Coláiste na hOllscoile Corcaigh

University College Cork

ucc

Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

Epitope Mapping of the E2 Glycoprotein, Including the Hypervariable Region 1, of the Hepatitis C Virus Genotype 3a, in the Context of Humoral Immune Pressure

Presented by

Nicole Ellen Walsh

for the degree of

Master of Science

University College Cork

Department of Medicine

Head of Department

Prof. Fergus Shanahan

Supervisor

Prof. Liam J. Fanning (Department of Medicine, University College Cork)

Advisor

Prof. Cliona O'Farrelly (School of Biochemistry and Immunology, Trinity College Dublin)

Submission Date

December 2020

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First and foremost, I want to thank my supervisor Liam Fanning. Without his endless patience and guidance this research project would never have been completed. It truly was a privilege to be part of the MVDRL group.

I want to thank ICORN, Bristol-Meyers Squibb, and the members of the ICORN committee for the opportunity to pursue this degree with the ICORN-BMS Research Scholarship.

I want to thank Brendan and Amruta for showing me everything I needed to know about the Department of Medicine, from operating the autoclave to decorating for Christmas. Their expertise and advice was invaluable. I want to thank Rita for helping me navigate everything and for always having her door open for a conversation. A special thank you to John for equal measures of entertainment and education. I want to thank Kevin, Dave, Bernie, Jacquie, Jill, Yvonne, Aileen, Beth, and everybody else at the Department of Medicine for everything they helped me with and for providing me with a wonderful and welcoming workplace. Of course, a special thanks to Emma, Shatabdi, and everybody else from Room 1.26 for keeping me company throughout the years.

I want to thank my parents and my sister for helping me through this project and for encouraging me during the rough patches. They're some of biggest supporters and I couldn't have done this without them.

My last 'thank you' goes to Frank. For everything.

DECLARATION

I hereby declare that all work presented in this thesis is original and entirely my own work, unless otherwise stated. This thesis has not been submitted in whole or in part for a higher degree to this or any other university. Any assistance and contribution by others to this work is acknowledged within the text.

Nicole Ellen Walsh, BSc.

ABSTRACT

The hepatitis C virus (HCV) is an enveloped +ssRNA virus, belonging to the family *Flaviviridae*. HCV is notable for displaying extraordinary genetic diversity and variability, having seven recognised genotypes and over sixty subtypes. HCV is responsible for the disease known as hepatitis C, which is associated with cirrhosis and hepatocellular carcinoma (HCC). The Global Hepatitis Report released by the World Health Organisation (WHO) in 2017 estimated that viral hepatitis was responsible for 1,340,000 deaths in 2015. The report also estimated that 71,000,000 people have ongoing HCV infections. HCV is largely transmitted *via* exposure to infected blood, with intravenous drug use accounting for approximately 55% of cases. HCV infections can be categorised as acute or chronic. During chronic HCV infections, antibodies (Abs) are produced against HCV - however, the host Abs are unable to neutralise HCV and only accelerate the evolution of circulating HCV variants. HCV variants resistant to the current generation of host Abs become the dominant variant through selective pressure. The variants of HCV within a host are known as quasispecies. Although the host Ab response is not able to resolve the chronic HCV infection, some Abs can bind to particular HCV variants. These Abs form complexes with virus particles and are known as AAVs (antibody-associated virus). AAVs are detectable in the blood of patients with chronic HCV infections and examination of these AAVs could reveal conserved viral structures and vulnerable HCV epitopes.

Twenty genotype 3a serum and plasma samples from patients with chronic HCV infections were obtained from the National Virus Registry Laboratory (NVRL) and from the Molecular Virology Research and Diagnostic Laboratory (MVDRL). HCV genotype 3a was chosen for this research project given its prevalence (estimated to account for 17.9% of chronic HCV infections), resistance to treatment, and increased risk of causing severe steatosis and HCC when compared to other genotypes. Building on previous research carried out by the MVDRL, the patient samples were screened for the presence of AAVs. Initially, AAV+ samples were going to be processed

and used to generate HCV pseudoparticles (HCVpp). The HCVpp system is a model system that incorporates the E1E2 glycoprotein from HCV into a plasmid. The E1E2 glycoprotein is responsible for HCV entry and infection, meaning the HCVpp can be used to infectivity and Ab neutralisation assays. However, the E1E2 glycoproteins could not be extracted from the AAV+ patient samples. Instead, the IgG from the AAV+ samples was extracted and used for a series of neutralisation experiments on HCV pseudoparticles generated using the HCV H77 isolate. H77 (GenBank: AAB67037.1) is an infectious genotype 1a isolate that has undergone complete genome sequencing. The Abs that showed the greatest neutralisation potential against the H77 pseudoparticles were selected for epitope mapping. The epitope mapping procedure tested the selected Ab samples against a synthesised H77 E2 glycoprotein structure, and characterised the sites where the patient Abs bound to the synthesised E2. This revealed vulnerable epitopes on the HCV E2 glycoprotein. The epitope mapping also revealed a large number of glycosylation sites around the vulnerable epitopes – a phenomenon known as glycan shielding. Glycan shielding is used by a number of viruses (including HCV and HIV) to protect conserved and vulnerable epitopes from Abs. However, strategies are being developed to counter viral glycosylation, including modifications to glycosylation sites and the use of polysaccharides derived from non-mammalian sources as therapeutic agents against glycosylated viruses.

Statement of Contribution

Samples from HCV genotype 3a patients were provided by the National Virus Registry Laboratory (NVRL), University College Dublin, by Prof. Colm Bergin (School of Medicine, TCD) and Dr. Cillian De Gascun (NVRL, UCD). Additional HCV genotype 3a samples were provided by the Pyro2 Project through the Molecular Virology Diagnostic and Research Laboratory (MVDRL), University College Cork. The Pepscan Presto B. V. peptide array was re-used from Dr. Amruta Naik's 2015 epitope mapping project (ES150122008A+B).

This work was funded by the Irish Hepatitis C Outcomes Research Network – Bristol-Myers Squibb (ICORN-BMS) Research Scholarship. The authors would like to acknowledge the contribution of the members of ICORN for their cooperation with this study, along with the employees and patients from the participating hepatology and infectious disease units

Ethical Approval and Consent to Participate

Patients were consented by Prof. Colm Bergin (MD FRCP FRCPI FIDSA Consultant Physician in Infectious Diseases, St. James' Hospital Dublin, Clinical Professor of Medicine, Trinity College Dublin), ethical approval was obtained from the St. James' Hospital/Tallaght Research Ethics Committee as part of the Irish National Hepatitis C Outcomes and Research Centre (ICORN) HCV registry.

CONTRIBUTIONS

Research Papers

1. *"Reverse Epitope Mapping of the E2 Glycoprotein in Antibody Associated Hepatitis C Virus"* Naik AS, Owsianka A, Palmer BA, O'Halloran CJ, **Walsh N**, Crosbie O, Kenny-Walsh E, Patel AH, Fanning LJ1. PLoS One, 2017 May 30

Research Presentations

1. Poster Presentation – **New Horizons in Medical Research** – December 2018, Cork, Ireland. "*Efficacy of Anti-HCV Antibodies in Reducing Pseudoparticle Entry and Infectivity".*

2. Poster Presentation – **New Horizons in Medical Research** – December 2017, Cork, Ireland. "*Isolation of Virus-Free Antigen Binding Fragments (VF-FAbs) from Antibody-Associated Hepatitis C Virus*"

1 INTRODUCTION

1.1 INTRODUCTION

The hepatitis C virus (HCV) is an enveloped +ssRNA (positive-sense (+) single-stranded (ss) ribonucleic acid (RNA)) virus, belonging to the family *Flaviviridae*. *Flaviviridae* is a family of +ssRNA viruses that includes the West Nile virus, Dengue virus, and Zika virus. The majority of *Flaviviridae* are spread through arthropod vectors. Viruses from the *Flaviviridae* family form 40 – 60nm virions, and tend to have a single core protein (except for the *Pegivirus* genus), and 2 – 3 envelope glycoproteins. Genomes are between 9 – 13kb. The *Flaviviridae* family contains four genera – *Flavivirus*, *Pestivirus, Hepacivirus*, and *Pegivirus* – and over sixty species of viruses. The *Hepacivirus* genus contains HCV, and a number of other hepaciviruses [1].

The whole HCV genome is 9,600 nucleotides, with a single opening reading frame (ORF) encompassing 9,000 nucleotides. The single ORF is surrounded by two untranslated regions (UTRs), termed the 5'-UTR and 3'-UTR. The HCV RNA genome serves as translational template for the synthesis of viral proteins. The translation of the HCV genome produces both structural and non-structural (NS) proteins – the structural proteins include the core, and the E1 and E2 glycoproteins, and the NS proteins include p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B [2]. The replication of HCV is consistent with the replication pathways of other +ssRNA viruses, with the process being cytoplasmic, and using membrane vesicles derived from the host endoplasmic reticulum (ER). The assembled HCV virions bud to the ER lumen, and are secreted through the vesicle transport pathway – see Figure 1 for an overview of the HCV lifecycle, and see Section 1.3 for a detailed description of the process [1]. HCV displays extraordinary genetic diversity, having seven recognised genotypes and >60 subtypes, with nucleotide sequences varying from 30% - 35% between genotypes, and ≤15% between subtypes. A meta-analysis of HCV genotype studies published between 1989

– 2013 estimated that genotype 1 accounts for 46.2% of HCV infections, genotype 3 accounts for 30.1%, genotype 2, genotype 4, and genotype 6 accounts for 22.8%, and genotype 5 accounts for the remaining <1% of HCV infections [4]. Genotype 7 is an emerging genotype, originating from Central Africa and accounting for a sparse number of cases [5].

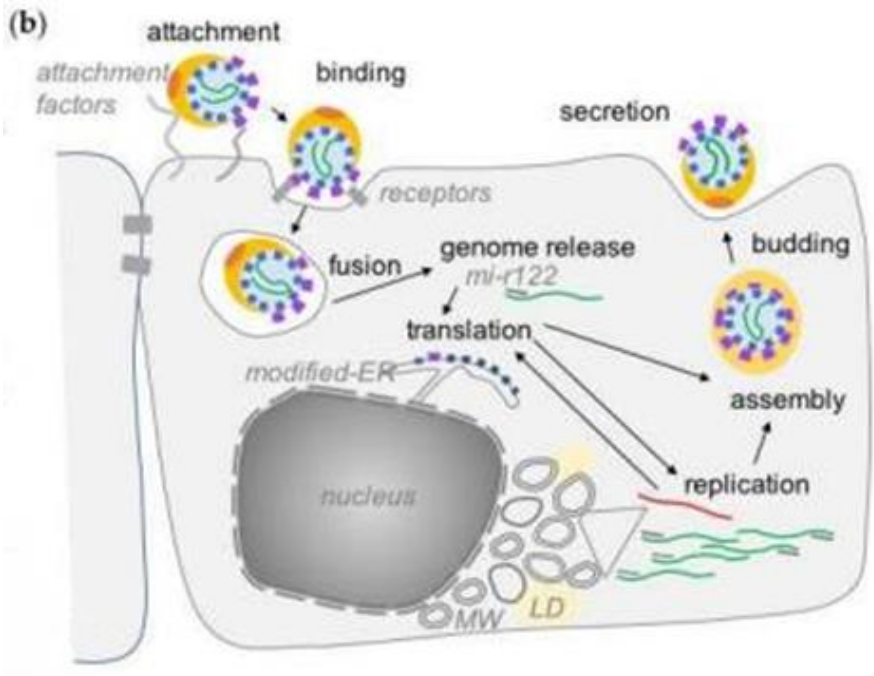


Figure 1 – A representation of the HCV lifecycle. HCV particles are bound through specific interactions with host surface receptors. Once the bound particles are internalised via endocytosis, the HCV particles undergo un-coating, and the genome is translated at the endoplasmic reticulum (ER). NS4B induces the formation of the membranous web (MW) – this acts as a framework for the virus replication complex. Once the genome has been translated and the appropriate HCV proteins expressed, progeny virions are assembled, and leave the host through the constitutive secretory pathway. "LD" indicates lipid droplets. Figure sourced from Alazard-Daly et al., 2019 [3].

