain the conserved disulfide bond of Ig formed by C23 and C104 that is guarded by W41, the modeled E2 structure has C41 adjacent to C104 that may form an alternative disulfide bond (data not shown), as occurs for some members of the Ig superfamily (Barclay, 2003).

3.3 Evolution of E2 is Directional with Respect to Ig

Our data showed that a high level of similarity existed between E2 and Ig, and this led us to reason that evolution of E2 could be directional under the immune selective

Fig. 4. (A) A three-dimensional model structure of the N-terminal region of HCV. The aa sequence of the N-terminal domain of E2 genotype 2a (AAF59944) was superimposed on the known tertiary structure of the variable region of humanized antibody (PDB:1FVD). Antibody is shown in yellow (ribbon for beta sheet) with the location of E2 shown in red for HVR1 (aa1–20 IMGT) or blue for the remainder of the Ig-like domain. All the framework regions align within 2 Å of the kappa antibody peptide chain. (B) The location of host protein sequence homology domains shown for the primary structure of E2. Ig-like domains and PKR/eIF2α phosphorylation domains are shown (numbering is from the amino terminus of mature E2 protein).

A

B

pressure. We postulated that immune evasion of persistent HCV variants was mainly due to substitutions in epitopes that raised the extent of similarity between E2 and Ig and that subsequently leads to a lack of recognition of the virus by the immune system. To assess this hypothesis, we analyzed the data from a previous study of immune escape in a HCV 1b infected patient, where two neutralizing epitopes were identified in HVR1 sequences, and aa substitution in these two epitopes lead to escape from recognition by preexisting anti-HVR1 antibodies (Kato et al, 1994). Using a scoring system based on a position-specific scoring of aa similarity using the updated Dayhoff matrix of Jones (Jones et al, 1992) that quantify aa similarity, the similarities of variant neutralizing epitopes in HVR1 sequences with respect to large groups of human Ig chains were determined based on the average sum of similarity score for each amino acid in the epitope region (aa4-17 Fig. 5) with respect to the corresponding amino acids of a large group of Ig (heavy, kappa, lambda) and TCR (α, β). Amino acids substitutions in epitopes that significantly increased Ig similarity scores resulted in the decreases in epitopes binding of patient sera and escape from immune recognition for the 11 month epitope ($P = 0.031$ by paired t test at 11 months relative to prior samples) (Fig. 5B), suggesting that under immune selective pressure, only the mutations of amino acids that are identical to, or more like, the corresponding amino acids in the majority of individual Ig and TCR molecules would be selected, and immune evasion and loss of immunogenicity was directional with respect to the extent of Ig molecular mimicry.

3.4 Evolution of E2 Mimicry During Seroconversion and Interferon Therapy

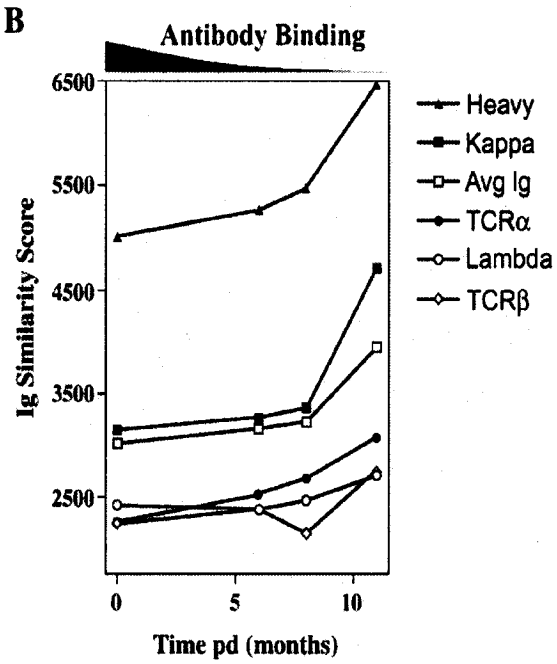
HCV quasispecies variants that competitively replace prior populations (typically

Fig. 5. Evolution of Ig mimicry during immune escape and loss of immunogenicity of neutralizing epitopes in a human HCV 1b patient. (A) The aa sequence of epitopes 1 and 2 (IMGT #4–14 and 7–17) in HVR1 of sequential isolates from a previous analysis of immune escape in a patient infected with HCV 1b (Kato et al., 1994). The sequences of the epitopes at diagnosis and 6, 8 and 11 months post-diagnosis (PD) are shown, identity to the original sequence is shown with dots. Serum antibody reactivity to synthetic peptides corresponding to each sequence is indicated as (++++) positive at 6, 8, 11 and 14 m PD.; (+++) positive at 8, 11, and 14 m; and (++) positive at 8 and 11, or 14 m PD.; (−) indicates a lack of reactivity for all time points. (B) Ig mimicry increases on immune escape and loss of immunogenicity. The sequence of aa4–17 encompassing both overlapping neutralizing epitopes in HVR1 was analyzed with respect to mimicry for variant sequences selected during escape immunity are shown for each Ig group and as an average relative to antibody binding to the variant epitopes. Mimicry scores were significantly elevated among Ig groups for the 11-m variants ($P = 0.031$ by t test).

A

Epitope #		Time (mo.PD)	Antibody binding
1	$HSVRGFTSLFS^{4}_{17}$	0	++++
2	RGFTSLFSAGS	0	++++
1	.....L.....	6	+++
2	..L.....	6	++
1	.G...L.....	8	++
2	..L.....	8	++
1	.G.S.L.....	11	-
2	S.L.....	11	-

B



comprising $>10^6$ HCV virions per ml in blood) are selected by virtue of their superior replicative fitness given selective conditions such as immune pressure and interferon (IFN) responses, and have not been selected by virtue of chance events such as population bottlenecks that trap unselected mutations. Currently, the only available therapy for HCV infection is the combination of IFN and ribavirin, but more than 60% of patients become IFN-resistant and this was linked to changes in HVR1. Therefore, it's good to study the evolution of E2 mimicry of Ig in serial samples from a patient following natural infection, seroconversion and subsequent IFN treatment so that adaptive mutations can be identified. Accordingly we analyzed the evolution of E2 sequences relative to Ig in a patient (patient A) with acute nosocomial HCV 1a infection transmitted from a defined source patient (S), before and during seroconversion and the establishment of persistent infection, followed by an unresponsive IFN treatment (Larke et al, 2002). Patient A had been diagnosed to be HCV infected by the presence of nucleic acid using RT-PCR and had remained serologically negative for a period of time (Larke, et al, 2002). On sequence comparison of a large number of molecular clones (n=137, Fig. 6) of the amino terminal region of E2 from serial patient samples collected at different time points post infection, relative to the direct sequencing of the source virus (S-D), Fig 6 and Fig 7 showed that multiple forms of variants with different Ig similarity scores appeared until seroconversion and persistent infection, when the most divergent viral variants (those with >5 mutations) disappeared to be replaced by viruses having Ig scores to the source virus. The sequence encompassing the amino terminal (8-72aa) region of E2 including major part (8-27, 20aa) of HVR1 is quite homogeneous in the quasispecies of the source virus (S) from the patient with chronic HCV infection, where 94 % (16/17) of the clones analyzed were almost identical, with the remaining 6%

of clones differing by more than six substitutions (Fig. 6). In comparison with the S-D isolate, there was significantly more sequence variation seen in the variants from the nosocomially infected patient where, by day 36 post infection (PI), 6 to 15 amino acid replacements were observed in HVR1 for 14 of 28 clones (50%) (sample A1, pre-seroconversion) and in 8 of 23 (35%) clones at day 46 PI (A2, early-seroconversion). Four weeks after seroconversion, (72 days PI), all the clones in A2.4m (29/29) were found to be identical or genetically similar in sequence to those clones from S, the source virus; there were no clones with more than 5 mutations in A2.4m. The composition of the qusispecies shifted significantly during seroconversion where the most divergent viral variants found in A1.2m, A1.6m, carrying multiple mutations in HVR1 had disappeared after the development of antibody and were replaced by viruses with sequences that were identical or genetically similar to the sequences in S, which had thus become predominant on persistent infection in the A2.4m sample ($P<0.0004$ by Chi-Square analysis) (Fig 6). The quasispecies shift suggests that two types of quasispecies variants co-existed during acute infection, one was restricted (i.e. non-persistent form) while the other became persistent after the selective pressure of seroconversion (i.e persistent forms of variants with few or no mutations). The disappearance of the majority of the more divergent variants coincided with a >10-fold decrease in viral load (Fig. 8) as HCV-specific antibodies were developed. Although the persistent variants in A2.4m sample didn't show an increased average Ig similarity in HVR1 sequences (IMGT aa 1-20) relative to the non-persistent variants in the A1.2m and A1.6m (Fig 7), the least variant-type viruses constituted a genetically distinct population that resists negative selection by the immune response. This suggests that the variants that had evaded the host immune response and became persistent in patient A

may have possessed currently undefined, qualitative, or quantitative properties of mimicry that mediated immune evasion. However, after IFN treatment, all the variants in A IFN+1m and A IFN+2m had shown a significantly increased average Ig similarity scores ($P = 4\times 10^{-12}$) suggesting a role for Ig mimicry in IFN resistance (Fig 7).

Another way to study the impact of immune response on the evolution of viral quasispecies is to examine HCV evolution in individuals with immune system defects that result in low or minimal immune pressure. Several studies showed that there were either no or fewer variations in the dominant amino acid sequence in HVR1 in immunocompromised patients (Odeberg et al, 1997; Ni et al, 1999). We analyzed the evolution of HCV in an immunocompromised lung transplant patient B (patient B) who was infected with HCV 3a via the transplant organ/donor (patient D). Sequential samples were obtained from patient B at different time points after transplant/infection. On analyzing HCV E2 sequences obtained from the donor as well as the sequential samples from the infected recipient, we found that fewer mutations occurred in the dominant HCV variants in patient B (Fig 9) relative to our previous observation in the immunocompetent patient A (Fig 6), and all the variants remained relatively constant average Ig similarity ($P>0.05$) (Fig 10). This suggests that the evolution of E2 mimicry of Ig is driven by immune selective pressure.

3.5 E2 is Antigenically Similar to IgG

To test whether E2 is antigenically similar to IgG, different forms of recombinant E2 proteins (the first 1-141 aa or full length) from patient A and B were expressed in E. coli or baculovirus infected Sf9 insect cells, and the reactivity of the recombinant E2 proteins

Fig. 6. Amino acid sequence analysis of cloned variants in HCV 1a populations in a nosocomial patient during the early phases of HCV primary infection and IFN treatment (Larke et al., 2002). Mutations in the E2 region (E2 aa8–72) are indicated relative to the source virus sequence (S1, direct sequence, (S-D)) for sequences of clones in sequential patient samples from patient A. Dots indicate amino acid identity to the source virus, (S-D). The percentage prevalence of each genotype is shown for each sequence in each population.

A

Direct sequence of patient S

```
E2 #          10        20        30        40        50        60        70
S-D           AAGRGVATITGLFSQGPQQKIQLVNTNGSWHINSTALNCNDSLKTGWIAGLFYRNNFNSSGCPER
```

S (17 clones) (from virus source patient)

```
64.7%         ...............................................................
23.5%         .........A.....................................................
 5.9%         VT.............................................................
 5.9%         SPA.AA...A.....................................................
```

A1.2m (29 clones) 36 days PI, (before seroconversion)

```
13.8%         VT.......A....S.............................N..........R..........
 6.9%         VT...........A..............................N..........R..........
 3.4%         VT.......A....P.............................................R...P......
 3.4%         VT...........P..............................N..........Y..........
 3.4%         ............A...............................................................
 3.4%         VT......N.....S.............................N..........R..........
 3.4%         .........A....S.............................................................
 3.4%         VT.......A....P.............................N.......L..R..........
 3.4%         VT.......A....S.............................N..........R..........
 3.4%         ............H...............................................................
 3.4%         T.A.AA.GLANI..P.AK.N........................N..........Y..........
 3.4%         .PP.AA.G.A....P.............................N..........R..........
 3.4%%        .PP.AA.G.A..S.S.............................N..........R..........
 3.4%         ............P...............................N..........Y..........
 3.4%         .PA.AA.G.A....S.............................N..........R..........
 3.4%         ............P...............................N..........R..........
 3.4%         .........A....S.............................N..........Y..........
 3.4%         VT.......A....S.............................N..........R..........
 3.4%         SPA.AA.G.A....S.............................N..........R..........
 3.4%         .PA.AA.AFA....S.............................................................
 3.4%         ............P..H...........................................Y..........
 3.4%         TPA.AA.G.A....S.............................N..........R..........
 3.4%         .PA.AA.G....L.S.............................N..........R..........
 3.4%         ..A.......A....P............................................R..........
 3.4%         VT...........P..............................N..........R..........
```

A1.6m (23 clones) 46 days PI, (early seroconversion)

```
 8.3%         T..RAA.GLANI..P.AK.N.....I..................N..........Y..........
 4.2%         .P.......A....P.............................N.....................
 4.2%         ............P..H.....................................................
 4.2%         VT.......A....P.............................N..........R..........
 4.2%         VT...........A..............................N..........R..........
 4.2%         ............A...............................N..........R..........
 4.2%         VT...........A..............................N..........Y..........
 4.2%         V...QT.GLAS...P..H.......I..................N..........Y..........
 4.2%         V...QT.GLAS...P..H.......I..................N..........R..........
 4.2%         VT.......A....S.............................N..........R..........
 4.2%         VT...........A..............................N..........Y..........
 4.2%         V...QT.GLAS...P.AH..........................N..........R..........
 4.2%         VT...........S..............................N..........Y..........
 4.2%         AP.......A....S.............................N..........R..........
 4.2%         VT.......A....S.............................N..........R..........
 4.2%         ............A...............................................Y..........
 4.2%         VT...........P..............................N..........R..........
 4.2%         VT...........A..............................N.....................
 4.2%         VT...A.G.AD...L.............................N..........R..........
 4.2%         .....................................................N.....................
 4.2%         ............A...............................................................
 4.2%         .........A....P.....................................................
 4.2%         .........A....S.............................N..........R..........
```