HCV is transmitted *via* exposure to contaminated blood. Intravenous drug use and the transfusion of contaminated blood products and organs are two of the main transmission routes. Ireland began testing donated blood and organs for HCV around October 1991, and the Health Service Executive (HSE) recommends those that received donations before 1989 to undergo HCV testing [6]. Other routes of transmission include sexual transmission, vertical transmission, and occupational exposure. According to the European

Centre for Disease Prevention and Control (ECDC)'s 2017 hepatitis C epidemiological report, 26% of cases reported had a known route of transmission. Intravenous drug use was the most common route of transmission, accounting for 40% of acute cases and 55% of chronic cases. Nosocomial transmission (17% of acute cases) and sexual transmission between men (15% of acute cases) were the second and third most common routes of transmission, respectively [7].

HCV infections are largely categorised as acute or chronic infections. The 2017 ECDC's HCV report found that 3% of the cases reported were acute, 22% were chronic, and 75% were "unknown" [7]. The ECDC defines acute infections as the patient having a recent HCV seroconversion (with a negative test for HCV during the previous twelve months) or the presence of HCV nucleic acid or core antigen without the presence of HCV antibodies. Individuals with acute infections can be asymptomatic, but symptoms (fever, nausea, jaundice, and vomiting) can manifest among 20% - 30% of patients. The ECDC defines chronic HCV infections as the detection of HCV nucleic acid or core antigen for two samples at least twelve months apart. The potential complications arising from chronic HCV infections include hepatic fibrosis, cirrhosis, and hepatocellular carcinoma (HCC). Occult HCV infection (OCI) was first described by Castillo *et al* in January 2004, and is defined as the presence of HCV RNA in hepatocytes or peripheral blood mononuclear cells (PBMCs) without detectable serum HCV RNA [8 9].

HCV underwent numerous names and characterisations before becoming known as the hepatitis C virus. Around the 1940s and 1950s, human viral hepatitis was thought to be caused by just two etiologic agents, a hypothesis formed from studies that examined the transmission and pathology of two infectious agents associated with hepatitis. The two variations of hepatitis were termed "infectious hepatitis" and "serum hepatitis" [10]. Later, these would become known as hepatitis A (HAV) and hepatitis B (HBV). The inoculation of subjects with blood taken from suspected hepatitis patients would provide evidence that viral hepatitis was caused by transmissible agents, [11]. However, these inoculation studies were unable to explain episodes of hepatitis that were not consistent with the presence of HAV or HBV. These

inconsistencies led to suggestions that there could be a third agent capable of causing hepatitis [12].

Around the 1970s a number of serological tests were developed for the detection of HAV and HBV - namely a radioimmunoassay capable of detecting the hepatitis B surface antigen (HB-sAg). Testing donated blood products for HB-sAg reduced the number of cases of HBV infections, but studies found that 90% of transfusion hepatitis infectious were not related to HBV or HAV. This furthered suggestions that these cases of hepatitis were being caused by an infectious agent that had not been characterised [13] [14]. In 1975, Alter *et al.* proposed the name “non-A, non-B” hepatitis for this mysterious infectious agent, observing that eight surgery patients had developed hepatitis from blood products that were negative for HB-sAg [15]. A hypothesis emerged that there could be multiple agents capable of causing NANBH, and that these agents were related to HBV, but induced a separate immune response [16] [17].

In 1989 Choo *et* al. generated cDNA from the plasma of a patient diagnosed with NANBH, and isolated a clone generated from an RNA molecule present in NANBH infections. The clone consisted of about 10,000 nucleotides and was found to be positive stranded [18]. The infectious agent was found to share features with *Togaviridae* and *Flaviviridae* – both members of Group IV of the Baltimore classification of viruses. The Baltimore classification categorises viruses into seven groups based on their genome and replication strategy, with Group IV comprising of +ssRNA viruses. Through a combination of genomic and taxonomic similarities, and the medical presentation of NANBH infection, NANBH became known as the hepatitis C virus. Figure 2 summarises the timeline of HCV, from the discovery of NANBH to the manufacturing and FDA ((United States) Food and Drug Administration) approval of modern therapeutics and treatments for chronic hepatitis C infection.

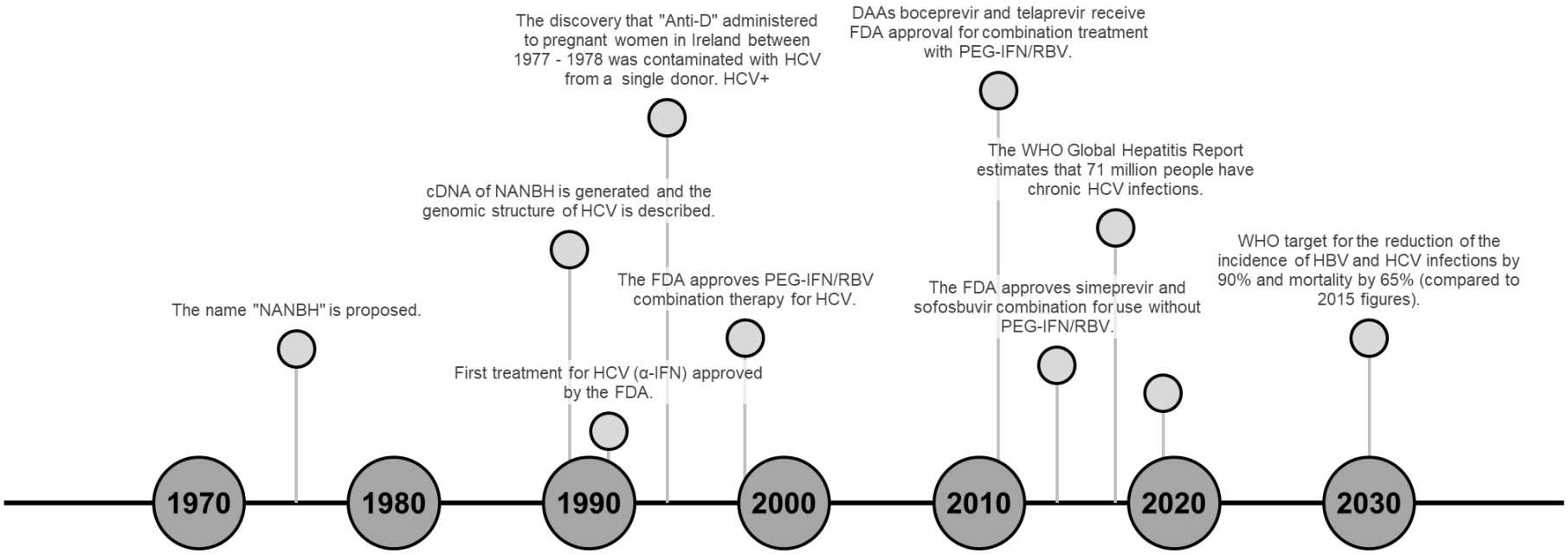


Figure 2 – A graphic showing significant events associated with the hepatitis C virus.

1.1.1 Epidemiology

Hepatitis C – the disease caused by infection with HCV - remains a major disease burden, with previous reports from the World Health Organisation (WHO) estimating that there are 130 – 150m people with chronic HCV infections. The WHO had estimated that 700,000 deaths, per annum, could be contributed to complications from HCV infection, but the release of the Global Hepatitis Report 2017 has revised a number of these statistics [19]. The Global Hepatitis Report 2017 found that viral hepatitis was responsible for 1.34m deaths during 2015 - 720,000 deaths were from cirrhosis and related conditions, and 470,000 deaths were caused by HCC. Estimates of the current number of individuals with chronic HCV infections have been revised from 130 – 150m people to 71m [20]. This decrease can be explained by changes to the methodologies used to evaluate the number of HCV cases. Previously, estimates focused on the presence of HCV antibodies as an indicator of HCV infection, but the mere presence of HCV antibodies is not enough to distinguish between previous and current infections. Instead, the presence of HCV RNA is used to determine the presence of a chronic HCV infection [21]. As per Section 1, the ECDC stated that 31,273 cases of hepatitis C were reported in 29 EU/EEA Member States in 2017 [7].

HCV genotypes have distinct regional groupings - for example, a recent epidemiology study of HCV genotype distribution across Europe found that genotype 2 infections were more prevalent in Western Europe (8.9%), whereas genotype 4 was more prevalent for Central Europe (4.9% for genotype 4 versus 3.2% for genotype 2) [22]. There are several population groups that are vulnerable to HCV infection. The CDC recommends that those born between 1945 – 1965, those who received blood transfusions and/or organ transplants before 1992, intravenous drug users (IDUs), HIV+ individuals, those who have had occupational HCV exposure, and children born to HCV+ mothers should be tested for HCV [23]. Exposure to HCV through blood transfusions and organ transplantations was one of the major routes of transmission, but the widespread screening of blood and organs

(from 1992 onwards in Ireland), has hugely decreased the number of cases – however, this problem does persist for developing countries [24].

There have been several incidents involving the iatrogenic transmission of HCV. The improper sterilisation of equipment, combined with the lack of knowledge about HCV, caused the extensive transmission of HCV genotype 4 during a vaccination campaign to combat schistosomiasis in Egypt during the 1950s – 1980s. A number of epidemiological studies have estimated that between 7.3% - 14.7% of the Egyptian population are HCV+, making Egypt the country with the highest prevalence of HCV [25]. During February 1994, the Irish Health Boards (now the Health Service Executive (HSE)) discovered that anti-D immunoglobulin used from 1977 to 1978 had been contaminated with HCV from a single HCV+ donor. Anti-D immunoglobulin is administered to pregnant women to prevent Rh (Rhesus) isoimmunisation and is derived from human plasma. In Ireland, the routine screening of blood donors for HCV antibodies started in October 1991, with a regional study of donors from 1991 to 1994 identifying 14 men and 15 women with HCV antibodies. The women did not match the typical donor demographic (older age and mostly Rh-negative) and it was found that 12 of these women had received anti-D immunoglobulin in 1997. Further investigations uncovered the HCV genotype 1b contamination of anti-D immunoglobulin distributed in 1997 and 1998, which was confirmed through reverse transcription PCR (polymerase chain reaction) of stored samples from the suspected infected donor. A chronology of the infected donor and contaminated plasma was established – from October 1976 to January 1977, plasma from the donor was used to prepare five batches of anti-D immunoglobulin, but in November 1976 the donor had an adverse reaction to a transfusion, became jaundiced, and was diagnosed with a hepatitis C infection. Additional samples of plasma were obtained from the donor in January 1977 and used to prepare more batches of anti-D immunoglobulin. In total, 16 batches of anti-D immunoglobulin were prepared from the infected donor in 1977, with eight of those batches testing positive for HCV RNA [26]. In March 1994 62,667 women who had received anti-D immunoglobulin between 1970 and 1994 were screened for HCV, and 704 women showed evidence for previous or ongoing HCV infections. 390 of the

704 women were positive for serum HCV RNA and were referred for further assessment and treatment. 376 women accepted the treatment, and biopsies showed hepatic inflammation for 98%, with moderate hepatic inflammation for 52%. 51% showed evidence of fibrosis, and 2% had confirmed cases of cirrhosis. The Hepatitis C Compensation Tribunal paid €310 million to victims by 1998, with the 2018 cost of the incident estimated to be about €1.5 billion for the compensation and legal costs of >4,500 claims [6] [27] [28].

IDUs represent a major risk group for contracting and spreading HCV. A study surveying HCV prevalence among IDUs across Asia suggested that HCV prevalence rates range from 41% of IDUs in Taiwan, to 74% of IDUs in Vietnam [29]. Also, because IDUs can become re-infected despite undergoing treatment and achieving a sustained virological response, several economic evaluations have been carried about the prioritisation of HCV treatment for IDUs. A study conducted by Martin *et al.* using a mathematical model concluded that treating infected IDUs with moderate and mild fibrosis was more economical than delaying treatment to the development of cirrhosis. The study estimated that treating a single infected IDU with DAAs (direct-acting antiviral) could prevent two HCV infections, for scenarios where the prevalence of HCV is 20% among the IDU population. However, for scenarios with higher HCV prevalence among IDUs - 40% and 60% - the number of prevented infections dropped to 0.78 and 0.23, respectively, because of the increased rate of secondary HCV infections – these predictions are graphed in Figure 3 below[30].