A2.4m (29 clones) 72 days PI (after seroconversion)

```
44.8%         VT.....................................................N.....................
13.8%         ............A...........................................N.....................
10.3%         ............A...........................................N.....................
10.3%         .....................................................N.....................
 7.0%         ...............................................................
 7.0%         VT.............................................................
 3.4%         VT.......A....S.............................N.....................
 3.4%         VT...........S..............................N..........R..........
 3.4%         ............A...............................N.....................
```

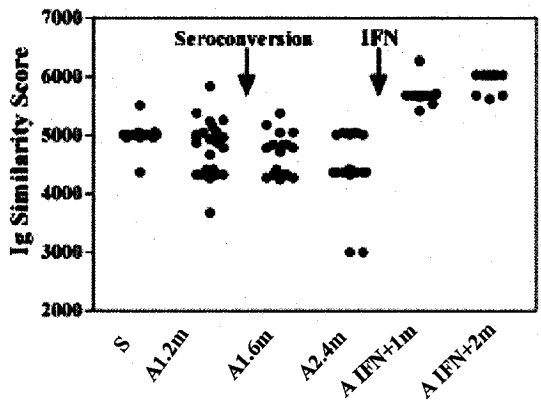
IFN +1m (22 clones) 1 month after IFN treatment

```
31.8%         T.A.AA.GFANI..P.AK.N........................N..........Y..........
22.7%         T.A.AA.GLATI..P.AK.N........................N..........Y..........
 4.5%         T.A.AA.GFANV..P.AK.N........................N..........Y..........
 4.5%         T.A.AA.GFANI..P.AR.N........................N..........Y..........
 4.5%         T.V.AA.GLATI..P.AE.N........................N..........Y..........
 4.5%         T.A.AA.GFANI..P.AK.N.....C..................N..........Y..........
 4.5%         T.V.AA.GFANI..P.AK.N........................N..........Y..........
 4.5%         T.A.VA.GFANI..P.AK.N........................N..........Y..........
 4.5%         T.AHAA.GFANI..P.AK.N........................N..........Y..........
 4.5%         T.A.AT.GFANI..P.AK.N........................N..........Y..........
 4.5%         T.A.AA.GLADIL.P.AK.N........................N..........Y..........
 4.5%         T.A.AA.GFANI..P.AK.N.......G.................N..........Y..........
```

IFN +2m (20 clones) 2 month after IFN treatment

```
55%           T.A.AA.GLATI..P.AK.N........................N..........Y..........
10%           T.A.AA.GIANI..P.AK.N........................N..........Y..........
 5%           T.A.AA.GFANI..P.AK.N.................D.....N..........Y..........
 5%           T.A.AA.GFANI..P.AK.N..I.....................N..........Y..........
 5%           T.A.AA.GLATI..P.AK.N........................N.........HY..........
 5%           T.A.AA.GLATI..P.AK.N..SA....................N..........Y..........
 5%           T.A.AA.GSANI..P.AK.N..............I.........N..........Y..........
 5%           T.A.AA.GLATI..P.AK.N.........D..............N..........Y..........
 5%           T.A.AA.RLATI..P.AK.N........................N..........Y..........
```

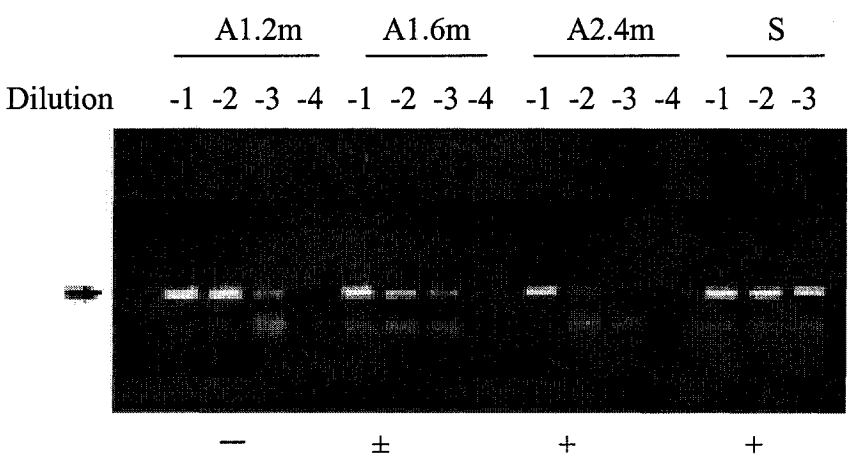
Fig.7 Evolution of Ig (variable region) similarity in HCV 1a patient A (Pat-A) infected by a defined source patient S (Pat-S). Each symbol represents the similarity score for HVR1 (aa1–20 IMGT) of individual molecular clones sampled at each time point. Ig similarity scores represent average scores among 5 groups of Ig and TCR molecules.



Viral populations in patients

Fig.8.Viral load decreased on seroconversion and the establishment of persistent infection in a HCV 1a infected patient A (pat-A). RT-PCR products of serially diluted viral RNA prepared as described in Methods are shown following gel electrophoresis; the location of the amplified DNA is indicated with an arrow. Seroconversion was monitored by EIA and indicated as negative (−), weak reactive (±) and positive (+).

Viral load (RT-PCR)



Immune response (EIA)

Fig. 9. Amino acid sequence analysis of cloned variants in an immunocompromised HCV 3a infected lung transplant patient (Pat-B) infected from patient D (Pat-D). Mutations in the E2 region (E2 aa8–72) are indicated relative to the sequence from patient D (direct sequence, (D-D)) for sequences of clones in sequential patient samples from patient B. Dots indicate amino acid identity to D-D. The percentage prevalence of each genotype is shown for each sequence in each population.

B

Direct sequence of patient D

```
E2 #          10        20        30        40        50        60        70
D-D           AAHRTSMFTSLFSQNTGASQKLQLVNTNGSWHINSTALNCNESINTGFIAGLLYYHKFNATGCPQR
```

Patient D (10 clones)

```
50%           ..........................................................
20%           ..........................................................S.....
10%           ..........................................................Y.....
10%           ..........................................V...............
10%           ................T.........................................S.....
```

Patient B

B1 (9 clones) 8 days

```
44.4%         ..........................................................S.....
22.2%         ..........................................................
11.1%         ................HE........................................S.....
11.1%         ..........................................S...............
11.1%         .....N....................................................S.....
```

B2 (10 clones) 78 days

```
20%           ..Q.......................................................S.....
10%           V....N....................................................S.....
10%           .....N....................................H.....S.....
10%           ................T.........................................S.....
10%           ..............D...........................................
10%           ..Q.......................................................SS....
10%           ..........................................H.....S.....
10%           ..........................................P...........S.....
10%           ..........................................................S.....
```

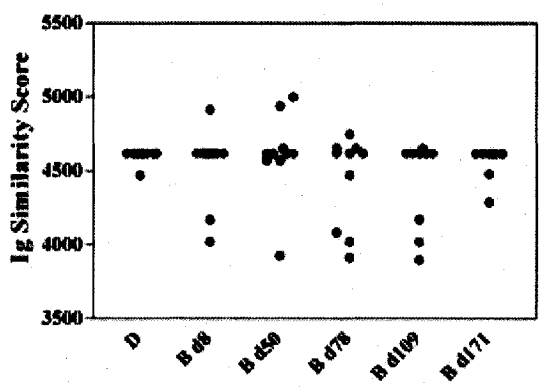
B3 (9 clones) 109 days

```
44.4%         ..........................................................S.....
11.1%         ..........................................................S..R...
11.1%         ..........................................S...........S.....
11.1%         ................C.........................................S.....
11.1%         .....R....................................................S.....
11.1%         ..Q.......................................................S.....
```

B4 (8 clones) 171 days

```
50%           ..........................................................S.....
12.5%         ..........................................Y.........S.....
12.5%         ..........................................................
12.5%         ........I.................................................S.....
12.5%         ..Q.............R.........................................S.....
```

Fig.10. Evolution of Ig similarity in an immunocompromised HCV 3a infected lung transplant patient (Pat-B) infected from patient D (Pat-D). Each symbol represents the average of Ig group similarity scores for HVR1 (aa1–20 IMGT) of individual molecular clones sampled at each time point. The patient letter and time of sampling is indicated for each viral population.



Viral populations in patients

with sera of patient A and anti-human IgG antibody was analyzed. Different forms of E2 fragments were constructed by PCR using colony PCR products (Fig 11). The recombinant plasmid expressing E2, designated pET17b-E2 and PVL1392-E2 were constructed as described in methods. Obtained clones were digested with NheI /Bgl II and EcoR1, 447/990 base pairs (bp) fragments of E2 with 6-His tag could be detected from each of them (Fig 12). Automatic sequencing confirmed that the inserted E2 fragments corresponded to our previously reported data (Fig 13). The recombinant E2 proteins expressed in transformed E. coli or transfected Sf9 insect cells were analyzed by SDS-PAGE and western blot. The recombinant proteins migrated as an approximately 16 kilodalton (kDa) (for the first 141aa of HCV 1a E2) (Fig 14A and 16A) or 36kDa (for the full length of HCV 3a E2) band (Fig 15A). Fig 14B showed the loading controls that were detected by binding to anti-his6 antibody. The concentrations of the purified recombinant E2 proteins estimated by relative to BSA standards were 1.5-1.7 mg/ml.

Western blot was performed to test the reactivity of the recombinant E2 protein with anti-human IgG antibody. All the recombinant E2 proteins (the first 141aa of HCV1a or full length of HCV 3a), whether expressed in E. coli or in baculovirus infected Sf9 insect cells, were found to bind with antibody against human IgG Fab fragment (Fig 14C, 15B and 16B). To assess the extent of the cross-reactivity between E2 and IgG, serially diluted recombinant E2 proteins (A13) and synthetic IgG were used to probe with anti-human IgG Fab fragment. The same amounts of human IgG (combination of IgH-G2a (G2a heavy chain) and IgL-K (kappa light chain)) were loaded together with E2 to compare the reactivity with anti-human-IgG. Dose dependent reactivity of E2 and IgG

was seen with anti-human IgG (Fig 17B). Fig 17A showed total protein staining as a sample loading control (Coomassie blue). On comparison of E2 and human IgG for the reactivity with anti-human IgG, recombinant E2 protein A13 was seen to bind to anti-human IgG slightly stronger than human immunoglobulin kappa light chain (IgL-K) and significantly stronger than immunoglobulin heavy chain (IgH-G2a) (Fig.17B, 17C). The reactivity of E2 and IgG with anti-human-IgG (Fab) was compared with respect to the binding density units per microgram, which were 2, 1.2 and 0.4 for E2, IgL-K and IgH-G2a respectively. However, if the protein amounts are adjusted for molecular weight (15 kDa, 30 kDa and 60 kDa for E2, IgL-K and IgH-G2a respectively), the binding density per molecule would be 2 : 2.4 : 1.6 for E2, IgL-K and IgH-G2a respectively, so E2 binds anti-human IgG (Fab) slightly weaker than IgL-K but stronger than IgH-G2a The low reactivity of anti-human IgG to IgH-G2a is probably due to the fact that anti-human-IgG (Fab fragment) is composed of a mixture of antibodies against the four different IgG isotypes (IgG1, IgG2, IgG3 and IgG4), which results in the mixture of anti-IgG-CH1 antibodies making up the total antibody preparation. Taken together, these data indicated that structural similarity existed between E2 and human Ig molecules, and the immunoglobulin antigenic cross-reactivity involved the protein component of the N-terminal portion of E2.

Fig. 11 Construction of E2 fragments that are expressed in E.coli and baculovirus infected SF9 cells. Different forms of E2 fragments were constructed by PCR mutagenesis using colony PCR products. TI and TT denote translation initiation and translation termination codons respectively.

Baculovirus expression system

E.Coli expression system

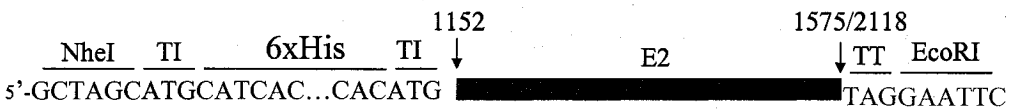


Fig. 12. Digestion of the recombinant plasmids with either BgLII and EcoRI (left panel), or NheI and EcoRI (right panel). Lane 1-4 shows four recombinant plasmid digestion results, PVU1392-A13, PVU-A19, PVU-A5, PVU-A9 respectively for left panel, and PET-A13, PET-A19, PET-A5, PET-A9 respectively for right panel.

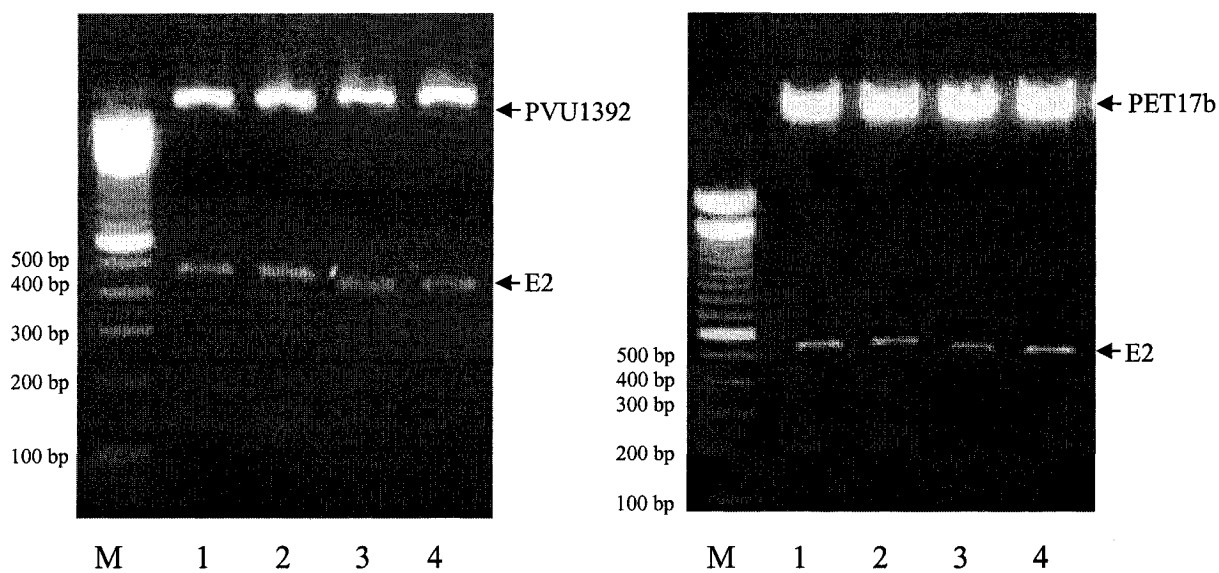

PVU1392
E2
500 bp
400 bp
300 bp
200 bp
100 bp
M 1 2 3 4
PET17b
E2
500 bp
400 bp
300 bp
200 bp
100 bp
M 1 2 3 4

Fig. 13. Amino acid sequences of E2 fragments expressed in E.coli or baculovirus infected SF9 cells. A: Amino terminal regions (aa1-141) of HCV 1a E2 isolated from patient A. Sequences are shown with a single letter amino acid code where residue is different from A13 sequence, and with a dot where residue is identical. B. Full length of HCV 3a E2 (aa1-322) isolated from patient B. Different amino acids in B2 are indicated with respect to B8 sequence. Dots indicate the amino acid identity to B8.

A:

```
                1         10        20        30        40        50        60        70
A13: MHHHHHHMETYVTGGAAGRGVATITGLFSHGPQQKIQLVNTNGSWHINSTALNCNDSLKTGWIAGLFYRNNFNSSGCPERLASC
A19: ................PP.AA.G.A..S.S.........................N.........R.......
A5:  ........T.DIS.RT.A.AA.GLATI..P.AK.N.....................N.........Y.......
A9:  ........T.DIS.KT.A.AA.GFANI..P.AK.N.....................N.........Y.......

          80        90        100       110       120       130       140
A13: RRLTDFAQGWGPITHANGSGPDERPYCWHYPPKPCGIVDAKTVCGPVYCFTPSPVVVGTTDRSGA
A19: ............................................................
A5:  ............................................................
A9:  ............................................................
```

B:

```
                1         10        20        30        40        50        60        70
B8:   MHHHHHHMFSGVDASTRTIAGNAAHRTSVFTSLFNTGASQKLQLVNTNGSWHINSTALNCNESINTGFIAGLLYYHKFNATGC
B109: ................H.........NM.....................................................S...
        80                                                                              160
B8:   PQRLSSCKPITFFKQGWGPLTDANISGPSDDKPYCWHYPPRPCATVKALSVCGPVYCFTPSPVVVGTTDSKGAPTYTWGENET
B109: ..T.......................................................................................
           170                                                                         240
B8:   DVFLLQSLRPPKGRWFGCTWMNSTGFVKTCGAPPCNIYGGGGDPKNESDSFCPTDCFRKHPEATYSRCGAGPWLTPRCMVDYP
B109: ........................V.................S.....................................
          250                                                                          320
B8:   YRLWHYPCTVNFTLFKVRMFVGGFEHRFTAACNWTRGERCDIEDRDRSEQNPLLHSTTEFAILPCSFTPMPALSTGLIHLH
B109: ................................................................................
```

Fig. 14. Anti-human-IgG binds N-terminal regions of HCV 1a E2 proteins expressed in E.coli. Amino terminal fragments (aa1-141) of HCV 1a E2 were cloned and expressed in E.coli. A: Recombinant E2 proteins were analyzed by SDS-PAGE and detected by Coomassie blue staining (sample loading control). B. Loading controls that were detected by binding to anti-his6 antibody. C. Anti-human-IgG (Fab fragment) bind to recombinant E2 proteins. Samples: U-uninduced HCV 1a clone A13-1, 1-HCV 1a-A13, 2-HCV 1a-A19, 3-HCV 1a-A5, 4-HCV 1a-A9, 5-human IgG (positive control).

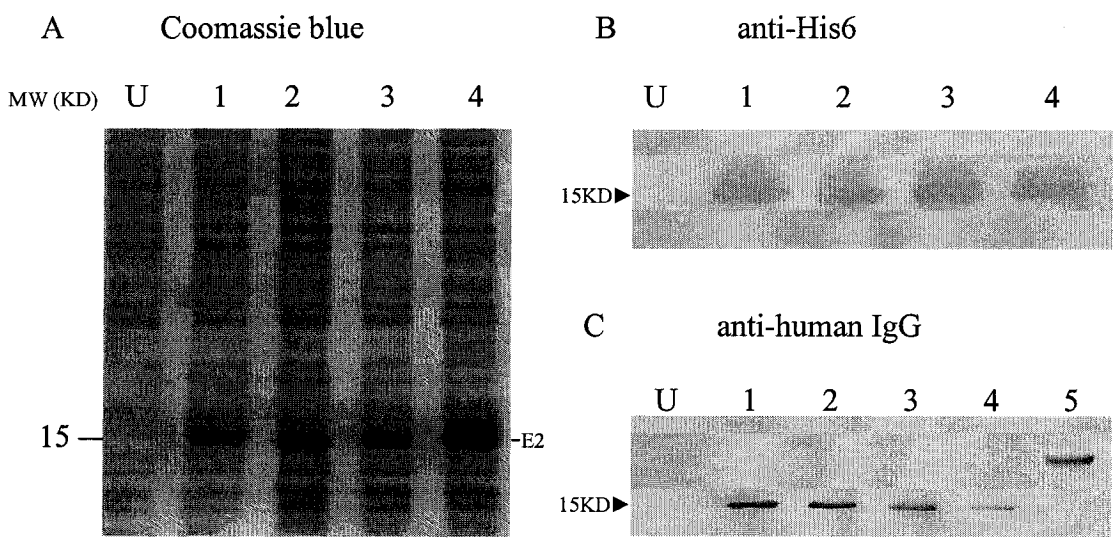
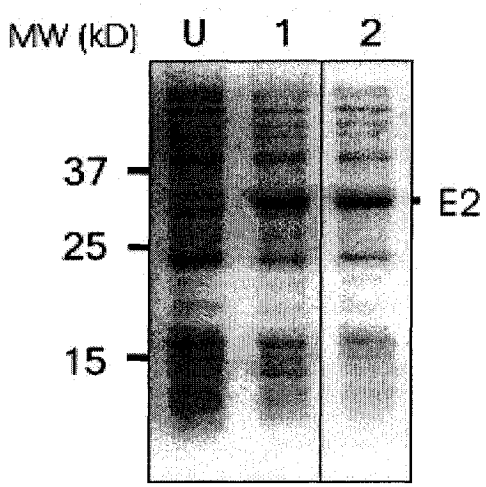

A
Coomassie blue
MW (KD)
U
1
2
3
4
15
E2
B
anti-His6
U
1
2
3
4
15KD
C
anti-human IgG
U
1
2
3
4
5
15KD

Fig. 15. Anti-human-IgG binds full length HCV 3a E2 proteins expressed in E.coli. Full length HCV 3a E2 (1-322) were cloned and expressed in E.coli. A: Recombinant E2 proteins were analyzed by SDS-PAGE and detected by Coomassie blue staining (sample loading control). B. Anti-human-IgG (Fab fragment) bind to recombinant E2 proteins. Samples: U-uninduced HCV 3a clone B-d8-1, 1-HCV 3a-B8, 2-HCV 1a-B109. Two E.coli protein bands were observed to bind anti-human IgG (arrows).

Coomasie blue



anti-human IgG

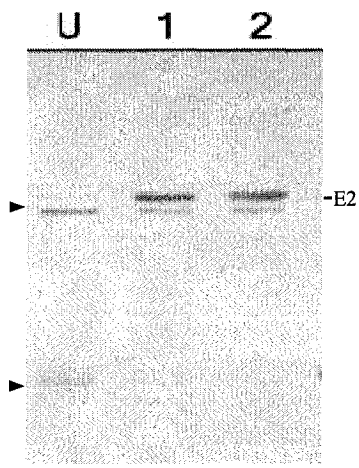


Fig 16. Anti-human-IgG binds N-terminal regions of HCV 1a E2 proteins expressed in baculovirus infected insect cells. Amino terminal fragments (aa1-141) of HCV 1a E2 were cloned and expressed in SF9 cells by using recombinant baculovirus. A: Recombinant E2 proteins were analyzed by SDS-PAGE and detected by Coomassie blue staining (sample loading control).B. Anti-human-IgG (Fab fragment) bind to recombinant E2 proteins. Samples: 1-HCV 1a-A13, 2-HCV 1a-A19, 3-HCV 1a-A5, 4-HCV 1a-A9, C-cell lysate from empty baculovirus vector.

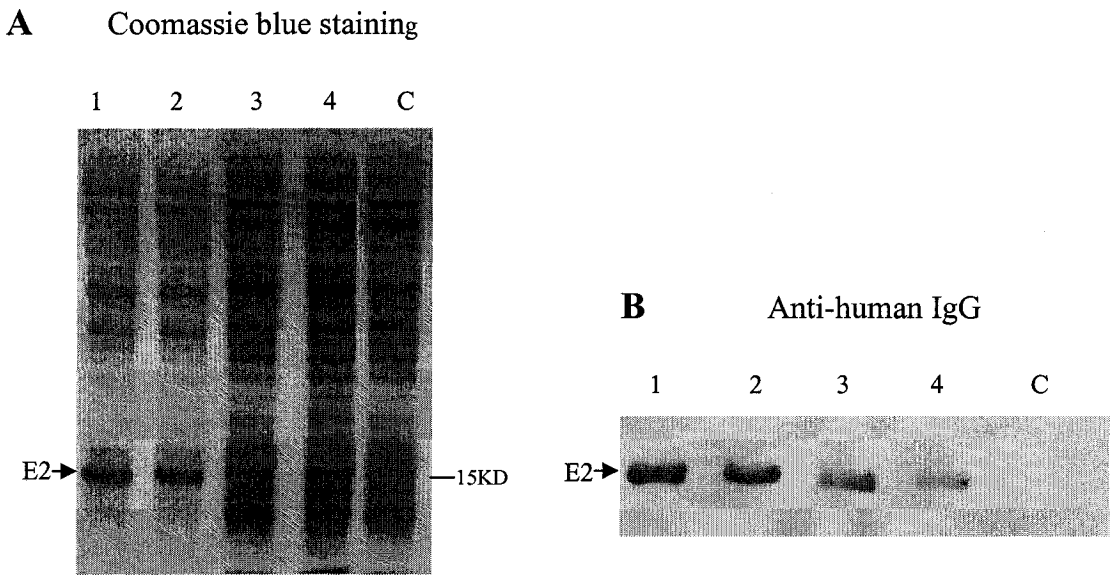
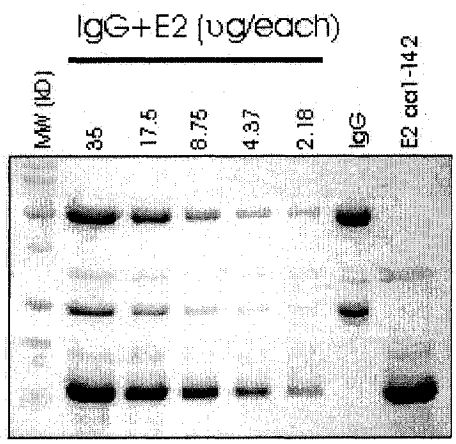
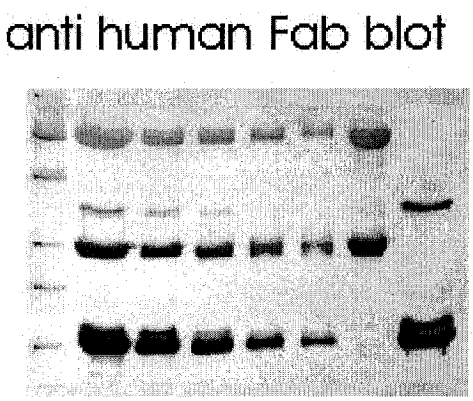

A
Coomassie blue staining
1
2
3
4
C
E2
15KD
B
Anti-human IgG
1
2
3
4
C
E2

Fig. 17. Reactivity of E2 with anti-human-IgG. Serially diluted E2 proteins (A13), together with same amounts of human IgG (combination of IgH-G2a (G2a heavy chain) and IgL-K (kappa light chain)), were subjected to SDS-PAGE and western blot using anti-human-IgG Fab fragment. A: Sample loading control that was detected by Coomassie blue staining. B. Dose dependent reactivity of E2 and IgG with anti-human-IgG. C: The reactivity of E2 and IgG with anti-human-IgG (Fab) was compared with respect to the density units per microgram.

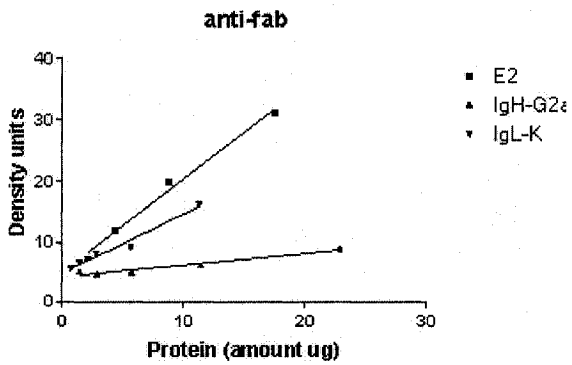
A


IgG+E2 (ug/each)
MW (kD)
35
17.5
8.75
4.37
2.18
IgG
E2 aa1-142

B


anti human Fab blot

C


anti-fab
Density units
Protein (amount ug)
E2
IgH-G2a
IgL-K
0
10
20
30
40

CHAPTER FOUR

DISCUSSION

HCV is the only RNA virus infecting humans (except retroviruses) that persists in the majority of infected individuals. The underlying mechanism for persistent infection is unknown. Molecular mimicry has been proposed as an immune evasion strategy used by the virus to promote persistent infection. Our study provides evidence that HCV E2 is a molecular mimic of immunoglobulins, and that evolutionary change in the extent of mimicry correlates with the host-dependent ability of HCV to escape immune recognition. Therefore we believe that HCV E2 mimicry may be one mechanism that supports persistent infection through immune evasion.We showed that N-terminal region of E2 is highly homologous to the variable region of human immunoglobulin in sequence and structure. This result strongly supports the conclusion that E2 possesses an immunoglobulin superfamily (IgSF) domain (Barclay, 2003), and especially that E2 had an IgSF domain in its N-terminal region. In agreement with the operational definition proposed for molecular mimicry (Cohen, 2004), both immunoglobulin and E2 are bound by common ligands, since the recombinant E2 proteins (aa1-141 or full length) were found to be recognized by anti-human-IgG. The retrospective analysis of immune evasion showed that loss of immunogenecity for the immune escape epitope coincide with the increases in similarity to Ig and TCR types, consistent with a model of immune evasion through mimicry. Thus, it is very reasonable to assume that HCV escape immune recognition by mimicry of structures of the most variable aspect of the immune system itself (i.e. the variable region). Molecular mimicry of immunoglobulins by HCV is a unique and efficient way to circumvent the immune response because it is focused

on the variable regions of immunoglobulin molecules which are themselves the effectors as well as possible targets for humoral immunity and thus must also represent tolerated antigens. Host-like antigenic structures are non-immunogenic due to tolerance mechanisms that prevent the synthesis of self-reactive antibodies (Rajewsky, 1996; Starr et al, 2003). The structural diversity of the Ig molecules makes them ideal targets for a flexible pattern of mimicry by forms of E2 which may somehow explain the substantial genetic diversity in HCV virus. Human cytomegalovirus (HCMV) has proved to be particularly inventive in subverting immune control through molecular mimicry (Reddehase , 2002), where the HCMV genome evolved with its host by incorporating genes that mimic major histocompatibility complex (MHC) genes and modulate the host immune response. It is possible to speculate that E2, as a viral homologue of immunoglobulins, may function as an imunomodulator and affect the nature of the immune response, including immmune dysfunction of T, NK, and dendritic cells (Eisen-Vandervelde et al, 2004) to further benefit HCV survival or replication. The evolution of Ig mimicry not only provides a mechanistic explanation of HVR1 variability and loss of immunogenicity, but also insight into host-dependent disease patterns.