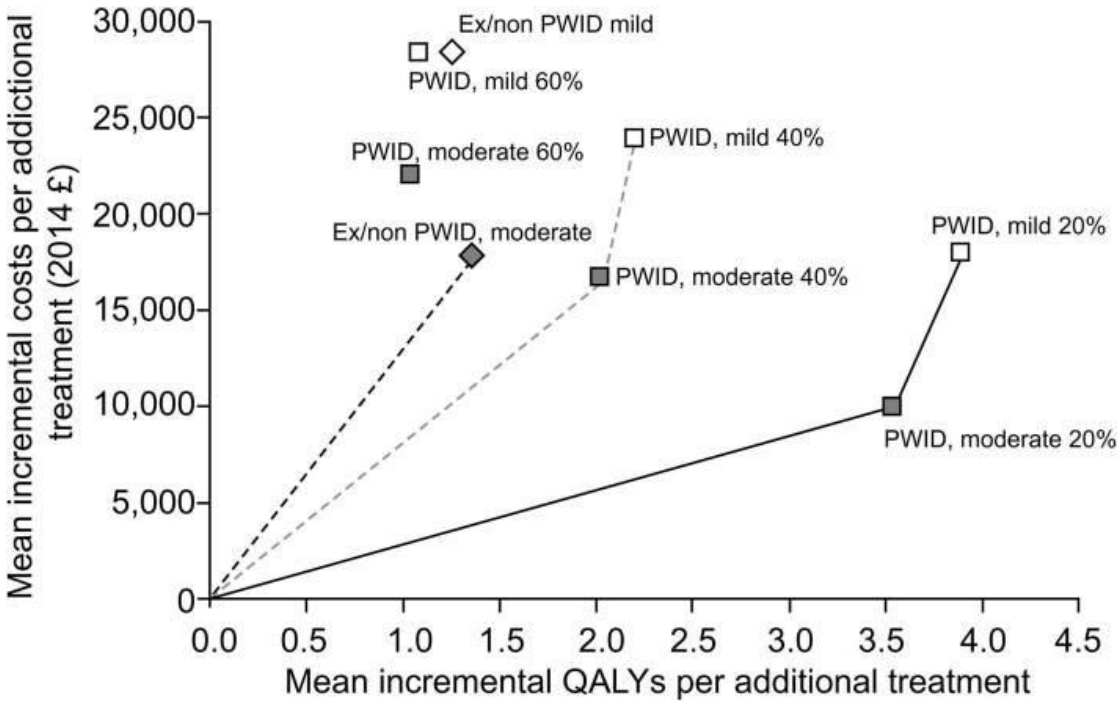


Figure 3 – A graph demonstrating the relationship between money spent on HCV treatment and the increase of QALYs per treatment. The most beneficial and economic treatment option is the treatment of IDUs with moderate disease amongst an IDU population with 20% HCV infection. Figure sourced from Martin et al.*, 2016* [30]*.*

In 2008 the Scottish government launched its Hepatitis C Action Plan with the goal of improving access to screening and treatment services and to prevent further HCV transmission - particularly among PWID (people who inject drugs). In Scotland, 85% of those infected with HCV were infected from injecting drugs [31]. The Action Plan was funded with £100m and utilised local and national monitoring services. It also used predictive models to estimate treatment outcomes and set government targets. The introduction of dried blood testing helped drive a 50% increase in the proportion of the infected population diagnosed (from 38% before 2008 to 55% in 2015), and a large increase in the numbers of people introduced to treatment for HCV (470 to 1,050 in 2015). There was a particular increase of treatment uptake among PWID, from 300 before the start of the Action Plan, to 840 as of 2015 [32]. A further study has been planned for 2019 onwards to evaluate the impact of DAA treatment for PWID involving an intervention site in Tayside, Scotland. In the Tayside site, HCV treatments have been expanded to include specialist treatment centres and needle and syringe exchange programmes (NSPs). From 2017 - 2018, 500 PWID were treated at the Tayside centre, and simulations have predicted that HCV prevalence among PWID in the

Tayside area should be reduced from 26% to 10%. Treatment responses and re-infection rates are going to be monitored and compared to control sites in Scotland and England to evaluate the success of the programme [33].

The worldwide prevalence of HCV infection among children is estimated to be 0.05% - 0.4% for developed countries, and 2% - 5% for developing countries, with vertical transmission and iatrogenic infection being the primary causes of paediatric HCV infections – however, intravenous drug use is a noted transmission route for adolescents in developed countries [34]. Vertical transmission refers to the passage of pathogens or infectious agents through the placenta, through direct contact during the birthing process, or through breastfeeding. The United States has an estimated 23,000 – 46,000 paediatric cases of HCV, with 34 iatrogenic HCV cases between January 2008 and April 2013 [35]. The vertical transmission rate is estimated to be 2% - 8%, with the risk factors being increased maternal viral load, intrapartum invasive procedures (foetal scalp blood sampling, etc.) and episiotomy procedures. Vertical transmission rates are increased for mothers with HIV/HCV co-infections, with the transmission rates estimated to be 15% - 20% [36] [37].

The prevalence of HCV is higher amongst the prison population than the general population. In a 2018 study four prisoner groups (n = 46) from two prisons in Dublin were interviewed and asked about their knowledge of hepatitis C and barriers they had experienced with accessing HCV screening and treatment [38]. IDUs are over-represented in the prison population compared to the wider community, with 26% of prisoners reporting having used intravenous drugs. A 2014 study of Irish prisoners found that HCV prevalence in Irish prisons was estimated to be 13% [39]. A major barrier was lack of knowledge - it was found that the participants were aware of HCV, but were also aware of their ignorance about the modes of HCV transmission and the differences between HBV and HCV. Other barriers included the fear of invasive treatments and biopsies, and the delayed and bureaucratic process of screening and treatment - even when the participants were enthusiastic about engaging with the services. The participants also mentioned a lack of confidentiality between prisoners and wardens, and

stigmatisation by other prisoners and the wider community when going to out-patient hospital appointments. The participants suggested an “opt-out” screening upon arrival as one method of improving access. The participants praised the fibroscanning (measurement of fibrosis and steatosis using ultrasound technology) procedure, stating that it was non-invasive and easy to access. Another major facilitator to HCV screening and treatment, was the removal of external stressors (eg. lack of accommodation and money) when incarcerated - once these needs were met by the prison service, participants were able to focus on treatment. DAA treatment has been found to be cost effective when used in the prison system and the prison system represents a valuable chance to treat and educate a large population of people with HCV [40].

1.1.2 Diagnosis

The diagnosis of HCV infections is done through the detection of HCV antibodies and HCV RNA. If HCV antibodies are detected during initial examinations, HCV infection can be confirmed using a nucleic acid test (NAT) to detect the HCV RNA from the provided serum or plasma sample. A NAT is necessary because the presence of HCV antibodies could indicate an ongoing *or* a resolved infection, whereas the presence of HCV RNA is associated with a current infection. HCV antibodies are detectable 4 - 10 weeks post-exposure, and HCV RNA becomes detectable at 2 - 3 weeks post-exposure [41]. A surrogate marker known as a sustained virologic response (SVR) is used to evaluate the effectiveness of treatments, and refers to having no detectable HCV RNA 12 or 24 weeks (SVR_{12} or SVR_{24}) post-treatment completion [42]. However, many modern DAA treatment regimens can produce a SVR before 12 weeks – for example, Ohya *et al.* saw a SVR achieved in three paediatric patients at just four weeks of DAA treatment [43]. The HSE website for the National Hepatitis C Treatment Programme states that most treatment regimens take between 8 – 12 weeks, with only a minority needing 24 weeks of treatment [44]. HCV diagnosis is largely carried out by the detection of HCV RNA from serum or plasma samples, but recent advancements have seen the development of screening using dried blood spot (DBS) samples. DBS samples are quicker and easier

to perform than a traditional serum/plasma sample, and are convenient for locations with lower resources and/or high numbers of potential HCV patients [45]. Additionally, a number of quick and automated serological assays have been developed for the detection of hepatitis C core antigen (HCVAg) instead of the detection of HCV RNA. HCV RNA and HCVAg become detectable at the same stage of HCV infection, and both the WHO and EASL (European Association for the Study of the Liver) has declared HCV RNA and HCVAg assays as comparable for the diagnosis of HCV infection. Instead of screening for anti-HCV antibodies and then screening for HCV RNA, Galli *et al*. proposes that solely screening for HCVAg could be used to diagnose HCV infection. Galli notes that for chronic untreated HCV infections, HCV RNA has been observed to be detectable when <0.5% - 7% of HCVAg tests have shown target-not-detected (TND) depending on the method used – but this limitation could be balanced by the greater stability of HCVAg in biological samples, making it a convenient strategy for locations where samples are taken at point-of-care (POC) and sent to a central laboratory for processing [46].

1.1.3 Treatment

Historically, HCV infection was treated with a combination of ribavirin (RBV) and pegylated interferon-α (PEG-IFNα), over 24 or 48 weeks. Ribavirin is ribonucleic analogue and nucleoside inhibitor used to inhibit vRNA synthesis and mRNA capping. It can mimic guanosine or adenosine, causing mutations upon vRNA incorporation and incapacitating the virus. RBV has also been found to promote p53 activity through the ERK1/2 pathway [47]. α-interferon, given as subcutaneous injections, was one of the first immunomodulatory agents used to treat HCV infection, and was approved as a treatment for NANBH in 1986 [48] [49]. Interferons (IFNs) belong to the cytokine group of proteins and are produced as a response to the presence of pathogens, upregulating and activating numerous immunological pathways to combat infection. IFN-α is largely produced by plasmacytoid dendritic cells (pDCs). Interferon Alfa is the International Non-Proprietary Name for the pharmaceutical drug composed of natural IFN-α produced from human leukocytes and is delivered through subcutaneous injection. Pegylated IFN

(PEG-IFN) refers to polyethylene glycol polymer chains that are added to the interferon molecule to improve various pharmacological properties, such as reducing the immunogenicity of IFN and improving circulatory time [50].

However, there were several problems associated with PEG-IFN/RBV combination therapy. SVR rates remained suboptimal at 54-56%, and numerous contraindications to RBV prevented a significant percentage of patients from receiving the combination therapy. These contraindications included pregnancy, asthma, diabetes, thyroid disease, pulmonary diseases, and various autoimmune disorders. > 50% of those with HCV genotype 1 infections had positive treatment outcomes after a course of the combination therapy. PEG-IFN became notorious for adverse events, with fever, headache, nausea, insomnia, alopecia, and musculoskeletal pain being common. There was also a risk of patients developing depression and psychosis, hepatic complications, and thrombi. It became apparent that numerous patients are unable to receive, tolerate, or complete PEG-IFN/RBV combination therapy, and that treatment regimens that did not require interferon-α needed to be developed [51].

Direct-acting antivirals (DAAs) represented a significant treatment advancement for those with chronic HCV infections. Once *in vitro* models for studying HCV were developed and the HCV genome could be characterised, a number of HCV non-structural proteins were identified as potential therapeutic targets. There are four categories of DAAs based on their mechanism of action and therapeutic targets – NS3/4A protease inhibitors (PIs), NS5B nucleoside polymerase inhibitors (NPIs), NS5B non-NPIs, and NS5A inhibitors [52].