Host-dependent evolution of Ig mimicry

Human Ig structure is dependent on both the expression of specific germ-line alleles as well as adaptive somatic mutations to result in multiple groups of v-gene variants that could be characterized as quasispecies groups. There are approximately 100 Ig v-genes which will exist as Ig variant groups that have sequence features depending on the antigens they bind. If there is a threshold level of Ig similarity that is required for immune evasion, and if E2 is adapting to a population of host molecular structures, it is

possible that E2 is either well or poorly matched to the immune repertoire of a given host. If E2 meets or exceeds the threshold of similarity needed to avoid detection, it will presumably not induce an immune response and thus not be driven by antibody selection and would be expected to stay relatively stable with respect to Ig mimicry. Alternatively, a poor match to a novel immune repertoire will result in either a clearance of virus if E2 cannot efficiently evolve to evade immune detection (as seen for 15-20% of human infections) or alternatively will adapt to more closely match the new host environment. We thus expect two patterns for the evolution of mimicry in persistence; either stasis with little or no evolution of Ig mimicry or increased mimicry following seroconversion. Indeed, we found both patterns of HVR1 evolution during persistent infections. Retrospective analysis of previous studies of HCV 1b evolution demonstrated that E2 evolved to become increasingly immunoglobulin-like with significantly increased Ig similarity score in the HVR1 region (Fig. 3). Consistent with our hypothesis we observed few variations and therefore constant average Ig similarity score in the HVR1 regions from the variants in immunocompromised patient B (Fig. 8).

During the progression from acute to chronic infection in immunocompetent patient A, a high frequency of sequence variation in E2, especially in HVR1 resulted in significant changes in the composition of quasispecies. The variants that differ most from the source virus disappeared, while variants having sequences identical or similar to source virus remained and became persistent after seroconversion, suggesting that the source variants represent the most fit or best adapted sequences with respect to viral survival and persistance and confer a selective advantage. Quasispecies complexity in the E2 region was observed to increase before seroconversion. This was beneficial for viral

persistence, since it increased the probability of mimicry to a given host where the greater the E2 structural diversity, the greater the chance that effective levels of mimicry exist to evade immune detection (Abbate et al, 2003; Arenas et al, 2004). Most of the variants that were selected at this time could not evade the immune response indicating that such variants had a selective advantage in establishing infection before seroconversion. The persistent variants selected after seroconversion in sample A 2.4m didn't show increased average Ig similarity scores relative to those present before seroconversion. One possibility is that the source virus sequence may have been well matched to the average immune structures in patient A. Another possibility is that there might be currently undefined, qualitative, or quantitative properties of mimicry that mediated immune evasion and that such properties are possessed by the patient's genotype. On comparison of the E2 sequences from patient A and constant regions of immunoglobulins we found that HCV possesses a sequence in E2 that is also homologous to the constant regions of human immunoglobulins (Fig. 1A). This would suggest that the E2 domain is an average of the structure of both Ig variable domains and Ig constant domains. When E2 is aligned with constant domains, the alignment includes the entire HVR1 (aa 1-27 of E2) instead of the alignment of aa 8-27 of HVR1 that aligns with the first 20 aa of v-genes. In an IMGT structural alignment of Ig V-genes and C-genes, the first region that encompasses beta strand 8 is up to 9 aa longer for C-genes. Both E2 and IgG CH2 domains are 7 aa longer than V-genes in this section. Our alignment not only identifies common aa sequence features, but also constitutes a structural alignment of corresponding beta strands. As the beta strands are different in length, and have low sequence identity, the exact alignment of all strands has not been confirmed. Eventually when the three dimensional structure of E2 and all variable and

constant domains have been determined it will be possible to produce a more refined alignment. Using the current alignment the average similarity score between E2 and the 38 constant genes of immunoglobulin and TCR was monitored using the same method as for variable regions. The similarity scores were found to increase in the persistent variants selected after seroconversion relative to those present before seroconversion in the A2.4m sample (Fig. 2A).

Ig mimicry and IFN

HCV has been shown to resist the inhibitory effects of IFN through the mechanism of molecular mimicry where E2 has a 12 amino acid sequence that is identical to dsRNA-dependent protein kinase (PKR) and eukaryotic initiation factor 2α (eIF2α) (Taylor et al, 2000). IFN treatment of patient A resulted in the selection of mutants with increased average Ig similarity scores within HVR1 suggesting a role for Ig mimicry in IFN resistance. Our data on retrospective analysis of immune escape epitopes (Fig. 3) showed that immune evasion and loss of immunogenicity was coincident with the acquisition of mutations that increased similarity to Ig and TCR types. A previous study showed that the frequency of HVR1-specific T cell responses was significantly higher in patients who recovered after IFN therapy than in those who did not (Del Porto et al, 2000), indicating the correlation between IFN resistance and a lack of immunogenicity. Our observations were consistent with this report since the lack of immunogenicity could be due to increased host mimicry. However, retrospective analysis of Ig mimicry in several IFN-treated HCV patients (Boulestin et al, 2002; Penin et al, 2001) showed both increased and decreased similarity scores in IFN resistant variants following multiple IFN treatments indicating that overall average levels of immune mimicry

(averaged among Ig and TCR types) does not simply explain the variation in HVR1 of IFN resistant variants (unpublished data E.G. Brown, M. Pelchat and Y.W. Hu). The increased average Ig similarity in patient A following IFN treatment suggested that some aspects of immune mimicry may be operating IFN resistance. However, the fact that not all patients show increased average mimicry suggests that it is maybe a particular relationship to specific Ig or TCR molecules that is important for IFN resistance. More detailed analysis of E2 evolution will be needed to assess the role of mimicry in IFN resistance.

Ig mimicry and autoimmune disease

Molecular mimicry has been proposed as one possible mechanism that contributes to autoimmune disease (Zhao et al, 1998; Oldstone et al, 1998). If viral determinants are similar to host antigens with respect to sequence and structure, but are sufficiently different to be recognized as foreign by the host's immune system, the immune response (both humoral and cellular) to the exogenous determinant could then crossreact with the host tissue and eventually lead to autoimmune diseases. Examples of this have been seen for HCMV that has been reported to induce endothelial cell damage in systemic sclerosis (SSc) and atherosclerosis through a mechanism of molecular mimicry of normally expressed endothelial cell surface molecules by HCMV UL94 protein (Lunardi et al, 2005). Similarly, a number of sequence similarities have been identified that may contribute to the pathology of AIDS. It was reported that a high degree of sequence homology existed between the proteins of HIV-1 gp120 and important immune biomolecules such as antigen receptors, immunoglobulins, and T cell receptors, and the cross-reactive antibodies herein generated promoted immune dysfunction (Susal et al,

1993). It is reasonable to speculate that the sequence similarity between E2 and immunoglobulin provides common immunological determinants, which may, through chronic stimulation, induce autoimmune diseases such as mixed type II cryoglobulinemia and others that are associated with 75% of chronic HCV infections (Dammacco et al, 2000; Major et al, 2004). In addition, there is a further association between HCV patients with mixed type II cryoglobulinemia and lymphoproliferative disease seen as non-Hodgkin's B-cell lymphoma, indicating a progression of the former to the latter and that both thus implicate chronic HCV antigenic stimulation (Dammacco et al, 2000). Evidence of Ig mimicry now incriminates E2 as a candidate for the induction of autoimmunity.

Ig mimicry as a model for HCV persistence

The unique discovery of Ig mimicry by E2 provides a model for virus-host co-evolution as a function of Ig mimicry, where HCV possessing threshold levels of mimicry can circumvent host reactions. This model not only advances our concept of viral immune escape and persistent infection but also provides insight into host-dependent disease patterns.

Future directions

We are encouraged by our observation that E2 and Ig molecules share similar structural features. In future work it will be important to characterize the structural similarities between E2 and both types of Ig domains, variable domain and constant domain. The biological relevance of E2 immune mimicry will need to be further studied using various approaches that measure the effect of E2 molecules on immune functions.

REFERENCES

Abbate, I., G. Cappiello, I. O. Lo, R. Longo, D. Ferraro, G. Antonucci, M. Di, V, R. Di Stefano, A. Craxi, M. C. Solmone, A. Spano, G. Ippolito, and M. R. Capobianchi. 2003. Heterogeneity of HVR-1 quasispecies is predictive of early but not sustained virological response in genotype 1b-infected patients undergoing combined treatment with PEG- or STD-IFN plus RBV. J.Biol.Regul.Homeost.Agents **17**:162-165.

Abdulkarim, A. S., N. N. Zein, J. J. Germer, C. P. Kolbert, L. Kabbani, K. L. Krajnik, A. Hola, M. N. Agha, M. Tourogman, and D. H. Persing. 1998. Hepatitis C virus genotypes and hepatitis G virus in hemodialysis patients from Syria: identification of two novel hepatitis C virus subtypes. Am.J.Trop.Med.Hyg. **59**:571-576.

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

Al, R. H., Y. Xie, Y. Wang, and C. H. Hagedorn. 1998. Expression of recombinant hepatitis C virus non-structural protein 5B in Escherichia coli. Virus Res. **53**:141-149.

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

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

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

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

ALTER, H. J., R. H. Purcell, J. W. Shih, J. C. Melpolder, M. Houghton, Q. L. Choo, and G. Kuo. 1989. Detection of antibody to hepatitis C virus in prospectively followed transfusion recipients with acute and chronic non-A, non-B hepatitis. N.Engl.J.Med. **321**:1494-1500.

Anwar, A., N. Ali, R. Tanveer, and A. Siddiqui. 2000. Demonstration of functional requirement of polypyrimidine tract-binding protein by SELEX RNA during hepatitis C virus internal ribosome entry site-mediated translation initiation. J.Biol.Chem. **275**:34231-34235.

Arenas, J. I., J. F. Gallegos-Orozco, T. Laskus, J. Wilkinson, A. Khatib, C. Fasola, D. Adair, M. Radkowski, K. V. Kibler, M. Nowicki, D. Douglas, J. Williams, G. Netto, D. Mulligan, G. Klintmalm, J. Rakela, and H. E. Vargas. 2004. Hepatitis C

virus quasi-species dynamics predict progression of fibrosis after liver transplantation. J.Infect.Dis. **189**:2037-2046.

Arima, N., C. Y. Kao, T. Licht, R. Padmanabhan, Y. Sasaguri, and R. Padmanabhan. 2001. Modulation of cell growth by the hepatitis C virus nonstructural protein NS5A. J.Biol.Chem. **276**:12675-12684.

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

Barclay, A. N. 2003. Membrane proteins with immunoglobulin-like domains--a master superfamily of interaction molecules. Semin.Immunol. **15**:215-223.

Bartenschlager, R., L. Ahlborn-Laake, J. Mous, and H. Jacobsen. 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. J.Virol. **67**:3835-3844.

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

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

Bell, H., K. Hellum, S. Harthug, A. Maeland, S. Ritland, B. Myrvang, L. B. von der, N. Raknerud, K. Skaug, B. G. Gutigard, R. Skjaerven, L. E. Prescott, and P. Simmonds. 1997. Genotype, viral load and age as independent predictors of treatment outcome of interferon-alpha 2a treatment in patients with chronic hepatitis C. Construct group. Scand.J.Infect.Dis. **29**:17-22.

Bergqvist, A. and C. M. Rice. 2001. Transcriptional activation of the interleukin-2 promoter by hepatitis C virus core protein. J.Virol. **75**:772-781.

Berman, M., H. J. ALTER, K. G. Ishak, R. H. Purcell, and E. A. Jones. 1979. The chronic sequelae of non-A, non-B hepatitis. Ann.Intern.Med. **91**:1-6.

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

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

Booth, J. C., U. Kumar, D. Webster, J. Monjardino, and H. C. Thomas. 1998. Comparison of the rate of sequence variation in the hypervariable region of E2/NS1 region of hepatitis C virus in normal and hypogammaglobulinemic patients. Hepatology **27**:223-227.

Borowski, P., K. Oehlmann, M. Heiland, and R. Laufs. 1997. Nonstructural protein 3 of hepatitis C virus blocks the distribution of the free catalytic subunit of cyclic AMP-dependent protein kinase. J.Virol. **71**:2838-2843.

Boulestin, A., K. Sandres-Saune, J. L. Payen, L. Alric, M. Dubois, C. Pasquier, J. P. Vinel, J. P. Pascal, J. Puel, and J. Izopet. 2002. Genetic heterogeneity of the envelope 2 gene and eradication of hepatitis C virus after a second course of interferon-alpha. J.Med.Virol. **68**:221-228.

Bradley, D. W., E. H. Cook, J. E. Maynard, K. A. McCaustland, J. W. Ebert, G. H. Dolana, R. A. Petzel, R. J. Kantor, A. Heilbrunn, H. A. Fields, and B. L. Murphy. 1979. Experimental infection of chimpanzees with antihemophilic (factor VIII) materials: recovery of virus-like particles associated with non-A, non-B hepatitis. J.Med.Virol. **3**:253-269.

Bradley, D. W., J. E. Maynard, H. Popper, E. H. Cook, J. W. Ebert, K. A. McCaustland, C. A. Schable, and H. A. Fields. 1983. Posttransfusion non-A, non-B hepatitis: physicochemical properties of two distinct agents. J.Infect.Dis. **148**:254-265.

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

Brass, V., E. Bieck, R. Montserret, B. Wolk, J. A. Hellings, H. E. Blum, F. Penin, and D. Moradpour. 2002. An amino-terminal amphipathic alpha-helix mediates membrane association of the hepatitis C virus nonstructural protein 5A. J.Biol.Chem. **277**:8130-8139.

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

Bruix, J., J. M. Barrera, X. Calvet, G. Ercilla, J. Costa, J. M. Sanchez-Tapias, M. Ventura, M. Vall, M. Bruguera, C. Bru, and . 1989. Prevalence of antibodies to hepatitis C virus in Spanish patients with hepatocellular carcinoma and hepatic cirrhosis. Lancet **2**:1004-1006.

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

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

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

Chambers, T. J., R. C. Weir, A. Grakoui, D. W. McCourt, J. F. Bazan, R. J. Fletterick, and C. M. Rice. 1990. Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. Proc.Natl.Acad.Sci.U.S.A **87**:8898-8902.

Chang, K. M., B. Rehermann, J. G. McHutchison, C. Pasquinelli, S. Southwood, A. Sette, and F. V. Chisari. 1997. Immunological significance of cytotoxic T lymphocyte epitope variants in patients chronically infected by the hepatitis C virus. J.Clin.Invest **100**:2376-2385.

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

Chothia, C. and A. M. Lesk. 1987. Canonical structures for the hypervariable regions of immunoglobulins. J.Mol.Biol. **196**:901-917.

Chung, K. M., J. Lee, J. E. Kim, O. K. Song, S. Cho, J. Lim, M. Seedorf, B. Hahm, and S. K. Jang. 2000. Nonstructural protein 5A of hepatitis C virus inhibits the function of karyopherin beta3. J.Virol. **74**:5233-5241.

Cocquerel, L., S. Duvet, J. C. Meunier, A. Pillez, R. Cacan, C. Wychowski, and J. Dubuisson. 1999. The transmembrane domain of hepatitis C virus glycoprotein E1 is a signal for static retention in the endoplasmic reticulum. J.Virol. **73**:2641-2649.

Cocquerel, L., J. C. Meunier, D. B. Op, D. Bonte, C. Wychowski, and J. Dubuisson. 2001. Coexpression of hepatitis C virus envelope proteins E1 and E2 in cis improves the stability of membrane insertion of E2. J.Gen.Virol. **82**:1629-1635.

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

Cohen, I.R. 2004. Principles of molecular mimicry and autoimmune disease. In: Cunningham, M.W., Fujinami, R.S. (Eds.), Molecular Mimicry, Microbes and Autoimmunity. ASM Press, pp. 17-26.

Colombo, M., G. Kuo, Q. L. Choo, M. F. Donato, E. Del Ninno, M. A. Tommasini, N. Dioguardi, and M. Houghton. 1989. Prevalence of antibodies to hepatitis C virus in Italian patients with hepatocellular carcinoma. Lancet **2**:1006-1008.

Crofts, N., D. Jolley, J. Kaldor, B. van, I, and A. Wodak. 1997. Epidemiology of hepatitis C virus infection among injecting drug users in Australia. J.Epidemiol.Community Health **51**:692-697.

Dammacco, F., D. Sansonno, C. Piccoli, V. Racanelli, F. P. D'Amore, and G. Lauletta. 2000. The lymphoid system in hepatitis C virus infection: autoimmunity, mixed cryoglobulinemia, and Overt B-cell malignancy. Semin.Liver Dis. **20**:143-157.

Del Porto, P., G. Puntoriero, C. Scotta, A. Nicosia, and E. Piccolella. 2000. High prevalence of hypervariable region 1-specific and -cross-reactive CD4(+) T cells in HCV-infected individuals responsive to IFN-alpha treatment. Virology **269**:313-324.

Deleersnyder, V., A. Pillez, C. Wychowski, K. Blight, J. Xu, Y. S. Hahn, C. M. Rice, and J. Dubuisson. 1997. Formation of native hepatitis C virus glycoprotein complexes. J.Virol. **71**:697-704.

Diamantis, I., S. Bassetti, P. Erb, D. Ladewig, K. Gyr, and M. Battegay. 1997. High prevalence and coinfection rate of hepatitis G and C infections in intravenous drug addicts. J.Hepatol. **26** :794-797.

Dienstag, J. L. 1983. Non-A, non-B hepatitis. I. Recognition, epidemiology, and clinical features. Gastroenterology **85**:439-462.

Dienstag, J. L. and H. J. ALTER. 1986. Non-A, non-B hepatitis: evolving epidemiologic and clinical perspective. Semin.Liver Dis. **6**:67-81.

Dubuisson, J. and C. M. Rice. 1996. Hepatitis C virus glycoprotein folding: disulfide bond formation and association with calnexin. J.Virol. **70**:778-786.

Duverlie, G., H. Khorsi, S. Castelain, O. Jaillon, J. Izopet, F. Lunel, F. Eb, F. Penin, and C. Wychowski. 1998. Sequence analysis of the NS5A protein of European hepatitis C virus 1b isolates and relation to interferon sensitivity. J.Gen.Virol. **79 (Pt 6)**:1373-1381.

Duvet, S., L. Cocquerel, A. Pillez, R. Cacan, A. Verbert, D. Moradpour, C. Wychowski, and J. Dubuisson. 1998. Hepatitis C virus glycoprotein complex localization in the endoplasmic reticulum involves a determinant for retention and not retrieval. J.Biol.Chem. **273**:32088-32095.

Eckart, M. R., M. Selby, F. Masiarz, C. Lee, K. Berger, K. Crawford, C. Kuo, G. Kuo, M. Houghton, and Q. L. Choo. 1993. The hepatitis C virus encodes a serine protease involved in processing of the putative nonstructural proteins from the viral polyprotein precursor. Biochem.Biophys.Res.Commun. **192**:399-406.

Eigen, M. 1996. On the nature of virus quasispecies. Trends Microbiol. **4**:216-218.

Eisen-Vandervelde, A. L., Z. Q. Yao, and Y. S. Hahn. 2004. The molecular basis of HCV-mediated immune dysregulation. Clin.Immunol. **111**:16-21.

Enomoto, N., I. Sakuma, Y. Asahina, M. Kurosaki, T. Murakami, C. Yamamoto, N. Izumi, F. Marumo, and C. Sato. 1995. Comparison of full-length sequences of

interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. J.Clin.Invest **96**:224-230.

Enomoto, N., I. Sakuma, Y. Asahina, M. Kurosaki, T. Murakami, C. Yamamoto, Y. Ogura, N. Izumi, F. Marumo, and C. Sato. 1996. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. N.Engl.J.Med. **334**:77-81.

Erickson, A. L., Y. Kimura, S. Igarashi, J. Eichelberger, M. Houghton, J. Sidney, D. McKinney, A. Sette, A. L. Hughes, and C. M. Walker. 2001. The outcome of hepatitis C virus infection is predicted by escape mutations in epitopes targeted by cytotoxic T lymphocytes. Immunity. **15**:883-895.

Esumi, M., T. Rikihisa, S. Nishimura, J. Goto, K. Mizuno, Y. H. Zhou, and T. Shikata. 1999. Experimental vaccine activities of recombinant E1 and E2 glycoproteins and hypervariable region 1 peptides of hepatitis C virus in chimpanzees. Arch.Virol. **144**:973-980.

Ezelle, H. J., S. Balachandran, F. Sicheri, S. J. Polyak, and G. N. Barber. 2001. Analyzing the mechanisms of interferon-induced apoptosis using CrmA and hepatitis C virus NS5A. Virology **281** :124-137.

Failla, C., L. Tomei, and R. De Francesco. 1994. Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. J.Virol. **68**:3753-3760.

Failla, C., L. Tomei, and R. De Francesco. 1995. An amino-terminal domain of the hepatitis C virus NS3 protease is essential for interaction with NS4A. J.Virol. **69**:1769-1777.

Farci, P., J. Bukh, and R. H. Purcell. 1997. The quasispecies of hepatitis C virus and the host immune response. Springer Semin.Immunopathol. **19**:5-26.

Farci, P., A. Shimoda, A. Coiana, G. Diaz, G. Peddis, J. C. Melpolder, A. Strazzera, D. Y. Chien, S. J. Munoz, A. Balestrieri, R. H. Purcell, and H. J. ALTER. 2000. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. Science **288**:339-344.

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

Farci, P., R. Strazzera, H. J. ALTER, S. Farci, D. Degioannis, A. Coiana, G. Peddis, F. Usai, G. Serra, L. Chessa, G. Diaz, A. Balestrieri, and R. H. Purcell. 2002. Early changes in hepatitis C viral quasispecies during interferon therapy predict the therapeutic outcome. Proc.Natl.Acad.Sci.U.S.A **99**:3081-3086.

Feinstone, S. M., A. Z. Kapikian, and R. H. Purceli. 1973. Hepatitis A: detection by immune electron microscopy of a viruslike antigen associated with acute illness. Science **182**:1026-1028.

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

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

Feng D.F., Johnson, M.S., Doolittle, R.F. 1984. Aligning amino acid sequences: comparison of commonly used methods. J.Mol. Evol. 21, 112-125

Fujita, T., S. Ishido, S. Muramatsu, M. Itoh, and H. Hotta. 1996. Suppression of actinomycin D-induced apoptosis by the NS3 protein of hepatitis C virus. Biochem.Biophys.Res.Commun. **229**:825-831.

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

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

Gale, M. J., Jr., M. J. Korth, N. M. Tang, S. L. Tan, D. A. Hopkins, T. E. Dever, S. J. Polyak, D. R. Gretch, and M. G. Katze. 1997. Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein. Virology **230**:217-227.

Garfein, R. S., D. Vlahov, N. Galai, M. C. Doherty, and K. E. Nelson. 1996. Viral infections in short-term injection drug users: the prevalence of the hepatitis C, hepatitis B, human immunodeficiency, and human T-lymphotropic viruses. Am.J.Public Health **86**:655-661.

Ghosh, A. K., M. Majumder, R. Steele, P. Yaciuk, J. Chrivia, R. Ray, and R. B. Ray. 2000. Hepatitis C virus NS5A protein modulates transcription through a novel cellular transcription factor SRCAP. J.Biol.Chem. **275**:7184-7188.

Ghosh, A. K., R. Steele, K. Meyer, R. Ray, and R. B. Ray . 1999. Hepatitis C virus NS5A protein modulates cell cycle regulatory genes and promotes cell growth. J.Gen.Virol. **80 (Pt 5)**:1179-1183.

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

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

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

Harpaz, Y. and C. Chothia. 1994. Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. J.Mol.Biol. **238**:528-539.

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

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

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

Ishikawa, T., K. Shibuya, K. Yasui, K. Mitamura, and S. Ueda. 2003. Expression of hepatitis C virus core protein associated with malignant lymphoma in transgenic mice. Comp Immunol.Microbiol.Infect.Dis. **26**:115-124.

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

Jackson, D., E. Tabor, and R. J. Gerety. 1979. Acute non-A, non-B hepatitis: specific ultrastructural alterations in endoplasmic reticulum of infected hepatocytes. Lancet **1**:1249-1250.

Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. Comput.Appl.Biosci. **8**:275-282.

Kaito, M., S. Watanabe, K. Tsukiyama-Kohara, K. Yamaguchi, Y. Kobayashi, M. Konishi, M. Yokoi, S. Ishida, S. Suzuki, and M. Kohara. 1994. Hepatitis C virus

particle detected by immunoelectron microscopic study. J.Gen.Virol. **75 (Pt 7)**:1755-1760.

Kaneko, T., Y. Tanji, S. Satoh, M. Hijikata, S. Asabe, K. Kimura, and K. Shimotohno. 1994. Production of two phosphoproteins from the NS5A region of the hepatitis C viral genome. Biochem.Biophys.Res.Commun. **205**:320-326.

Kato, N. 2001. Molecular virology of hepatitis C virus. Acta Med.Okayama **55**:133-159.

Kato, N., K. H. Lan, S. K. Ono-Nita, Y. Shiratori, and M. Omata. 1997. Hepatitis C virus nonstructural region 5A protein is a potent transcriptional activator. J.Virol. **71**:8856-8859.

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

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

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

Katze, M. G., B. Kwieciszewski, D. R. Goodlett, C. M. Blakely, P. Neddermann, S. L. Tan, and R. Aebersold. 2000. Ser(2194) is a highly conserved major phosphorylation site of the hepatitis C virus nonstructural protein NS5A. Virology **278**:501-513.

Kim, D. W., Y. Gwack, J. H. Han, and J. Choe. 1995. C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. Biochem.Biophys.Res.Commun. **215**:160-166.

Kim, D. W., J. Kim, Y. Gwack, J. H. Han, and J. Choe. 1997. Mutational analysis of the hepatitis C virus RNA helicase. J.Virol. **71**:9400-9409.

Kim, J. L., K. A. Morgenstern, J. P. Griffith, M. D. Dwyer, J. A. Thomson, M. A. Murcko, C. Lin, and P. R. Caron. 1998. Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. Structure. **6**:89-100.

Kim, J. L., K. A. Morgenstern, C. Lin, T. Fox, M. D. Dwyer, J. A. Landro, S. P. Chambers, W. Markland, C. A. Lepre, E. T. O'Malley, S. L. Harbeson, C. M. Rice, M. A. Murcko, P. R. Caron, and J. A. Thomson. 1996. Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. Cell **87**:343-355.

Kittlesen, D. J., K. A. Chianese-Bullock, Z. Q. Yao, T. J. Braciale, and Y. S. Hahn. 2000. Interaction between complement receptor gC1qR and hepatitis C virus core protein inhibits T-lymphocyte proliferation. J.Clin.Invest **106**:1239-1249.