2002 saw the development of the first NS3 PI (BILN 2061, CILUPREVIR®), which showed promise at the start of the testing period. However, a number of adverse outcomes prevented further development [53]. 2011 saw the approval of two oral DAA treatments – boceprevir and telaprevir. These protease inhibitor DAAs were combined with peg-IFN-α and RBV to form a “triple therapy” for patients with genotype 1 infections, with SVR rates increasing to 75% - nearly double the SVR rate for the PEG-IFN/RBV

combination treatment [54]. However, there were several problems associated with this “triple therapy”. The treatment was expensive, demanded strict dietary requirements, and patients needed to take 9 – 18 tablets per day [55]. In 2013, simeprevir (SMV) and sofosbuvir (SOF) were introduced to the market, becoming the first DAA treatments that only required a single tablet per day. Simeprevir and sofosbuvir were tolerable and produced SVR rates of >90% when combined, or with simeprevir/sofosbuvir and PEG-IFN/RBV. Simeprevir is a NS3/4A PI that was commonly used for genotype 1 and genotype 4 infections but demonstrated some effectiveness against genotype 2 and genotype 6. Sofosbuvir is a nucleotide analogue that targets the NS5B polymerase and was approved by the FDA for the treatment of genotype 1, 2, 3, and 4 infections when combined with RBV – PEG-IFN was optional depending on the genotype and patient. The release of simeprevir and sofosbuvir marked the advancement towards treatment regimens that no longer need PEG-IFN [56].

At present, many therapeutic regimens are based on the combination of two or more DAAs, depending on the genotype and subtype of the HCV infection and the patient’s hepatic condition. For example, in 2019 the American Association for the Study of Liver Diseases (AASLD) has issued a simple algorithm for treating treatment-naïve patients that do not have cirrhosis which recommends glecaprevir (300mg)/pibrentasvir (120mg) taken for eight weeks in conjunction with sofosbuvir (400mg)/velpatasvir (100mg) taken for twelve weeks [57]. RBV is not commonly used in treatment regimens, but there are circumstances where RBV may be added to a course of DAA treatments – for example, the 2018 EASL HCV treatment recommendations states that patients with decompensated cirrhosis without HCC and who are awaiting liver transplantation be given RBV based on their body weight for 12 weeks. Modern HCV treatment regimens are estimated to achieve >95% SVR, and represent a huge improvement when compared to the RBV and PEG-IFN-α combination treatment [55].

Despite the vast improvements DAAs have made to HCV treatment, there are several outstanding problems. A notable criticism of DAAs has been the cost. For example, sofosbuvir was launched at US$1,000/tablet. The

development of generic DAAs has improved access to DAA treatments, but the WHO emphasises that access must improve further to have a meaningful impact on the global burden of HCV [21]. According to the Comptroller and Auditor General 2017 report, the cost of hepatitis C treatment in Ireland ranges from €23,000 - €92,000 – however the HSE's National Hepatitis C Treatment Programme provides DAA treatment for free for patients [28] [44].

The considerable cost of DAA treatments has necessitated the prioritisation of particular patients. As described during Section 1.1.1, the treatment of chronic HCV among individual IDUs can reduce the incidence of HCV among the broader IDU population – but this is balanced by the risk that the treated individuals could relapse and become re-infected with HCV without proper support and treatment for drug dependencies. The 2016 European Association for the Study of the Liver (EASL) recommendations state that treatment should immediately be considered for patients with significant fibrosis (measured using the METAVIR system for assessing inflammation and fibrosis, with fibrosis being defined as a METAVIR score >F2) or cirrhosis (METAVIR score >F4), patients with significant extra-hepatic complications, patients with HCV reoccurrence post-transplantation, and patients that risk transmitting HCV to other individuals (IDUs, MSM (men-who-have-sex-with-men), incarcerated individuals, and women who wish to become pregnant). IDUs and MSM should be encouraged to use preventative measures alongside DAA treatment. The recommendations also state that treatment is not advised for patients "with limited life expectancy due to non-liver-related co-morbidities) [58].

The EASL recommendations also compare and contrast drug combinations for various HCV genotypes. For example, for genotype 3 infections, sofosbuvir + ribavirin is considered sub-optimal, and instead sofosbuvir + velpatasvir or sofosbuvir + daclatasvir is recommended. Ribavirin can be added to the above combinations, and the combination of peg-IFN/RBV + sofosbuvir remains acceptable when latest-generation DAAs are not accessible. Lepidasvir is not recommended for genotype 3 [58].

Another consideration for DAAs is that several resistance-associated variants (RAVs) have been observed *in vitro* and during treatment. A 2016 study examined GenBank HCV sequences and found that 58.7% of sequences examined had one dominant RAV. The highest RAV frequency was observed for genotype 6 sequences, (99%) with 50% of genotype 3 sequences displaying a known RAV. 40% of the sequences examined had RAVs associated with NS5A inhibitors, and 29.6% had RAVs associated with NS3 protease inhibitors. That said, RAVs associated with NS5B inhibitors and nucleotide inhibitors (NI) were only detected among <4% of the sequences, supporting the suggestion that regimens using combinations of NI-based DAAs had the lowest prevalence of RAVs [59]. The 2019 AASLD recommends NS5A RAV screening for genotype 3a patients before treatment commences, and states that those without the Y93H RAV can be treated with 12 weeks of sofosbuvir/velpatasvir, but patients with the Y93H RAV should be treated with glecaprevir/pibrentasvir or an alternative regimen [57].

Genotype 3 infections show greater resistance to DAA treatments when compared to non-genotype 3 infections. During the period of the PEG-IFN-α/RBV combination treatment, genotype 3 was considered a treatable genotype, and ~70% of genotype 3a patients achieved SVR. However, genotype 3 has proven to be resistant to PIs and SOF. This poses a significant problem, as SOF forms the foundation of many treatment strategies, and was considered a pan-genotypic drug. SOF was thought to have a “genetic barrier” to drug resistance, but Ramirez *et al.* proved that substantial SOF resistance could be induced using a JFH1-based genotype 3a model system. The RAVs demonstrated increased HCV persistence, and increased DAA resistance, including resistance to nucleoside analogues and non-nucleoside analogues. [60]. A study conducted by Lawitz *et al.* demonstrated that genotype 3 infections treated with a SOF-RBV combination had a lower response rate compared to genotype 2 infections – 56% versus 97% [61]. Genotype 3a is associated with frequent and severe steatosis, increased fibrosis progression rates, and elevated incidences of HCC – factors that could exacerbate drug resistance [62]. Genotype 3 is

estimated to account for 10% – 15% of chronic HCV infections, and is prevalent across European and IDU populations – therefore, the development of effective treatment strategies for genotype 3 infections is crucial.

One potential strategy for the effective prevention and eradication of HCV is the development of a HCV vaccine – this is discussed in further detail in Section 1.5.

1.2 HCV Genome and Proteins

The HCV genome consists of approximately 9,600 nucleotides and encodes a polyprotein that is cleaved by virus and host proteases to produce three structural proteins (the core protein and the envelope glycoproteins E1 and E2) and seven non-structural proteins (p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B). The 5'-UTR and 3'-UTR are located at the N-terminal and C-terminal of the genome – a genomic conformation common among members of the *Flaviviridae* family.

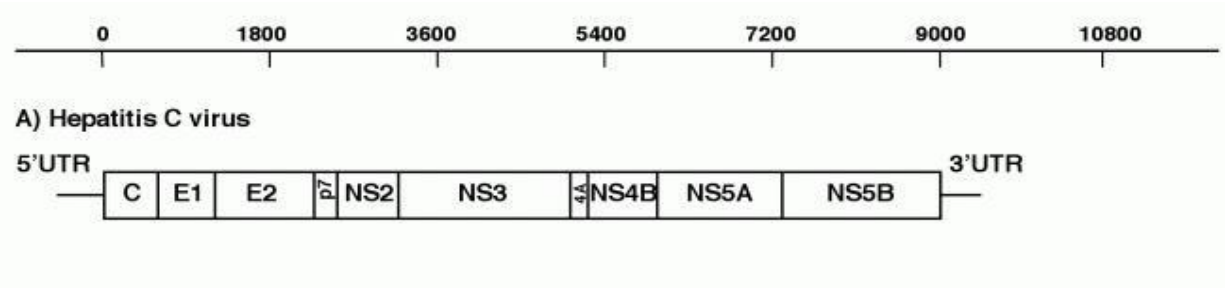


Figure 4 – The organisation of the HCV genome, demonstrating the position of the structural and non-structural proteins. The scale at the top of the figure indicates the number of nucleotides [63].

1.2.1 Untranslated Regions

The 5'-UTR and 3'-UTR are regions of vRNA that do not code for proteins but do have structural elements and sequences important for HCV translation and replication. The 5'-UTR is located at the N-terminal of the genome, and the 3'-UTR is located at the C-terminal.

The 5'-UTR consists of 341 nucleotides located upstream of the translation initiation site of the ORF. The 5'-UTR is one of the most conserved regions of the HCV genome, showing ~90% nucleotide sequence identity among genotypes, and sharing ~60% sequence identity with GBV-B. GBV-B is a

member of the *Flaviviridae* family and was previously known as the hepatitis G virus (HGV) [63] [64]. The 5'-UTR has four structural domains – Domain I, II, III, and IV. These domains have a number of conserved structural features, such as pseudoknots and stem-loops. The initial 40 nucleotides constitute Domain I and are needed for RNA replication. The remaining 5'-UTR domains and the start of the core protein constitute an internal ribosome entry site (IRES), which directs the cap-independent translation of HCV RNA [65]. Cap-independent translation refers to the fact that eukaryotic mRNAs have a 5' cap structure to bind translation initiation factors, whereas virus mRNAs tend not to have a 5' cap. Instead, these viruses using cap-independent translation elements to initiate translation, and recruit translation initiation factors and ribosomal subunits to vRNA.

Domain II and III contain multiple stem-loop structures, which are conserved across HCV genotypes and subtypes. The base of Domain III forms a structural pseudoknot – this pseudoknot is important for IRES activity, and forms part of the binding site for the 40S ribosome subunit. Domain IV has a stem-loop structure that houses a translation initiation codon, but this stem-loop structure is not needed for the internal entry of ribosomes. In fact, one paper found that the stability of this structure was *negatively* correlated with the translation of HCV RNA, and proposed that the Domain IV stem-loop structure could instead regulate translation through a feedback mechanism [66]. The HCV IRES forms a stable pre-initiation complex through binding the 40S ribosomal subunit without the need for typical translation initiation factors – this event probably constitutes the first stage of HCV polyprotein translation [63].

The 3'-UTR contains 200 – 235 nucleotides and contains three distinct regions – a non-conserved variable region of 30 – 40 nucleotides, a polyuridine (polyU/UC) tract of 20 – 200 nucleotides, and a conserved region that forms three stem-loop structures and contains 98 nucleotides (known as the "X" region) [65] [67]. The variable region is located next to the termination codon of the HCV polyprotein. The variable region sequence composition varies between genotypes but is conserved among HCV isolates of the same genotype. The HCV genome does not have a poly(A) tail (a feature of

eukaryotic mRNA needed for production of mature mRNA) but uses the poly(U/C) tract. The 3'-UTR is thought to be important for the *in vivo* infectivity of HCV. A "knockout" study of the 3'-UTR using infectious cDNA clones, deletion mutants, and the chimpanzee model found that HCV mutants lacking the poly(U/C) region and the "X" region were unable to infect the chimpanzee, implying that these two regions are critical for infection [68].

1.2.2 Core Protein

The mature form of the core protein is a 21-kDa protein that forms the HCV capsid. The capsid protects the genomic RNA as the virus moves throughout the host and can modulate several host pathways. The core protein has a α-helical conformation for roughly 50% of the protein and can have two conformations based on the presence or absence of detergents – with detergents the protein assembles as dimers, and lacking detergents the core protein forms soluble aggregates. The core protein has two domains – D1 and D2 – which are separated by their amino acid compositions and hydrophobicity. The D1 domain is located at the N-terminal and is hydrophilic, whereas the D2 domain encompasses the C-terminal and has hydrophobic properties. D1 can be further divided into BD1, BD2, and BD3, (BD = basic domain), and D1 is thought to be involved with the oligomerisation needed for particle formation and RNA binding. The D2 domain contains two α-helices, and is responsible for the core's association with the endoplasmic reticulum (ER) and lipid droplet (LD) membranes [69 70]. The synthesis of the core protein involves a number of stages. First, the core protein is cleaved at position 191 by host signal peptidase (SP) to separate the protein from the precursor polyprotein. The immature form of the core protein is cleaved by host SP peptidase (SPP) at the C-terminal transmembrane region – this generates the mature form of the core protein. The cleavage of the HCV core protein by host signal peptidases is essential for proper HCV assembly [71]. The HCV core protein has a major impact on HCV pathogenesis, and has been shown to impact transcription regulation and signal transduction [72].