Kiyosawa, K., T. Sodeyama, E. Tanaka, Y. Gibo, K. Yoshizawa, Y. Nakano, S. Furuta, Y. Akahane, K. Nishioka, R. H. Purcell, and . 1990. Interrelationship of blood transfusion, non-A, non-B hepatitis and hepatocellular carcinoma: analysis by detection of antibody to hepatitis C virus. Hepatology **12**:671-675.

Knodell, R. G., M. E. Conrad, J. L. Dienstag, and C. J. Bell. 1975. Etiological spectrum of post-transfusion hepatitis. Gastroenterology **69**:1278-1285.

Koch, J. O. and R. Bartenschlager. 1999. Modulation of hepatitis C virus NS5A hyperphosphorylation by nonstructural proteins NS3, NS4A, and NS4B. J.Virol. **73**:7138-7146.

Kolykhalov, A. A., S. M. Feinstone, and C. M. Rice. 1996. Identification of a highly conserved sequence element at the 3' terminus of hepatitis C virus genome RNA. J.Virol. **70**:3363-3371.

Kolykhalov, A. A., K. Mihalik, S. M. Feinstone, and C. M. Rice. 2000. Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3' nontranslated region are essential for virus replication in vivo. J.Virol. **74**:2046-2051.

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

Kuo, G., Q. L. Choo, H. J. ALTER, G. L. Gitnick, A. G. Redeker, R. H. Purcell, T. Miyamura, J. L. Dienstag, M. J. Alter, C. E. Stevens, and . 1989. An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. Science **244**:362-364.

Kwong, A. D., J. L. Kim, and C. Lin. 2000. Structure and function of hepatitis C virus NS3 helicase. Curr.Top.Microbiol.Immunol. **242**:171-196.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685.

Laporte, J., I. Malet, T. Andrieu, V. Thibault, J. J. Toulme, C. Wychowski, J. M. Pawlotsky, J. M. Huraux, H. Agut, and A. Cahour. 2000. Comparative analysis of translation efficiencies of hepatitis C virus 5' untranslated regions among intraindividual quasispecies present in chronic infection: opposite behaviors depending on cell type. J.Virol. **74** :10827-10833.

Larke, B., Y. W. Hu, M. Krajden, V. Scalia, S. K. Byrne, L. R. Boychuk, and J. Klein. 2002. Acute nosocomial HCV infection detected by NAT of a regular blood donor. Transfusion **42**:759-765.

Lawal, Z., J. Petrik, V. S. Wong, G. J. Alexander, and J. P. Allain. 1997. Hepatitis C virus genomic variability in untreated and immunosuppressed patients. Virology **228**:107-111.

Lefranc, M. P., C. Pommie, M. Ruiz, V. Giudicelli, E. Foulquier, L. Truong, V. Thouvenin-Contet, and G. Lefranc. 2003. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. Dev.Comp Immunol. **27**:55-77.

Li, C., D. Candotti, and J. P. Allain. 2001. Production and characterization of monoclonal antibodies specific for a conserved epitope within hepatitis C virus hypervariable region 1. J.Virol. **75**:12412-12420.

Li, X., L. J. Jeffers, L. Shao, K. R. Reddy, M. de Medina, J. Scheffel, B. Moore, and E. R. Schiff. 1995. Identification of hepatitis C virus by immunoelectron microscopy. J.Viral Hepat. **2**:227-234.

Lin, C., J. A. Thomson, and C. M. Rice. 1995. A central region in the hepatitis C virus NS4A protein allows formation of an active NS3-NS4A serine proteinase complex in vivo and in vitro. J.Virol. **69**:4373-4380.

Lohmann, V., F. Korner, U. Herian, and R. Bartenschlager . 1997. Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. J.Virol. **71**:8416-8428.

Lohmann, V., F. Korner, J. Koch, U. Herian, L. Theilmann, and R. Bartenschlager. 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science **285**:110-113.

Love, R. A., H. E. Parge, J. A. Wickersham, Z. Hostomsky, N. Habuka, E. W. Moomaw, T. Adachi, and Z. Hostomska. 1996. The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. Cell **87**:331-342.

Lozach, P. Y., H. Lortat-Jacob, d. L. de Lacroix, I. Staropoli, S. Foung, A. Amara, C. Houles, F. Fieschi, O. Schwartz, J. L. Virelizier, F. Arenzana-Seisdedos, and R. Altmeyer. 2003. DC-SIGN and L-SIGN are high affinity binding receptors for hepatitis C virus glycoprotein E2. J.Biol.Chem. **278**:20358-20366.

Lu, W., S. Y. Lo, M. Chen, K. Wu, Y. K. Fung, and J. H. Ou. 1999. Activation of p53 tumor suppressor by hepatitis C virus core protein. Virology **264**:134-141.

Lunardi, C., C. Bason, R. Corrocher, and A. Puccetti. 2005. Induction of endothelial cell damage by hCMV molecular mimicry. Trends Immunol. **26**:19-24.

Major, M.E., Rehermann, B., Feinstone, S.M. 2004. Hepatitis C viruses. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology. Lipincott Williams and Wikins, Philadelphia, pp. 1127-1162

Majumder, M., A. K. Ghosh, R. Steele, R. Ray, and R. B. Ray. 2001. Hepatitis C virus NS5A physically associates with p53 and regulates p21/waf1 gene expression in a p53-dependent manner. J.Virol. **75**:1401-1407.

Marrone, A. and R. Sallie. 1996. Genetic heterogeneity of hepatitis C virus. The clinical significance of genotypes and quasispecies behavior. Clin.Lab Med. **16**:429-449.

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

Marusawa, H., M. Hijikata, T. Chiba, and K. Shimotohno. 1999. Hepatitis C virus core protein inhibits Fas- and tumor necrosis factor alpha-mediated apoptosis via NF-kappaB activation. J.Virol. **73**:4713-4720.

Matsumoto, M., T. Y. Hsieh, N. Zhu, T. VanArsdale, S. B. Hwang, K. S. Jeng, A. E. Gorbalenya, S. Y. Lo, J. H. Ou, C. F. Ware, and M. M. Lai. 1997. Hepatitis C virus core protein interacts with the cytoplasmic tail of lymphotoxin-beta receptor. J.Virol. **71**:1301-1309.

McLauchlan, J., M. K. Lemberg, G. Hope, and B. Martoglio . 2002. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. EMBO J. **21**:3980-3988.

Mondelli, M. U. and E. Silini. 1999. Clinical significance of hepatitis C virus genotypes. J.Hepatol. **31 Suppl 1**:65-70.

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

Nakano, I., Y. Fukuda, Y. Katano, S. Nakano, T. Kumada, and T. Hayakawa. 1999. Why is the interferon sensitivity-determining region (ISDR) system useful in Japan? J.Hepatol. **30**:1014-1022.

Neddermann, P., A. Clementi, and R. De Francesco. 1999. Hyperphosphorylation of the hepatitis C virus NS5A protein requires an active NS3 protease, NS4A, NS4B, and NS5A encoded on the same polyprotein. J.Virol. **73**:9984-9991.

Ni, Y. H., M. H. Chang, P. J. Chen, H. Y. Hsu, T. W. Lu, K. H. Lin, and D. T. Lin. 1999. Decreased diversity of hepatitis C virus quasispecies during bone marrow transplantation. J.Med.Virol. **58**:132-138.

Njoh, J. and S. Zimmo. 1997. Prevalence of antibodies to hepatitis C virus in drug-dependent patients in Jeddah, Saudi Arabia. East Afr.Med.J. **74**:89-91.

Nousbaum, J., S. J. Polyak, S. C. Ray, D. G. Sullivan, A. M. Larson, R. L. Carithers, Jr., and D. R. Gretch. 2000. Prospective characterization of full-length hepatitis C virus NS5A quasispecies during induction and combination antiviral therapy. J.Virol. **74**:9028-9038.

Odeberg, J., Z. Yun, A. Sonnerborg, K. Bjoro, M. Uhlen, and J. Lundeberg. 1997. Variation of hepatitis C virus hypervariable region 1 in immunocompromised patients. J.Infect.Dis. **175**:938-943.

Ogata, N., H. J. ALTER, R. H. Miller, and R. H. Purcell. 1991. Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. Proc.Natl.Acad.Sci.U.S.A **88**:3392-3396.

Oh, J. W., T. Ito, and M. M. Lai. 1999. A recombinant hepatitis C virus RNA-dependent RNA polymerase capable of copying the full-length viral RNA. J.Virol. **73**:7694-7702.

Ohkoshi, S., H. Kojima, H. Tawaraya, T. Miyajima, T. Kamimura, H. Asakura, A. Satoh, S. Hirose, M. Hijikata, N. Kato, and . 1990. Prevalence of antibody against non-A, non-B hepatitis virus in Japanese patients with hepatocellular carcinoma. Jpn.J.Cancer Res. **81**:550-553.

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

Oldstone, M. B. 1998. Molecular mimicry and immune-mediated diseases. FASEB J. **12**:1255-1265.

Op, D. B., R. Montserret, S. Duvet, L. Cocquerel, R. Cacan, B. Barberot, M. Le Maire, F. Penin, and J. Dubuisson. 2000. The transmembrane domains of hepatitis C virus envelope glycoproteins E1 and E2 play a major role in heterodimerization. J.Biol.Chem. **275**:31428-31437.

Otsuka, M., N. Kato, K. Lan, H. Yoshida, J. Kato, T. Goto, Y. Shiratori, and M. Omata. 2000. Hepatitis C virus core protein enhances p53 function through augmentation of DNA binding affinity and transcriptional ability. J.Biol.Chem. **275**:34122-34130.

Pavio, N. and M. M. Lai. 2003. The hepatitis C virus persistence: how to evade the immune system? J.Biosci. **28** :287-304.

Penin, F., C. Combet, G. Germanidis, P. O. Frainais, G. Deleage, and J. M. Pawlotsky. 2001. Conservation of the conformation and positive charges of hepatitis C

virus E2 envelope glycoprotein hypervariable region 1 points to a role in cell attachment. J.Virol. **75**:5703-5710.

Pereira, B. J., E. L. Milford, R. L. Kirkman, and A. S. Levey. 1991. Transmission of hepatitis C virus by organ transplantation. N.Engl.J.Med. **325**:454-460.

Pfeifer, U., R. Thomssen, K. Legler, U. Bottcher, W. Gerlich, E. Weinmann, and O. Klinge. 1980. Experimental non-A, non-B hepatitis: four types of cytoplasmic alteration in hepatocytes of infected chimpanzees. Virchows Arch.B Cell Pathol.Incl.Mol.Pathol. **33**:233-243.

Ploegh, H. L. 1998. Viral strategies of immune evasion. Science **280**:248-253.

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

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

Pozzato, G., S. Kaneko, M. Moretti, L. S. Croce, F. Franzin, M. Unoura, L. Bercich, C. Tiribelli, M. Crovatto, G. Santini, and . 1994. Different genotypes of hepatitis C virus are associated with different severity of chronic liver disease. J.Med.Virol. **43**:291-296.

Prince, A. M., B. Brotman, G. F. Grady, W. J. Kuhns, C. Hazzi, R. W. Levine, and S. J. Millian. 1974. Long-incubation post-transfusion hepatitis without serological evidence of exposure to hepatitis-B virus. Lancet **2**:241-246.

Rajewsky, K. 1996. Clonal selection and learning in the antibody system. Nature **381**:751-758.

Ray, R. B., L. M. Lagging, K. Meyer, R. Steele, and R. Ray. 1995. Transcriptional regulation of cellular and viral promoters by the hepatitis C virus core protein. Virus Res. **37**:209-220.

Ray, R. B., K. Meyer, and R. Ray. 1996. Suppression of apoptotic cell death by hepatitis C virus core protein. Virology **226**:176-182.

Ray, R. B., K. Meyer, R. Steele, A. Shrivastava, B. B. Aggarwal, and R. Ray. 1998. Inhibition of tumor necrosis factor (TNF-alpha)-mediated apoptosis by hepatitis C virus core protein. J.Biol.Chem. **273**:2256-2259.

Ray, R. B., R. Steele, K. Meyer, and R. Ray. 1997. Transcriptional repression of p53 promoter by hepatitis C virus core protein. J.Biol.Chem. **272**:10983-10986.

Ray, R. B., R. Steele, K. Meyer, and R. Ray. 1998. Hepatitis C virus core protein represses p21WAF1/Cip1/Sid1 promoter activity. Gene **208**:331-336.

Reddehase, M. J. 2002. Antigens and immunoevasins: opponents in cytomegalovirus immune surveillance. Nat.Rev.Immunol. **2**:831-844.

Reed, K. E. and C. M. Rice. 1999. Identification of the major phosphorylation site of the hepatitis C virus H strain NS5A protein as serine 2321. J.Biol.Chem. **274**:28011-28018.

Reed, K. E. and C. M. Rice. 2000. Overview of hepatitis C virus genome structure, polyprotein processing, and protein properties. Curr.Top.Microbiol.Immunol. **242**:55-84.

Reynolds, J. E., A. Kaminski, A. R. Carroll, B. E. Clarke, D. J. Rowlands, and R. J. Jackson. 1996. Internal initiation of translation of hepatitis C virus RNA: the ribosome entry site is at the authentic initiation codon. RNA. **2**:867-878.

Rijnbrand, R. C. and S. M. Lemon. 2000. Internal ribosome entry site-mediated translation in hepatitis C virus replication. Curr.Top.Microbiol.Immunol. **242**:85-116.

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

Saito, I., T. Miyamura, A. Ohbayashi, H. Harada, T. Katayama, S. Kikuchi, Y. Watanabe, S. Koi, M. Onji, Y. Ohta, and . 1990. Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. Proc.Natl.Acad.Sci.U.S.A **87**:6547-6549.

Sakamuro, D., T. Furukawa, and T. Takegami. 1995. Hepatitis C virus nonstructural protein NS3 transforms NIH 3T3 cells. J.Virol. **69**:3893-3896.

Santolini, E., L. Pacini, C. Fipaldini, G. Migliaccio, and N. Monica. 1995. The NS2 protein of hepatitis C virus is a transmembrane polypeptide. J.Virol. **69**:7461-7471.

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

Sauder, J. M., J. W. Arthur, and R. L. Dunbrack, Jr. 2000. Large-scale comparison of protein sequence alignment algorithms with structure alignments. Proteins **40**:6-22.

Scarselli, E., H. Ansuini, R. Cerino, R. M. Roccasecca, S. Acali, G. Filocamo, C. Traboni, A. Nicosia, R. Cortese, and A. Vitelli. 2002. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. EMBO J. **21**:5017-5025.

Schreiber, G. B., M. P. Busch, S. H. Kleinman, and J. J. Korelitz. 1996. The risk of transfusion-transmitted viral infections. The Retrovirus Epidemiology Donor Study. N.Engl.J.Med. **334**:1685-1690.

Schwede, T., Kopp, J., Guex, N., Peitsch, M.C., 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res.31, 3381-3385

Seet, B. T., J. B. Johnston, C. R. Brunetti, J. W. Barrett, H. Everett, C. Cameron, J. Sypula, S. H. Nazarian, A. Lucas, and G. McFadden . 2003. Poxviruses and immune evasion. Annu.Rev.Immunol. **21**:377-423.

Selby, M. J., Q. L. Choo, K. Berger, G. Kuo, E. Glazer, M. Eckart, C. Lee, D. Chien, C. Kuo, and M. Houghton. 1993. Expression, identification and subcellular localization of the proteins encoded by the hepatitis C viral genome. J.Gen.Virol. **74 (Pt 6)**:1103-1113.

Shang, D., W. Zhai, and J. P. Allain. 1999. Broadly cross-reactive, high-affinity antibody to hypervariable region 1 of the hepatitis C virus in rabbits. Virology **258**:396-405.

Shih, C. M., S. J. Lo, T. Miyamura, S. Y. Chen, and Y. H. Lee. 1993. Suppression of hepatitis B virus expression and replication by hepatitis C virus core protein in HuH-7 cells. J.Virol. **67** :5823-5832.

Shih, J. W., J. I. Mur, and H. J. ALTER. 1986. Non-A, non-B hepatitis: advances and unfulfilled expectations of the first decade. Prog.Liver Dis. **8**:433-452.

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

Shrivastava, A., S. K. Manna, R. Ray, and B. B. Aggarwal . 1998. Ectopic expression of hepatitis C virus core protein differentially regulates nuclear transcription factors. J.Virol. **72**:9722-9728.

Simmonds, P., E. C. Holmes, T. A. Cha, S. W. Chan, F. McOmish, B. Irvine, E. Beall, P. L. Yap, J. Kolberg, and M. S. Urdea. 1993. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. J.Gen.Virol. **74 (Pt 11)**:2391-2399.

Simmonds, P. 2000 Hepatitis C virus genotypes. In: Liang T.J. and Hoofnagle, J.H. (Eds.), Hepatitis C. Academic Press, London, pp. 53-83

Spahn, C. M., J. S. Kieft, R. A. Grassucci, P. A. Penczek, K. Zhou, J. A. Doudna, and J. Frank. 2001. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. Science **291**:1959-1962.

Spangberg, K., L. Wiklund, and S. Schwartz. 2001. Binding of the La autoantigen to the hepatitis C virus 3' untranslated region protects the RNA from rapid degradation in vitro. J.Gen.Virol. **82**:113-120.

Squadrito, G., F. Leone, M. Sartori, B. Nalpas, P. Berthelot, G. Raimondo, S. Pol, and C. Brechot. 1997. Mutations in the nonstructural 5A region of hepatitis C virus and response of chronic hepatitis C to interferon alfa. Gastroenterology **113**:567-572.

Starr, T. K., S. C. Jameson, and K. A. Hogquist. 2003. Positive and negative selection of T cells. Annu.Rev.Immunol. **21**:139-176.

Susal, C., M. Kropelin, V. Daniel, and G. Opelz. 1993. Molecular mimicry between HIV-1 and antigen receptor molecules: a clue to the pathogenesis of AIDS. Vox Sang. **65**:10-17.

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

Tabor, E., R. J. Gerety, J. A. Drucker, L. B. Seeff, J. H. Hoofnagle, D. R. Jackson, M. April, L. F. Barker, and G. Pineda-Tamondong. 1978. Transmission of non-A, non-B hepatitis from man to chimpanzee. Lancet **1**:463-466.

Tan, S. L., H. Nakao, Y. He, S. Vijaysri, P. Neddermann, B. L. Jacobs, B. J. Mayer, and M. G. Katze. 1999. NS5A, a nonstructural protein of hepatitis C virus, binds growth factor receptor-bound protein 2 adaptor protein in a Src homology 3 domain/ligand-dependent manner and perturbs mitogenic signaling. Proc.Natl.Acad.Sci.U.S.A **96**:5533-5538.

Tanaka, T., N. Kato, M. J. Cho, and K. Shimotohno. 1995. A novel sequence found at the 3' terminus of hepatitis C virus genome. Biochem.Biophys.Res.Commun. **215**:744-749.

Tanaka, Y., T. Shimoike, K. Ishii, R. Suzuki, T. Suzuki, H. Ushijima, Y. Matsuura, and T. Miyamura. 2000. Selective binding of hepatitis C virus core protein to synthetic oligonucleotides corresponding to the 5' untranslated region of the viral genome. Virology **270**:229-236.

Tanimoto, A., Y. Ide, N. Arima, Y. Sasaguri, and R. Padmanabhan. 1997. The amino terminal deletion mutants of hepatitis C virus nonstructural protein NS5A function as transcriptional activators in yeast. Biochem.Biophys.Res.Commun. **236**:360-364.

Tanji, Y., M. Hijikata, S. Satoh, T. Kaneko, and K. Shimotohno. 1995. Hepatitis C virus-encoded nonstructural protein NS4A has versatile functions in viral protein processing. J.Virol. **69** :1575-1581.

Tanji, Y., T. Kaneko, S. Satoh, and K. Shimotohno. 1995. Phosphorylation of hepatitis C virus-encoded nonstructural protein NS5A. J.Virol. **69**:3980-3986.

Taylor, D. R., S. T. Shi, and M. M. Lai. 2000. Hepatitis C virus and interferon resistance. Microbes.Infect. **2** :1743-1756.

Taylor, D. R., S. T. Shi, P. R. Romano, G. N. Barber, and M. M. Lai. 1999. Inhibition of the interferon-inducible protein kinase PKR by HCV E2 protein. Science **285**:107-110.

Tellinghuisen, T. L. and C. M. Rice. 2002. Interaction between hepatitis C virus proteins and host cell factors. Curr.Opin.Microbiol. **5**:419-427.

Thompson, J. D., F. Plewniak, R. Ripp, J. C. Thierry, and O. Poch. 2001. Towards a reliable objective function for multiple sequence alignments. J.Mol.Biol. **314**:937-951.

Tsai, S. L., Y. M. Chen, M. H. Chen, C. Y. Huang, I. S. Sheen, C. T. Yeh, J. H. Huang, G. C. Kuo, and Y. F. Liaw. 1998. Hepatitis C virus variants circumventing cytotoxic T lymphocyte activity as a mechanism of chronicity. Gastroenterology **115**:954-965.

Tseng, C.T., Climpel, G.R. 2002 Binding of the hepatitis C virus envelope protein E2 to CD81 inhibit natural killer cell functions. J.Exp.Med 195, 43-49

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

Tu, H., L. Gao, S. T. Shi, D. R. Taylor, T. Yang, A. K. Mircheff, Y. Wen, A. E. Gorbalenya, S. B. Hwang, and M. M. Lai. 1999. Hepatitis C virus RNA polymerase and NS5A complex with a SNARE-like protein. Virology **263**:30-41.

Van Regenmortel, M. H., M. A. Mayo, C. M. Fauquet, and J. Maniloff. 2000. Virus nomenclature: consensus versus chaos. Arch.Virol. **145**:2227-2232.

Vossen, M. T., E. M. Westerhout, C. Soderberg-Naucler, and E. J. Wiertz. 2002. Viral immune evasion: a masterpiece of evolution. Immunogenetics **54**:527-542.

Wasley, A. and M. J. Alter. 2000. Epidemiology of hepatitis C: geographic differences and temporal trends. Semin.Liver Dis. **20**:1-16.

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

Weiner, A. J., M. J. Brauer, J. Rosenblatt, K. H. Richman, J. Tung, K. Crawford, F. Bonino, G. Saracco, Q. L. Choo, M. Houghton, and . 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.

WHO, 1999. Global surveillance and control of hepatitis C. Report of a WHO Consultation organized in collaboration with the Viral Hepatitis Prevention Board, Antwerp, Belgium. J.Viral Hepat. **6**:35-47.

Witherell, G. W. and P. Beineke. 2001. Statistical analysis of combined substitutions in nonstructural 5A region of hepatitis C virus and interferon response. J.Med.Virol. **63**:8-16.

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

Wyke, R. J., K. N. Tsiquaye, A. Thornton, Y. White, B. Portmann, P. K. Das, A. J. Zuckerman, and R. Williams. 1979. Transmission of non-A non-B hepatitis to chimpanzees by factor-IX concentrates after fatal complications in patients with chronic liver disease. Lancet **1**:520-524.

Yamada, N., K. Tanihara, A. Takada, T. Yorihuzi, M. Tsutsumi, H. Shimomura, T. Tsuji, and T. Date. 1996. Genetic organization and diversity of the 3' noncoding region of the hepatitis C virus genome. Virology **223**:255-261.

Yanagi, M., M. St Claire, S. U. Emerson, R. H. Purcell, and J. Bukh. 1999. In vivo analysis of the 3' untranslated region of the hepatitis C virus after in vitro mutagenesis of an infectious cDNA clone. Proc.Natl.Acad.Sci.U.S.A **96**:2291-2295.

Yasui, K., T. Wakita, K. Tsukiyama-Kohara, S. I. Funahashi, M. Ichikawa, T. Kajita, D. Moradpour, J. R. Wands, and M. Kohara. 1998. The native form and maturation process of hepatitis C virus core protein. J.Virol. **72**:6048-6055.

Yoshida, T., T. Hanada, T. Tokuhisa, K. Kosai, M. Sata, M. Kohara, and A. Yoshimura. 2002. Activation of STAT3 by the hepatitis C virus core protein leads to cellular transformation. J.Exp.Med. **196**:641-653.

Yuan, Z. H., U. Kumar, H. C. Thomas, Y. M. Wen, and J. Monjardino. 1997. Expression, purification, and partial characterization of HCV RNA polymerase. Biochem.Biophys.Res.Commun. **232**:231-235.

Zein, N. N., J. Rakela, E. L. Krawitt, K. R. Reddy, T. Tominaga, and D. H. Persing. 1996. Hepatitis C virus genotypes in the United States: epidemiology, pathogenicity, and response to interferon therapy. Collaborative Study Group. Ann.Intern.Med. **125**:634-639.

Zeuzem, S., J. H. Lee, and W. K. Roth. 1997. Mutations in the nonstructural 5A gene of European hepatitis C virus isolates and response to interferon alfa. Hepatology **25**:740-744.

Zhao, Z. S., F. Granucci, L. Yeh, P. A. Schaffer, and H. Cantor. 1998. Molecular mimicry by herpes simplex virus-type 1: autoimmune disease after viral infection. Science **279**:1344-1347.

Zhu, N., A. Khoshnan, R. Schneider, M. Matsumoto, G. Dennert, C. Ware, and M. M. Lai. 1998. Hepatitis C virus core protein binds to the cytoplasmic domain of tumor necrosis factor (TNF) receptor 1 and enhances TNF-induced apoptosis. J.Virol. **72**:3691-3697.

APPENDIX A

Fig.A1 IMGT sequence alignment of the N-terminal domain of HCV E2 and the constant region of human immunoglobulins and T cell receptor genes. Groups of 10 sequences representing the major genotypes of E2, as well as groups of constant gene sequences representing light kappa, light lambda, T cell receptor and the four heavy chain domains (CH1, CH2, CH3 and CH4)(genes are named with accession numbers). The numbering of the HCV polyprotein is shown at the top along with the location of HVR1. The location of the strands (A, B and C) and AB, BC and CD turns are shown within the IMGT numbered region at the bottom of the alignment. Amino acid sites with ≥70% shared identity between both E2 and the most common Ig domain (ie, IGKC and IgG, or 75% of all antibodies from IGKC and IgG) are masked in red with alternative common substitutions at these sites (of similar chemistry) shown in blue. All other instances of shared aa identity with E2 are masked in black.