1.2.3 E1 and E2 Glycoproteins

The HCV genome encodes two envelope glycoproteins – E1 and E2. These HCV glycoproteins contain N-terminal ectodomains and C-terminal hydrophobic transmembrane regions. E1 and E2 are Type I transmembrane proteins. Once cleaved by host signal peptidases from the precursor polyprotein, the ectodomains are targeted to the ER lumen, and the transmembrane domains are inserted into the ER membrane. The C-terminal of the immature core protein contains a signal sequence that assists the translocation of the E1 ectodomain to the ER lumen. Once the HCV glycoproteins have been synthesised and translocated, the glycoprotein transmembrane domains form a non-covalent heterodimer – E1E2. E1E2 is present on the surface of the HCV particle, and can be considered as the functional subunit of the HCV envelope, facilitating HCV entry and infectivity through interaction with surface molecules and receptors [73].

The ectodomains of the glycoproteins are glycosylated - a property that impacts the folding and entry functions of E1 and E2, and can modulate host responses. E1 and E2 demonstrate N-linked glycosylation, which occurs when oligosaccharides are transferred from lipid intermediates to Asn residues within Asn-X-Thr/Ser sequences, where the X can be any amino acid except proline. E1 has six glycosylation sites, and E2 has eleven. A number of these N-glycosylation sites are conserved across HCV genotypes [74 75]. The physical and chemical structure of the glycoproteins has been hard to characterise. Separating E2 from the E1E2 heterodimer causes E1 to form misfolding aggregates, complicating structural analysis. Incomplete crystal structures of the E1 and E2 ectodomains have been proposed, and a number of research groups have used computational models to explore the unresolved sections. A number of disulphide bridges across the E1 and E2 glycoproteins have been discovered, and a “heterohexamer/trimer” model was proposed by Falson *et al*. that supported by other groups and immunological and computational data [76 77].

The sequence variation among the E1 and E2 glycoproteins can exceed 37% between isolates. Because the glycoproteins are used for HCV entry and infection, this genetic heterogeneity poses a major obstacle to the

development of preventative treatments. The extreme sequence variation can be attributed to the presence of three hypervariable regions (HVRs) on the E2 glycoprotein. HVR1 displays the greatest sequence diversity across the entire genome and consists of a 26 – 27 aa region at the N-terminus of E2. HVR2 and the inter-genotypic variable region (IgVR) are located proximal to the CD81-binging region and the transmembrane domain of E2 respectively [78].

1.2.4 Glycosylation of E1 and E2

The ectodomains of E1 and E2 are highly glycosylated, with E1 having five conserved glycosylation sites and E2 having eleven glycosylation sites conserved across genotypes. Glycans account for a third of the molecular mass of the E1E2 heterodimer. Despite the large genetic diversity between genotypes and subtypes, the *N*-glycosylation sites tend to be conserved across genotypes and have been shown to modulate several functions of the envelope glycoproteins. *N*-linked glycosylation is the attachment of an oligosaccharide (a carbohydrate consisting of a number of sugar molecules and known as a glycan) to a nitrogen atom. HCV particles are assembled in the ER, meaning that they must cross the secretory pathway before being released. As the HCV particle crosses the secretory pathway, the HCV glycoproteins are modified by glycosyl transferases and glycosidases. The first evidence of glycan modifications being related to the secretion of HCV particles came from the HCV pseudoparticle (HCVpp) and HCV cell culture (HCVcc) systems, with HCVpp E1E2 glycoproteins displaying complex glycans and HCVcc E1E2 glycoproteins displaying high mannose and complex glycans. The presence of complex glycans is associated with the movement of protein through the Golgi apparatus. HCVcc particles display immature glycans (implying that some glycans are not accessible to Golgi enzymes) and HCVpp particles display mature glycans. This corresponds to the fact that HCVcc articles assemble near the ER, whereas HCVpp particles assemble post-Golgi apparatus. [79] [80]

Glycosylation is crucial for the proper assembly and release of HCV. For example, site-directed mutagenesis (SDM) was performed by Helle *et al*. to examine the roles of the E2 glycans – see Section 3.5.5 for further

information. It was found that E2N8 was important for E1E2 folding and heterodimerisation, with E2N8 knockouts have poor infectivity. E2N6 knockouts had the same infectivity and core release as the wildtype HCV, but were more sensitive to antibody neutralisation – see Table 1 for the breakdown of how glycan knockouts impacted HCVcc/HCVpp infectivity, core protein release, and sensitivity to antibody neutralisation.

Virus/Glycan Mutant	HCVcc Infectivity	HCVpp Infectivity	Core Release	Sensitivity to Neutralisation
Wild Type	+++	+++	++	+
E2N6	+++	++	++	++
E2N7	+/-	+++(+)	+	ND
E2N8	-	-	+/-	ND

Table 1 – Findings by Helle et al. regarding the importance of E2 glycans using SDM techniques. Percentage of infectivity: +++ = >90%, ++ = 30% - 90%, + = 10% - 30%, +/- = 2% - 10%. Percentage of core release: ++ = >75%, + = 30% - 75%, +/- = 12% - 30%. Sensitivity to antibody neutralisation: + = similar to wild type, ++ = >5x more sensitivity to antibodies tested. ND = not determined.[81]

Besides serving a number of functions for the assembly of the virus, glycosylation of the HCV envelope glycoproteins protects epitopes from neutralising antibodies (nAbs). Known as “glycan shielding”, this strategy is used by several viruses to prevent nAbs from accessing sensitive epitopes. For HCV, the virus uses the host to synthesise the glycoproteins, meaning that the viral glycoproteins are recognised as “self” structures by the immune system. The physical and biochemical structure of the glycans also “block” access to the vulnerable antigens, providing a two-pronged defence.

1.2.5 p7

The HCV p7 protein is 63aa residue located between the E2 glycoprotein and the NS2 protein. p7 is Class IIA viroporin, meaning that p7 has two transmembrane regions that are separated by a cytoplasmic loop, and has both its N-terminal and C-terminal orientated towards the ER lumen. The transmembrane α-helices are known as TM1 and TM2. Because p7 displays high hydrophobicity and forms oligomers, it has been proposed that p7 could act as a pH-activated ion channel for HCV-infected hepatocytes. The equilibrated proton gradient on the membranes of acidic organelles causes

alkalinisation of those organelles, protecting the HCV glycoproteins from mis-folding and inactivation, and enabling the final stages of HCV assembly and escape [82] [83]. Experiments using mutational analysis have shown that p7 is needed for host infection [84]. p7 is also involved with the process of capsid assembly and envelopment [85].

1.2.6 NS2 and NS3

The HCV NS2/3 protein is a hydrophobic protease responsible for the cleavage of the HCV polypeptide between NS2 and NS3. Extending from positions 810 – 1206, the NS2/3 protease is the first protease needed for HCV polyprotein processing, and the first NS protein that undergoes translation. The N-terminal of NS2 is cleaved from p7 using host signal peptidases, and the serine protease located at the C-terminus of NS3 is responsible for cleavage at the NS3/4A junction [86]. HCV mutants without NS2/3 proteases were unable to cause persistent infections on the chimpanzee model, indicating that the NS2/3 protease is needed for HCV replication and persistence [87]. Once cleavage at the NS2 and NS3 junction has occurred, the cleaved NS2 protein is moved to the ER membrane. From there, NS2 has been proposed to modulate gene expression and apoptosis, and contribute to the phosphorylation of HCV NS5A [88].

NS3 protease activity depends on the presence of zinc ions – the ions are needed for hydrolytic activity and contribute to the structural integrity of the NS3 protein. Additionally, NS3 cooperates with NS4A to induce conformational changes that enhance hydrolytic activity [89]. These hydrolytic and enzymatic changes enable NS3 to mediate the proteolytic release of mature NS4A, NS4B, NS5A, and NS5B proteins [90]. A number of PIs targeting the NS3 and NS4A region have been developed, but amino acid substitutions (polymorphisms) across the NS3 domain can counteract these PIs [91].

1.2.7 NS4A and NS4B

NS3 and NS4A combine to create the NS3-4A complex, with NS3 providing the enzymatic activity and NS4A functioning as the co-factor. The NS3-4A complex cleaves host proteins – such as TC-PTP (T cell protein tyrosine phosphate) and enhances epithelial growth factor (EGF)-induced signal

transduction, which is needed for HCV replication. The NS3-4A complex can block innate immune responses by cleaving CARDIF (CARD adaptor inducing IFN-β) and TRIF (Toll-interleukin-1 receptor domain-containing adaptor inducing IFN), which are important for RIG-I (retinoic acid inducible gene I) and TLR (Toll-like receptor) cascades. The RIG-I and TLR pathways are crucial for the production of IFNs, and disruption of these pathways hampers the innate immune response [92].

NS4B is a membrane protein possessing four transmembrane domains, and is cleaved from the HCV precursor polyprotein by NS3-4A serine protease activity. NS4B induces the membrane changes needed to form the membranous web. Once the changes are complete, the HCV replication complex associates and assembles using the membranous web. NS4B can also interact with other NS proteins, and displays NTPase activities [93].

1.2.8 NS5A and NS5B

NS5A consists of approximately 447 amino acids, and contains an N-terminal amphipathic alpha helix and three structural domains – Domain I, II, and III. Two low-complexity sequences are found between the structural domains, and are known as LCS1 and LCS2. NS5A is a phosphoprotein, with hyperphosphorylation thought to act as the “switch” that regulates its functions (replication and formation of the HCV particle). Domain I contains Zn^{2+}-binding and RNA-binding regions that are important for replication. Domain II has been associated with HCV replication. Domain III interacts with the core protein and is necessary for HCV assembly. Domains II and III appear to be unfolded domains, and can interact with several virus and host proteins [94 95].

NS5B is a membrane-associated phosphoprotein and an RNA-dependent RNA polymerase (RdRp) [96]. Because the HCV genome does not enter the host nucleus, an RNA polymerase must be encoded by the HCV genome to permit RNA synthesis. NS5B controls the synthesis of negative strand RNA from a positive strand template, and the negative strand becomes the template for RNA synthesis. The RNA templates can be stored for future RNA synthesis, protein translation, or packaging. NS5B represents an

important target for drug development. The catalytic site of NS5B does not tolerate many amino acid substitutions, being conserved throughout genotypes and subtypes. That said, a HCV isolate resistant to SOF (which targets NS5B) has already been described *in vitro* [97].

1.3 HCV LIFECYCLE

The entry and infection of HCV is thought to be mediated by interactions with host surface molecules, including CD-81, SR-B1, and LDLR, and other host receptors and proteins. CD-81 (cluster of differentiation – 81) is a member of the tetraspanin family of membrane proteins, and contains four hydrophobic transmembrane domains. The LEL (large extracellular loop) of CD-81 has been found to bind the E2 glycoprotein, and can mediate E2 binding to hepatocytes. The four cysteine residues present on the LEL create two disulphide bridges with the E2 glycoprotein, stabilising the interaction. A 2000 study found that only 30% of the experimental CD81-HCV complexes were internalised – a low uptake suggesting that CD-81 was not the only host receptor responsible for HCV entry [98]. SR-B1 (scavenger receptor class B – member 1), is a 509-aa integral membrane protein that enables the uptake of cholesteryl esters from HDLs (high-density lipoproteins). Because HDL originates from the liver, SR-B1 is found on hepatocytes – the primary target of HCV. HDLs and cholesterol are manipulated by a range of viruses to facilitate pathogenesis, including Dengue virus, another member of the *Flaviviridae* family. HCV E2 binds to hepatocytes *via* SR-B1, through the interaction of the SR-B1 LEL and the HVR1. SR-B1 and CD-81 have been found to cooperate to mediate HCV entry [99]. CD81 and SR-B1 can be considered as the primary host receptors for HCV, but several other host receptors are thought to participate, including LDL-R, glycosaminoglycans, DC-SIGN and L-SIGN, and asialoglycoprotein receptors [63].

The next stage covers involves HCV entry and fusion. The nucleocapsid of enveloped viruses is released into the cytoplasm through a fusion process that combines the virus envelope and endosome membrane. First, HCV creates an endocytic vesicle to enable the virus particle to enter the host.

These vesicles have clathirn and actin components. The actin components form the framework of the vesicle and clathirn components produce the coating. The endocytic vesicles mature and become endosomes, and the endosomes undergo the acidification process – this acidification process is what categorises HCV fusion as being pH-dependent. The acidic environment is thought to induce structural changes to the E1 and E2 glycoproteins, exposing regions to host fusion peptides. The fusion of the virus membrane and endosome membrane causes the decapsidation of the HCV particle, and enables the HCV +ssRNA to enter the host cytoplasm [100].

The HCV +ssRNA serves as mRNA for the synthesis of the HCV polyprotein. The IRES and a number of host proteins regulate the synthesis. Once the genome has been translated and the precursor polyprotein produced, the precursor polyprotein is targeted to the ER. There, the core protein and the E1 glycoprotein translocate the polyprotein to the ER lumen, where the polyprotein undergoes further processing by host signal peptidases. The signal peptidases cleave various segments of the polyprotein, causing the maturation and development of functional HCV proteins. This leads to the formation of the HCV replication complex. NS4B recruits the proteins needed to produce the membranous web, and NS5B RdRp initiates the replication process. The +ssRNA serves as a template for the negative strand intermediate. The intermediate becomes the template for positive strands, and the positive strands are used for polyprotein translation or packaged to form virus particles [101] [63].

The mechanism of HCV assembly and release has not been completely characterised, but the assembly and release process does appear to involve lipoproteins. LDs are storage organelles used for metabolic functions and membrane synthesis, and been shown to associate with the HCV core protein. Experiments that removed LDs from the model system found that virus production was hampered, and another experiment proposed that the core protein captured new RNA from surface of LDs, using the RNA from the LDs to complete the assembly process [102] [103].

1.4 HCV Infection and the Immune Response

HCV targets hepatocytes. Hepatocytes can be categorised as parenchymal hepatocytes and non-parenchymal cells (NPCs). NPCs include LSEC (sinusoidal endothelial cells), KCs (Kupffer cells), HSCs (hepatic stellate cells), and the population of immune cells that circulate throughout the hepatic system – such as DCs (dendritic cells), NKs (natural killer cells), lymphocytes, neutrophils, and monocytes. LSECs, HSCs, and KCs can express pathogen recognition receptors (PRRs) and respond to HCV infection [104] [105].

1.4.1 The Innate Immune Response to HCV Infection

The initial detection of HCV infection is carried out by the innate immune system, which is composed of various receptors capable of recognising structures present on pathogens. These structures are known as PAMPs (pathogen-associated molecular patterns) and are recognised by various host receptors, such as TLRs (Toll-like receptors), RLRs (RIG-I-like receptors), and NLRs (NOD-like receptors). Once the specific PAMP has been recognised, these PRRs trigger the expression of genes for the inflammatory and immunological response. RIG-I or MDA5 (melanoma differentiation antigen 5) can recognise cytoplasmic ssRNA produced by HCV, and the HCV 5'-UTR and the polyU/UC region of the 3'-UTR are also considered to be HCV PAMPs.

Once the PRRs have recognised the HCV PAMPs, the innate response is triggered. PAMP recognition by PRRs induces the activation of downstream pathways, which activate the production of cytokines, chemokines, IFN-I, and IFN-λ. IFN molecules are drawn to surface receptors and start a signal cascade that activates the JAK-STAT (Janus kinase – signal transducer and activator of transcription) pathway. The JAK-STAT pathway triggers the transcription and expression of genes that regulate apoptosis, oncogenesis, and immunological functions. The JAK-STAT pathway involves interactions between three components - a surface receptor, the JAK proteins, and the STAT proteins. The JAK proteins phosphorylate IFN receptors, and the STAT proteins are phosphorylated and translocate to the nucleus, where the

STAT proteins act as transcriptional activators for ISGs (IFN-stimulated genes). The major advantage of using IFN molecules to activate responses is that the molecular components needed for the process are present from the start, reducing the delay between pathogen recognition the host response [106]. ISGs have range of functions, from enhancing the detection and recognition of PAMPs by PRRs, to triggering an "antiviral state" that targets the ssRNA produced by HCV.

Several virus species have evolved to target the JAK-STAT response and interrupt ISG activation, but activation of STING (stimulator of interferon genes) and/or MAVS (mitochondrial antiviral-signalling protein) can be used as another means of ISG production *via* the NF-κβ pathway. Once NF-κβ has been translocated to the nucleus, NF-κβ binds to specific sites on the IFN-β promoter and to promotors of ISGs.

1.4.2 HCV Innate Immune Response Evasion

HCV implements several strategies to modulate the host immunological response. MAVS is found throughout the membranes of mitochondria near the ER and forms an aggregate that recruits TRAFs (TNF receptor associated factors) to activate the downstream IRF (interferon regulatory factor) and NF-κβ processed required for IFN production. HCV NS3-4A protease can cleave MAVS near the transmembrane domain, preventing MAVS from starting IFN production [107].

IFN receptors interact with JAK-STAT to transmit signals to the nucleus. Type I and Type III IFNs induce the transcription factors STAT1, STAT2 and IRF9. HCV can block the JAK-STAT pathway by having various regions (see Table 2) of the HCV genome bind STAT1, preventing STAT1 from activating the relevant ISGs [108]. Table 2 summarises the proposed mechanisms employed HCV to evade the innate immune response, and further shows that the majority of the HCV genome has methods of inhibiting and suppressing the host response to infection.

Viral Elements	Possible Mechanisms
Unknown	Impair JAK/STAT signalling by inducing miR-373 to target JAK1 and IRF9 mRNA.

Core	Impair JAK/STAT signalling by binding STAT-1 to block its phosphorylation and dimerization with STAT-2. Impair JAK/STAT signalling by inducing SOCS3. Suppress IRF1 transcription.
E2	Inhibit PKR-dependent phosphorylation of eIF2a.
NS2	Inhibit IRF3 phosphorylation by interacting with IKKε and TBK1
NS3	Inhibit IRF3 activation by interacting with TBK1. Inhibit TNF-α induced NF-κβ activation *via* binding to LUBAC.
NS3/4A	Impair RLR signalling by cleaving MAVS. Impair TLR3 signalling by cleaving TRIF.
NS4B	Impair RLR signalling by disrupting interactions of STING and MAVS. Impair RLR signalling by disrupting interactions of STING and TBK1. Inhibit RLR signalling by suppressing STING accumulation.
NS5A	Inhibit IRF-7 activation via interacting with IRF-7. Impair JAK/STAT signalling by binding STAT-1 to block its function. Impair TLR-MyD88 signalling via binding to MyD88. Inhibit PKR activity via interacting with PKR. Interact with 2′–5′ OAS/RNase L to block its function.

Table 2 – HCV strategies to avoid and modulate the innate immune response [109]*.*

1.4.3 The Adaptive Immune System

Adaptive immunity can be considered the second stage of the immunological response and provides defence against specific antigens and pathogens. Adaptive immunity creates an immunological memory, which protects against subsequent infections from the specific pathogen. Adaptive immunity has two components – humoral immunity and cellular immunity. Humoral immunity refers to the aspects of the immune response that are mediated by antibodies (Abs), complement proteins, and antimicrobial peptides. Cellular

immunity refers to the responses generated by phagocytes, cytotoxic T lymphocytes (CTLs), and cytokines.

The adaptive immune response to HCV infection can take weeks to develop. HCV RNA can be detected around 1 – 3 weeks post-exposure, but HCV-specific antibodies and CTLs can only be observed at around 1 – 2 months post-exposure.

1.4.4 The Cellular Immune Response to HCV Infection

CD4+ T helper cells and CTLs are both produced as a response to HCV. Macrophages, B lymphocytes, TNF-α, and DCs produce various types of interleukin (IL-12, IL-18, IL-6, etc.) which promote the growth and development of CD4+ T helper 2 (Th2) cells. Th2 mediates B lymphocyte activation, antibody production, and the regulation of Th1 responses.

CTLs are generated from CTL precursors (naïve T_C (cytotoxic T cells)) using three sequential signals. The majority of CTLs are CD8+ and target MHC Class I molecules. CTLs have two methods of triggering apoptosis. The first method is the cytotoxic protein pathway and uses granzymes and perforins to trigger endogenous apoptosis. The second method is the Fas ligand pathway and uses Caspase 1 to trigger DNA fragmentation. Cytokines are a group of soluble proteins that include chemokines, interferons, TNFs, and interleukins. Cytokines act as receptors that help coordinate the innate and adaptive responses to infection and are important for the clearance of HCV and removing damaged tissue.

Depicted in Figure 5, a study of patients who resolved acute HCV infections found that a large amount of IFN-γ was detectable at the early disease stage. IFN-γ is a cytokine needed for the activation of macrophages and the induction of the Class II major histocompatibility complex (MHC), and is produced by NKs, NKTs (natural killer T cells), CD4+ Th1 and CTLs. The study observed that IFN-γ produced by CD4+ Th1 peaked at the very start of the disease state and continued once HCV had been cleared, becoming a persistent defence against infection [110]. Those who resolved acute HCV infections or became “target not detected” showed broader CTL responses compared to those with chronic HCV infections [111].

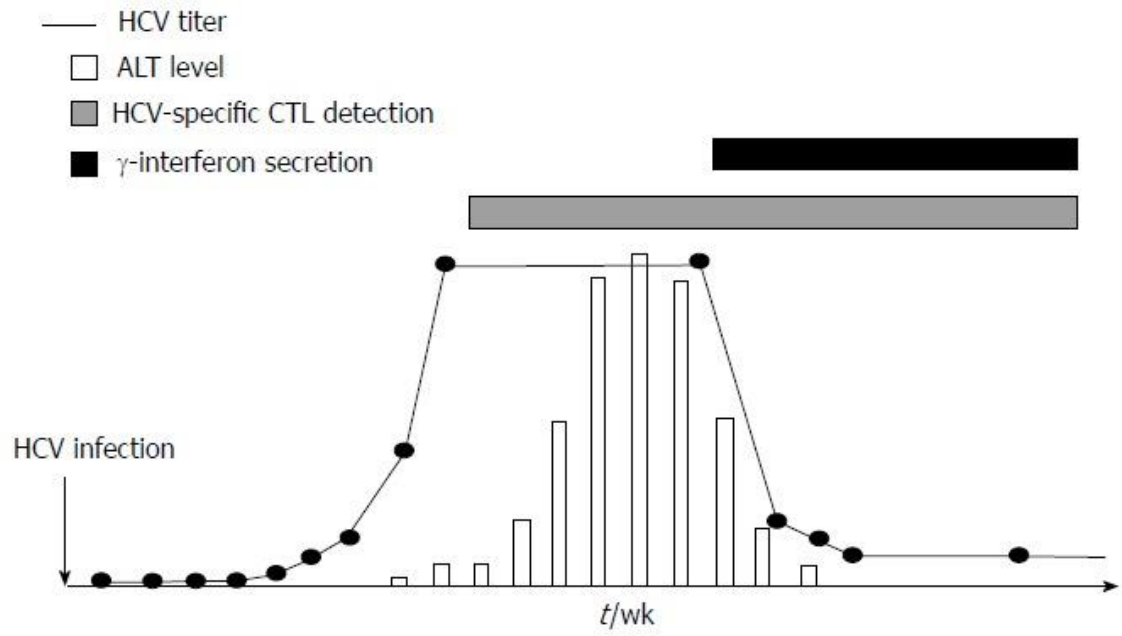


Figure 5 – A figure demonstrating HCV viral load versus ALT levels, presence of specific CTLs, and the presence of detectable IFN-γ over the course of an acute HCV infection. IFN-γ production persists even after HCV decreases [112].

1.4.5 HCV Cellular Immune Response Evasion

One of the main reasons for chronic HCV persistence is thought to be the rapid development of HCV escape mutations, but there are several other factors that contribute to HCV persistence – namely CD8+ exhaustion and the HCV-mediated secretion and induction of regulatory cytokines and modulatory Tregs (regulatory T cells) [111].