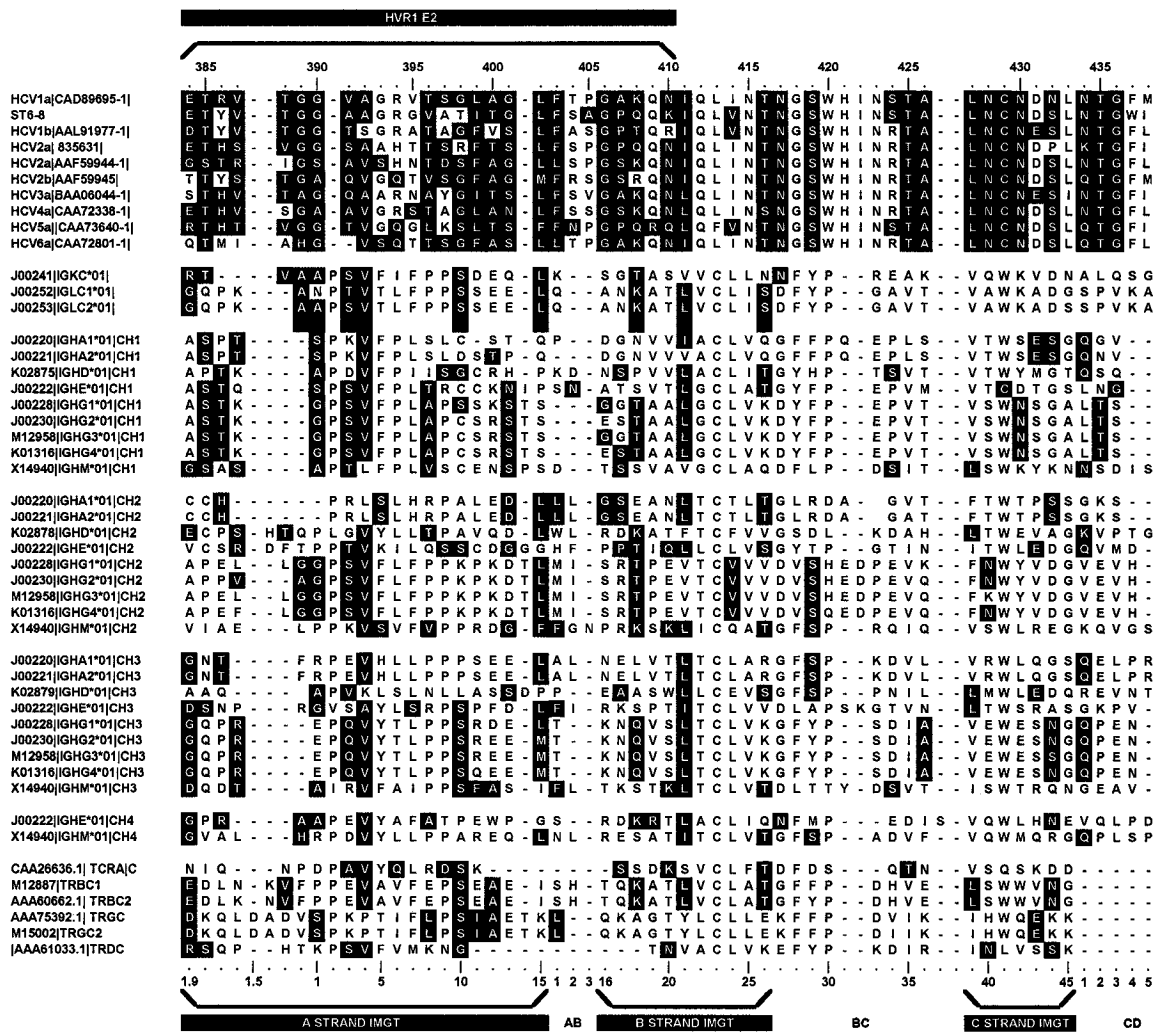
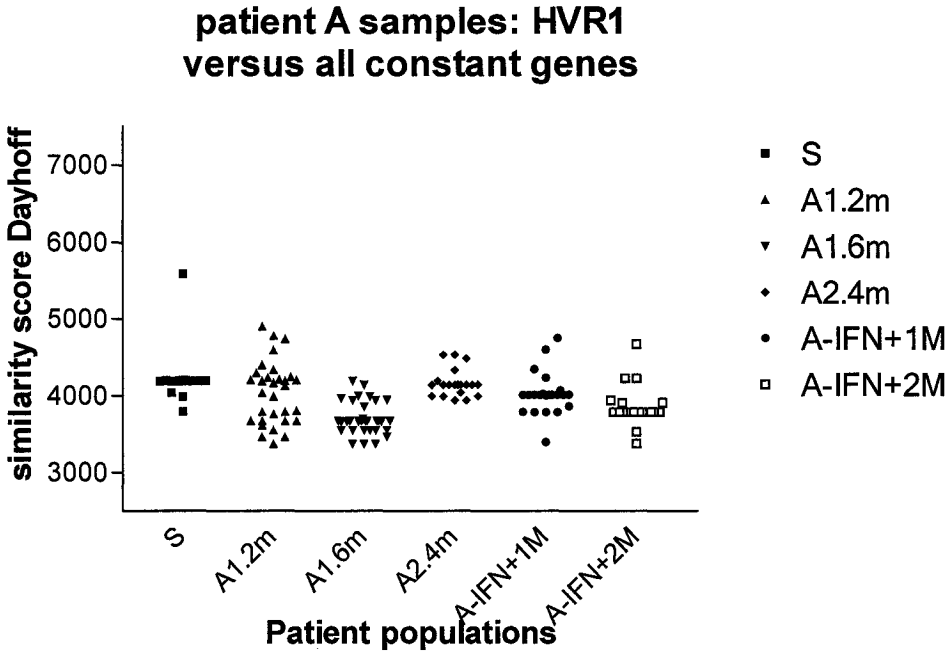

HVR1 E2
385
390
395
400
405
410
415
420
425
430
435
HCV1a|CAD89695-1|
ST6-8
HCV1b|AAL91977-1|
HCV2a| 835631|
HCV2a|AAF59944-1|
HCV2b|AAF59945|
HCV3a|BAA06044-1|
HCV4a|CAA72338-1|
HCV5a||CAA73640-1|
HCV6a|CAA72801-1|
J00241|IGKC*01|
J00252|IGLC1*01|
J00253|IGLC2*01|
J00220|IGHA1*01|CH1
J00221|IGHA2*01|CH1
K02875|IGHD*01|CH1
J00222|IGHE*01|CH1
J00228|IGHG1*01|CH1
J00230|IGHG2*01|CH1
M12958|IGHG3*01|CH1
K01316|IGHG4*01|CH1
X14940|IGHM*01|CH1
J00220|IGHA1*01|CH2
J00221|IGHA2*01|CH2
K02878|IGHD*01|CH2
J00222|IGHE*01|CH2
J00228|IGHG1*01|CH2
J00230|IGHG2*01|CH2
M12958|IGHG3*01|CH2
K01316|IGHG4*01|CH2
X14940|IGHM*01|CH2
J00220|IGHA1*01|CH3
J00221|IGHA2*01|CH3
K02879|IGHD*01|CH3
J00222|IGHE*01|CH3
J00228|IGHG1*01|CH3
J00230|IGHG2*01|CH3
M12958|IGHG3*01|CH3
K01316|IGHG4*01|CH3
X14940|IGHM*01|CH3
J00222|IGHE*01|CH4
X14940|IGHM*01|CH4
CAA26636.1| TCRA|C
M12887|TRBC1
AAA60662.1| TRBC2
AAA75392.1| TRGC
M15002|TRGC2
|AAA61033.1|TRDC
1.9
1.5
1
5
10
15
1
2
3
16
20
25
30
35
40
45
1
2
3
4
5
A STRAND IMGT
AB
B STRAND IMGT
BC
C STRAND IMGT
CD

Fig.A2 Evolution of Ig (constant region) similarity in HCV 1a patient A (Pat-A) infected by a defined source patient S (Pat-S). Each symbol represents the similarity score for HVR1 (aa 1-20 IMGT) of individual molecular clones sampled at each time point. Ig similarity scores represent average scores among the constant genes of light kappa, light lambda, heavy chain and T cell receptors.


patient A samples: HVR1
versus all constant genes
similarity score Dayhoff
7000
6000
5000
4000
3000
S
A1.2m
A1.6m
A2.4m
A-IFN+1M
A-IFN+2M
Patient populations
S
A1.2m
A1.6m
A2.4m
A-IFN+1M
A-IFN+2M

APPENDIX B

Preparation of Culture Medium, Buffers and Solutions

LB Medium (luria-Bertani Medium)
Per liter:
Tryptone 10g
Yeast Extract 5g
Sodium Chloride 10g
Adjust pH to 7.0

LB Agar Containing Ampicillin
Per liter:
Tryptone 10g
Yeast Extract 5g
Sodium Chloride 10g
Bacto Agar 15g
Add ampicillin 200mg

Insect Cell Culture Medium
TNM-FH insect medium
1% penicillin-streptomycin
10% Fetal bovine serum (FBS)

6x DNA Loading Buffer
0.25% (w/v) bromophenol blue
0.25% (w/v) xylene cyanol FF
15% (w/v) Ficoll (Type 400) in water
6 mM EDTA (pH 8.0) (dilution from 0.5 M stock)

2x SDS-PAGE Sample Buffer
Per 8.0 ml
dd H_2O 3.8 ml
0.5 M Tris-HCl (pH 6.8) 1.0 ml
Glycerol 0.8 ml
10% (w/v) Sodium dodecyl sulfate (SDS) 1.6 ml
2-mercaptoethanol 0.4 ml
1% (w/v) Bromophenol Blue 0.4 ml

12% Separation SDS-Polyacrylamide Gel
Per 20 ml
40% Acrylamide/Bis solution 6.0 ml
1.5 M Tris-HCl (pH 8.8) 5.0ml
dd H_2O 8.7 ml
10% (w/v) SDS 200μl
10% (w/v) ammonium persulfate (APS) 100 μl

TEMED 10μl

4% Stacking Gel:
Per 10 ml:
40% Acrylamide/Bis solution 1.0 ml
0.5 M Tris-HCl (pH 6.8) 2.5ml
dd H_2O 6.4 ml
10% (w/v) SDS 100μl
10% (w/v) APS 50 μl
TEMED 10μl

5x SDS Electrode Running Buffer (pH 8.3)
Per liter:
Tris Base 15g
Glycine 72g
SDS 5g
Dilute to 1x when using.

Protein Transfer Buffer – A (pH 8.3)
Per liter:
Tris Base 3.03g
Glycine 14.4g
Methanol 200ml

Western Blocking Buffer
Phosphate-Buffer Saline (PBS) (pH 7.4)
1% Bovine Serum Albumin (BSA)
0.1% Tween-20

Alkaline Phosphatase (AP) Buffer (pH 9.5)
Per liter:
Tris Base 12.1g
Sodium Chloride 5.8g
Magnesium Chloride 1.0g
Adjust pH to 9.5

Coomassie Blue Staining Solution
Per liter:
Methanol 500 ml
Glacial acetic acid 100 ml
Coomassie Blue 2.5g
dd H_2O 400 ml

De-staining Solution
35% Methanol
7% Acetic acid
58% dd H_2O

Protein Lysis Buffer
50 mM NaH_2PO_4
300 mM sodium chloride
1 protease inhibitor tablet
1 mM phenylmethylsulfonylfluoride (PMSF)
10 μM β-mercaptoethanol
Adjust pH to 8.0

Soluble Fraction Protein Buffer
50 mM NaH_2PO_4
300 mM sodium Chloride
10 μM β-mercaptoethanol
Adjust pH to 8.0, 6.3 or 4.5

Insoluble Fraction Protein Buffer
50 mM NaH_2PO_4
300 mM sodium Chloride
10 μM β-mercaptoethanol
8 M Urea

CURRICULUM VITAE

Song Liu
1118-1343 Meadowlands Dr.
Nepean, ON K2E 7E8
(613) 739-2461
E-mail: liu.song@bloodservices.ca

Objective:

Seeking a research oriented position that utilizes my expertise in molecular biology, microbiology, biochemistry and immunology.

Summary:

- 11 years of research and clinical experience in the fields of molecular biology and clinical medicine
- Solid background in molecular biology, microbiology, biochemistry and immunology
- Extensive experience in molecular biology techniques used in the medical research and molecular diagnostics laboratories
- Experience in developing clinical molecular diagnostic assays
- Good computer skills in the applications of Microsoft Office, basic statistical software, BioNumerics software and bioinformatics software
- Good animal manipulating skill
- Ability to work independently and as part of a team
- Good attention to detail and concern for production of reliable results
- Ability to learn and implement new skills quickly
- Good organizational, interpersonal and communication skills
- Strong timing and organizing ability to meet the deadline for clinical and other requirement

Education:

Master of Science in Molecular Microbiology 2005
University of Ottawa, Ottawa

M.D. in Clinical Medicine 1992
Dalian Medical University, Dalian, China

Experience:

Graduate Student 2002-2005

Department of Biochemistry, Microbiology and Immunology, University of Ottawa
Department of Research and Development, Canadian Blood Services

Thesis title: " Hepatitis C virus persistence -immunoglobulin mimicry by E2 protein"

- Extracted viral RNA from HCV infected patient serum samples.
- Amplified E2 genes by PCR, real-time PCR.
- Cloned and sequenced E2 genes.
- Analyzed E2 sequences derived from patient samples, Genbank or the LANL HCV database and tested immunoglobulin similarity of E2 by using bioinformatics software, such as Cluster W, BLAST, IMGT numbering scheme, DNA star, MEGA, SWISS-MODEL server, VMD, Raster 3D etc.
- Expressed recombinant HCV E2 proteins in E. coli, Baculovirus/sf9 cells system.
- Investigated the antibody reactivity to the recombinant E2 proteins in sequential sera by western blot, flow cytometry, and ELISA.

General lab duty:

- Maintained laboratory equipment, supplies and manuals.
- Trained and supervised undergraduates.

Physician 1992- 2000

Department of internal medicine
Liaoning Provincial Hospital Shenyang, China

- Clinical diagnosis and treatment of a variety of diseases of internal medicine
- Design, develop, and carry out Q-PCR for quantitation of hepatitis C virus RNA.
- Participated in research on HCV genotyping

Lab Skills:

Nucleic acids:
DNA and RNA extraction, PCR, RT-PCR, real time PCR, Cloning, DNA purification and sequencing, genotyping, Southern and Northern blotting.
Proteins:
Protein expression in E. coli, Baculovirus/sf9 and mammalian cells system, protein purification, SDS-PAGE, western blot.
Immunology:
Immunocytochemistry, immunoprecipitation, immunofluorescence, flow cytometry, ELISA.
Cell biology:
Tissue culture techniques, microscopy, transfection, virus infection, virus purification and plaque assay.
Animal model:
Small animal manipulating skills.

Computer Skills:

Applications:
Microsoft Word, Excel, Powerpoint, Photoshop.
Bioinformatics:
PubMed, NCBI, GenBank, DNA Star, BioNumberics, NPS@, JustBio, Clustal W, BLAST, Multalin, SSEARCH, MEGA, IMGT numbering scheme, SWISS-MODEL server, VMD Raster 3D, SMS.

Honours:

- Graduate student admission scholarship from university of Ottawa in 2003.
- Excellent achievement award from Shenyang Science and Technology Committee in 1999

Publications:

- Immunoglobulin Mimicry by Hepatitis C Virus Envelope Protein *Virology 332 (2005) 538-549*

References:

Dr. Earl G. Brown, Professor
Department of Biochemistry, Microbiology and Immunology
University of Ottawa
Ottawa, ON K1H 8M5
E-mail: Earl.Brown@uottawa.ca
Tel: (613) 562-5800 X-8310

Dr. Yu-Wen Hu, Prefessor
Senior scientist, Department of research and development
Canadian Blood Services (Head Office)
1800 Alta Vista, Ottawa, ON K1G 4J5
E-mail: yu-wen.hu@bloodservices.ca
Tel: (613) 739-2439