HCV-specific CD8+ responses are thought to be necessary for the control of HCV infection, as evidenced by *in vivo* CD8+ studies using chimpanzee models and functional analysis of the CD8+ responses of patients undergoing acute HCV infections [113]. During a chronic HCV infection, the continuous exposure to HCV antigens is thought to contribute to the exhaustion of CD8+ functions. The study examining CD8+ responses during chronic HCV infections found that the expression of granzyme A (a serine protease released by NK and CTL cytoplasmic granules to trigger apoptosis) was decreased, TNF-α production was decreased, and CTLs were unable to properly degranulate when stimulated by peptides. The CTL effector functions were further hampered by decreased IFN-γ production. These

impairments are not direct reason for the persistence of HCV, but contribute to the weakened immunological state that enables persistence [114].

HCV interferes with the cytokine network by inducing Th2/Tc2 cytokine responses, causing inflammatory molecules to migrate to damaged tissue. Promoted by the presence of IFN-γ, hepatocytes infected with HCV secrete chemokines (such as the CXC chemokine ligands CXCL9, CXCL10, and CXCL11) that trigger the migration of non-specific mononuclear cells. CXC chemokines are a type of chemokine where the two N-terminal cysteines are separated by a single amino acid, represented by the “X”. The presence of these causes chronic inflammation and activates hepatic myofibroblasts and fibroblasts, leading to the development of fibrosis and cirrhosis. A number of extrahepatic diseases associated with HCV infection – such as thyroid disorders, type 2 diabetes, and mixed cryoglobulinemia – have been found to be connected to the dysregulation of the cytokine network and the induction of inflammatory chemokines [115].

1.4.6 The Humoral Immune Response to HCV Infection

As previously stated, humoral immunity refers to the aspects of the immune response that are mediated by antibodies (also known as immunoglobulins (Ig)), complement proteins, and antimicrobial peptides. Antibodies are Y-shaped glycoproteins produced as a response to specific antigens being presented to plasma cells (also known as plasma B cells). An antibody consists of two Ig heavy chains and two Ig light chains bound together by disulphide bonds. Both heavy and light chains feature constant and variable domains. Two fragment antigen-binding (Fab) variable regions form the arms of the Y-shaped structure and – as the name suggests – are used to bind antigens. The end of the Fab arms contain the antigen-binding site (also known as a paratope) and are considered to be hypervariable regions due to the vast range of antibody paratopes that can be produced. The process that drives paratope diversity is known as V(D)J or VJ recombination. Three genes (variable (V), joining (J), and diversity (D)) are responsible for the antibody paratope, and these genes randomly recombine to produce a wide array of paratopes. The fragment crystallisable region (Fc) forms the stem of the Y-shaped structure and interacts with Fc receptors and the complement

system through a conserved glycosylation site. The heavy Ig chain (encompassing the Fc region and the lower segment of the Fab regions) determine the isotype of the antibody - the isotype determines the broad function and location of the antibody. In humans, there are five major isotypes: IgA, IgD, IgG, IgE, and IgM.

IgM is the first antibody produced as a response to antigen exposure and activates the classical complement pathway. The HCV capsid antigen is the first target of IgM [116]. IgM can re-emerge during chronic HCV infections, as the body responds to what it interprets as a new HCV infection – instead, this re-emergence of IgM is evidence of the continual evolution of HCV antigenic epitopes [117]. IgG first targets the HCV core protein, emerging almost at the same time as IgM. IgG accounts for about 75% of human serum antibodies, and is required for complement pathway activation, agglutination, toxin neutralisation, and several other immunological functions.

During the HCV disease state, both neutralising antibodies (nAbs) and non-nAbs are produced, targeting structural and NS HCV proteins. nAbs bind to surface structures of virus and other infectious agents to prevent the interaction with host surface receptors. nAbs “prevent” infection from viruses, and immunity from nAbs is also known as sterilising immunity. Non-nAbs do not neutralise infectious agents, but can contribute to the humoral immune response by flagging the infectious agent for downstream neutralisation by other components of the immune system. For example, a 2017 study performed by Long *et al.* found that non-nAbs induced a strong HCV-specific antibody-dependent CD56+ NK response in chronic HCV patients. NK antibody-dependent cellular cytotoxicity (NK-ADCC) is caused by the activation of the CD16 region on NKs by the Fc region of IgG bound to infectious agent. NKs bind the infectious agent and degranulate, releasing granzymes and perforin to lyse the target. Long *et al.* examined the NK-ADCC responses of 31 chronic and untreated HCV patients and found that their NK-ADCC responses were significantly impaired compared to healthy controls – however Long *et al.* acknowledged that there was ongoing controversy about whether or not NK-ADCC responses and NK natural cytotoxicity are impaired during chronic HCV infection, with other studies

demonstrating unchanged or increased responses [118]. 17 linear epitopes across the E1E2 glycoprotein for HCV genotypes 1b and 2a were generated for the study, with fiver epitopes shown to reduce a robust NK-ADCC response. Four of these epitopes did not overlap with known any putative neutralising epitopes, suggesting that non-nAbs form a part of the humoral response to HCV infection [119].

The antibodies produced during acute HCV infections tend to target epitopes on the structural and NS HCV proteins, but during chronic HCV infections nAbs tend to only target epitopes on the HCV envelope glycoproteins and the E1E2 heterodimer. The ectodomain of E2 contains three regions that are targeted by nAbs; HVR1, HVR2, and the IgVR. HVR1 contains the dominant nAb epitopes, and the strongest E2 nAbs target the C-terminus of E2, where the E2 glycoprotein interacts with host SR-BI. Regions of the E2 glycoprotein that interact with entry factors (CD81 and SR-BI, etc.) are prime targets for nAbs. A percentage of nAbs that recognise conformational epitopes on the E1E2 heterodimer recognise epitopes for a range of HCV genotypes – these are known as broadly-neutralising antibodies (bnAbs). Because the HVR1 mutates and produces numerous variants, nAbs that target the HVR1 tend to be target a single genotype. [120]. The role of nAbs during acute and chronic HCV infections has not been completely characterised, but there are suggestions that nAbs can prevent subsequent infections. Chronic HCV patients have been found to show delayed nAb development, but the target range of nAbs does increase as time passes. Individuals that have detectable bnAbs at the acute stages of HCV infection have a greater chance of spontaneous clearance, and those who have experienced spontaneous clearance of one HCV infection have a greater chance of clearing subsequent infections [121] [122].

1.4.7 HCV Humoral Immune Response Evasion

HCV evades the humoral immune response through genetic variation, the protection of vulnerable epitopes using glycosylation, the production of LVPs, and stimulating the production of competing antibodies. It has been proposed that the CD81-binding loop of the E2 glycoprotein can be targeted by nAbs

and non-nAbs, creating a form of competitive inhibition which prevents nAbs from accessing vulnerable epitopes.

The extraordinary replication rate of HCV – estimated to be around 10^{12} particles/day for chronic HCV infections – enables the virus to produce a huge range of variants resistant to nAbs. It has been shown that the evolution of the E1 and E2 glycoproteins (and HVR1) is strongly shaped by immunological pressure from nAbs, with Liu *et al.* demonstrating that for one individual the HVR1 remained stable for almost two years, but begun mutating once nAbs were introduced [123]. It has been suggested that the HVR1 could act as an immunological decoy – the presence or absence of HVR1 does not impact entry and infection, but HVR1 is highly immunogenic and could serve to "distract" host nAbs from epitopes that are necessary for HCV [124].

In 2014 Palmer *et al.* – using ultra-deep pyrosequencing (UDPS) – observed the emergence and disappearance of resistant HCV quasispecies over the course of 9.6 years of chronic and untreated HCV genotype 4a infection, identifying two primary lineages – L1 and L2. Ten samples were taken over the course of the 9.6 years (RL1 through RL10). L1 was further divided into three sub-lineages – L1a, L1b, and L1c. The group tracked the dominance of L1 over L2 for the first few years of the study, but L2 dominance began to occur as the L1 variants were suppressed by the host humoral system. It was found that L2 variants had a distinct evolutionary trajectory when compared to L1 variants. L2 variants showed a higher rate of synonymous (a single base substitution that does not change the amino acid produced) mutations, with dominant L2 variants demonstrating lower diversity but higher genetic stability. In contrast, L1 variants seemed to target IgG as their primary method of immune response evasion, with no IgG-targeting observed in the dominant L2 virions. The dominant L2 variants sampled from RL4, RL7, RL8, and RL10 varied by just a single synonymous mutation, implying the existence of host humoral immune pressure under which only a narrow range of L2 HVR1 variants could survive. Coupled with the L1 variants undergoing an extinction event (driven by nAb evolution and dependence on IgG-targeting) in RL7 and the strict purifying selection of L2

variants, L2 emerged as the dominant quasispecies from RL7 to RL10 [27]. Naik *et al.* added to this research by examining four serum samples taken from the same untreated HCV genotype 4a patient, extending the study to thirteen years. The four samples taken showed the continued dominance of the L2 lineage, but that a new HVR1 variant had superseded the primary HVR1 variant [125].

A number of conserved nAb epitopes have been discovered on the E1 and E2 glycoproteins, and two immunogenic domains covering the N-terminal and C-terminal regions of the HVR1 have been described, with the C-terminal region being the primary HVR1 neutralisation determinant [126]. However, glycosylation of the envelope glycoproteins appears to protect these conserved epitopes from nAbs – see Section 1.2.4 for further information [107] [126].

HCV particles circulate throughout the host as a mixture of LVPs and apolipoprotein B (apoB) subviral particles. These apoB particles only express the HCV envelope glycoproteins, meaning that they can become a target and distraction for nAbs and prevent the nAbs from reaching proper HCV particles [127]. A number of studies have suggested that LVPs are resistant to nAbs and that LVPs are needed for the establishment of chronic infections – this suggestion is supported by research showing that low levels of LVPs during acute HCV infections are a predictor of spontaneous HCV clearance [128]. Chronic HCV infections are associated with reduced amounts of cholesterol and apoB (because they have been 'hijacked' by HCV to produce LVPs and subviral particles) causing increased rates of steatosis and fibrosis, and resistance to IFN-based therapies.

1.4.8 HCV and Hepatic Immunity

The liver has several unique immunological properties. For example, hepatocytes produce 80 – 90% of circulating innate immunity proteins. The liver must tolerate food antigens and molecules produced by the gut microbiome without wrongly producing an inflammatory response - but be able to respond to potential pathogens before they spread to the rest of the body. The liver houses the single largest population of macrophages (Kupffer cells), which accounts for 80 – 90% of the total population of fixed

macrophages in the human body. The liver is also enriched with NKs. Liver sinusoidal endothelial cells (LSECs) are responsible for removing soluble macromolecular waste from circulation and can detect microbial infection using PRRs [104]. Blood arrives through the portal vein, where PRRs on KCs and hepatocytes bind to MAMPs and DAMPs and phagocytose and degrade the products to prevent the body from mounting an excessive immune response [129].

KCs are equipped with a large range of PRRs, Fc receptors (a receptor that recognises the Fc fragment of antibodies) and complement receptors. KCs are derived from circulating monocytes, and have varied physiological and cytological characteristics depending on their location. For example, large KCs are present near the periportal region, where they have access to incoming pathogens. These large KCs also show higher production of TNF-α and IL-1 when compared to other KCs [130]. TLR3 is an endosomal sensor of dsRNA expressed by KCs. TLR3 signals induce IRF-3 and NF-κβ to produce IFN, chemokines, and cytokines. Chemokine and cytokine induction from TLR3-HCV interactions has been shown to be delayed when compared to TLR3 interactions with other pathogens – HCV is an ssRNA virus, and that the dsRNA that TLR3 detects are replication intermediates produced late during the HCV lifecycle [131]. TLR3 sensing of HCV could serve as an important secondary detection system once the primary RIG-1 detection of HCV has been completed [132].

Dendritic cells (DCs) are antigen-presenting cells (APCs), and act as a bridge between the innate and adaptive immune systems by processing antigens and presenting them to T-cells to induce the adaptive immune response. Both myeloid (mDCs) and plasmacytoid (pDCs) are present in the healthy liver. mDCs secrete IL-12, IL-6, TNF, and chemokines, and display TLR2 and TLR4, whereas pDCs produce large amounts of IFN-α and display TLR7 and TLR9. Previously, the hepatic DC population was considered to be immature, but evidence has emerged that DCs can induce strong T-cell and cytokine responses. CD141(+) – a subtype of hepatic mDCs – has shown to produce increased levels of CD141 when compared to circulating mDCs (almost one third of CD141(+) mDCs producing CD141 versus <5% of

circulating mDCs). The CD141(+) mDC population was also found to be severely depleted during hepatic disease, suggesting a role for CD141(+) in maintaining normal function [133]. pDCs detect HCV PAMPs and trigger the IFN response. The pre-treatment ISG mRNA expression patterns of patients with chronic HCV infections were examined by Stone *et al.* It was found that IFN was being produced endogenously. Patients with the favourable *IFNL3* SNPs tended to have higher levels of *IFNL3* and lower levels of hepatic ISG - but increased macrophage ISG levels, suggesting that "bystander" uninfected hepatocytes can detect infected hepatocytes and produce IFN – "bystander" hepatocytes such as pDCs [134]. This response can be triggered by sub-genomic HCV particles, implying that a type of RNA transfer process exists between pDCs and other hepatocytes, circumventing the "typical" pathogen recognition pathways [135].

LSECs (also known as hepatic sinusoids) account for about 3% of the total hepatocyte volume, but are important components of the innate immune system. LSECs are responsible for the clearance of blood-borne waste, the mediation of lymphocyte migration across the sinusoidal endothelium and the prevention of fibrosis and steatosis. LSECs are characterised by their unique fenestrae (open pores), which are arranged across the LSEC surface in groups known as "sieve plates". These fenestrae are crucial for filtration activities, but can become damaged during hepatic disease. Damaged LSECs lose their fenestrae, develop a continuous basal membrane (absent from undamaged LSECS), and acquire pro-inflammatory features [136]. The loss of fenestrae is known as "capillarisation", and is associated with the development of fibrosis [137]. It has been demonstrated that the capillarisation of LSECs is furthered by the increased levels of liver X receptor alpha (LXRα) (a regulator of macrophage function and a modulator of steatosis development) – HCV core and NS5A proteins have both been shown to indirectly upregulate LXRα, meaning that HCV contributes to LSEC capillarisation and therefore the development of steatosis [138] [139].

1.4.9 Summary of HCV Infection and the Immune Response

The immune response to HCV infection can be divided between the innate and adaptive responses. PRRs recognise HCV PAMPs and initiate the

innate immune response. RIG-I or MDA5 recognise cytoplasmic ssRNA are crucial for the early detection of HCV. IFN is produced and activates the JAK-STAT pathway – responsible for the transcription and expression of genes that regulate apoptosis, oncogenesis, and immunological functions. APCs – such as the hepatic DCs – act as the bridge between the innate and adaptive immune response, and present antigens to T-cells. The adaptive response can be separated into cellular and humoral responses. The cellular response includes the production of phagocytes, CTLs, and cytokines. The humoral response refers to the antibody response to HCV infection. HCV has a number of methods to evade both the innate and adaptive immune responses – for example, blocking part of the JAK-STAT pathway, upregulating inflammatory responses to damage host tissues, and producing numerous HCV variants to evade nAbs, among other methods.

1.5 HCV AND VACCINATION

The development of both a prophylactic and a therapeutic HCV vaccine is important for combating the global HCV epidemic. Key to the development of an HCV vaccine is the identification of the immunological factors that enable ~15% of people to experience only acute HCV infections and prevents the ~85% of those from clearing chronic HCV infections. Therapeutic HCV vaccine development can be divided between those that concentrate on inducing strong CD4+ helper T and CTL responses during the acute stages, and those that focus on promoting the generation of functional nAbs. There are several obstacles impeding the development of an HCV vaccine – namely HCV's extraordinary genetic diversity. There exists a broad range of vaccine strategies, such as recombinant, DNA, peptide, and vector designs.

A recombinant protein vaccine is produced by isolating and purifying the pathogen proteins expressed by a model system. This system has the advantage of not having the actual pathogen. Experimental vaccines have attempted to use the E1 and E2 glycoproteins to generate the desired immunological response – for example, Innogenetics tested a prophylactic vaccine using recombinant E1 and an aluminium hydroxide adjuvant but

ceased development around 2007. Novartis tested a prototype vaccine combining the HCV core protein and the ISCOMATRIX adjuvant [140]. Another study examining a vaccine generated from the HCVcc system found that an HCVcc particle combined with the K3-SPG adjuvant triggered the production of nAbs against HCV genotypes 1a, 1b, 2a, and 3a. The production of IFN-γ was also triggered as a response to the presence of the HCV core protein [141]. A research group examining a vaccine containing genotype 1 E1E2 and a MF59C.1 adjuvant found that the vaccine induced a “significant” response [142]. During the course of prophylactic vaccine studies on chimpanzees, researchers found that out of the chimpanzees that had been vaccinated but remained infected, 45.8% of the chimpanzees had been administered a vaccine using HCV NS proteins, compared to 13.8% that were administered vaccines using HCV structural proteins [143].

An epitope (antigenic determinant) is the part of an antigen to which antibodies bind. Epitopes can be linear or conformational: linear epitopes are epitopes that are recognised through their primary protein structure (amino acid sequence), whereas conformational epitopes are recognised using their tertiary protein structure (3D structure). Identifying conformational HCV epitopes provides a means of avoiding the problem of HCV’s genetic diversity, as 3D structures have a greater chance of being conserved across genotypes and subtypes. An investigation of bnAbs derived from B lymphocytes taken from subjects with chronic HCV infections has identified a number of clusters of epitopes on E1E2. Four clusters are designated as antigenic domains A – D and are conformational epitopes on E2. There are two conformational epitope clusters defined as antigenic regions (AR) 4 and 5, and antigenic domain E is another cluster containing linear epitopes located adjacent to the HVR1 [144]. A 2020 study examined the use of monomeric forms of E2 as a potential vaccine candidate. Previous work from the same group used an HCVcc-derived high molecular weight (HMW) form of E2 and found that it elicited a broad neutralisation response against the seven major HCV genotypes. The HMV form lacked the three variable regions of E2 (HVR1, HVR2 and IgVR) and was thus named Δ123-HMW for the study. Despite the promising neutralisation response, Δ123-HMW was

produced at low yields and did not have a homogenous composition. The 2020 update to this study used sequential reduction and oxidation to improve Δ123-HMW yield and homogeneity, naming the new structure Δ123r. Δ123r largely retained the same immunogenic properties of Δ123-HMW – both showed strong reactivity to bNAbs with linear epitopes (those targeting AS412 and AS434, for example), but showed reduced reactivity to bNAbs with epitopes from the LEL of the CD81-binding domain and conformation-dependent epitopes that are occluded by the assembled glycoprotein [145] [146]. A study focusing on HIV – another virus with notable antigenic diversity that uses glycan shielding to evade antibodies – found that they were able to "reverse engineer" a nanoparticle that improved B lymphocyte activation and guided bnAb maturation. The nanoparticle used the atomic, hydrophobic, and electrostatic features of promising bnAbs that targeted the HIV gp120 and originated from the VH1-2*02 variable heavy gene. Further, the VH1-2 genes are present across ~2% of the human Ab repertoire [147]. This approach - targeting the structural conformation of the virus - represents a way of overcoming the sequence variations and evasion strategies of a broad range of viruses.

For prophylactic HCV vaccines, a meta-analysis has been conducted on experiments using chimpanzee models for vaccine development. The study noted that humans and chimpanzees had a spontaneous clearance rate of 83% for secondary HCV infections, compared to ~25% for humans and ~38% chimpanzees for primary HCV infections. Prophylactic vaccines tested against chimpanzees were able to induce a memory response upon exposure to HCV and were somewhat able to control HCV replication. However, the duration of viremia was not shortened by vaccination, suggesting that the "natural" immunity produced during a secondary infection was superior to vaccine immunity. The vaccines that seemed to have the best rates of resolving infection included parts of E1E2 glycoprotein that induced nAbs, T cell responses, or a combination of both. That said, the induction of nAbs does not guarantee the success of the treatment. The analysis concluded that vaccines containing HCV structural proteins (core and the envelope glycoproteins) appeared to be the most promising, but that

a prophylactic vaccine inducing complete immunity against HCV might not be possible – however, if the prophylactic HCV vaccine could reduce HCV persistence, natural immunity could be capable of clearing the virus [143] [148].

The advent of DAAs as an effective treatment for the majority of HCV patients has raised the question of whether an HCV vaccine – therapeutic or prophylactic – is needed. As Rosen argues, having a treatment or "cure" for a disease does not mean that the disease has been eradicated, as evidenced by the persistence of syphilis throughout the decades. Further, only widespread vaccination was able to completely eradicate smallpox and polio [149]. The subsequent question arises of which patient populations should be chosen for expensive DAA treatment and potential vaccination. Among developed countries, IDUs continue to represent the largest source of chronic HCV infections, but this population is vulnerable to subsequent infections due to the use and sharing of drug paraphernalia. The use of vaccination once DAA therapy has been completed could be a powerful method for protecting against secondary infections and slowing the spread of HCV among the IDU population. Hahn *et al*. proposed that widespread vaccination could reduce chronic HCV incidence among IDUs to 2-7% thirty years after the vaccine's introduction, contrasted to the 2009 chronic HCV incidence rate of 14-22% [150]. In 2019 Scott *et al.* conducted a study examining the impact of a HCV vaccine on the feasibility and cost of achieving the WHO HCV elimination target of an 80% reduction in the incidence of HCV by 2030 in the age of DAA treatment. Projections were made for 2018 – 2030 for 167 countries and involved two population groups – PWID and the general community. Scott *et al*. found that only 0 – 48 of the countries modelled would be able to achieve the 80% reduction of HCV incidences by 2030 without a vaccine, whereas 15 – 113 countries would be able to achieve this goal if a 75% efficacious vaccine with a 10-year duration was available. The study found that the optimal way of deploying a potential vaccine would be alongside test-and-treat programmes and vaccinating adolescents from communities with high levels of transmission. The study also predicted that if a course of vaccination cost US$200, the cost of elimination for 40% of countries would be reduced by an aggregate of

US$7.4 billion [151]. Despite the effectiveness and increasing accessibility and affordability of modern DAA treatments, predictive models such as Scott *et al.*'s suggest that widespread elimination of HCV is unlikely without vaccination.

1.6 METHODS OF STUDYING HCV

1.6.1 In Vivo Systems

One of the major obstacles to studying HCV infection has been the lack of appropriate systems that can support the infection and replication of HCV. The past decades have shown a distinct movement away from animal systems and towards *in vitro* systems. The chimpanzee system was among the first systems used for the studying of HCV – the 1989 isolation of a cDNA clone of HCV was derived from infected chimpanzee serum, for instance [18]. Humans and chimpanzees are the only species that are susceptible to HCV infection. Despite the valuable information obtained from the chimpanzee system, growing financial, ecological, and ethical concerns about the usage of this system has led to many countries banning the use of chimpanzees for research purposes. In 2013, the European Union issued Directive 2010/63/EU which restricts the use of great apes for research, "only for the purposes of research aimed at the preservation of those species and where action in relation to a life-threatening, debilitating condition endangering human beings is warranted, and no other species or alternative method would suffice in order to achieve the aims of the procedure".

One *in vivo* alternative to chimpanzees is the usage of murine models that have been altered and made susceptible to HCV infection. Rodents are resistant to HCV infection, but there are several ways to overcome this resistance, including genetic approaches, xenotransplantation, and the use of HCV homologs. Researchers investigated viruses that were related to HCV and capable of causing hepatitis among various species. GBV-B can be used to infect various species of New World monkeys, but the establishment of a persistent infection has proven troublesome. Recent deep sequencing studies have discovered a number of HCV-related pegiviruses and

hepaciviruses among dogs, horses, rodents, bats, and non-human primates – these new viruses could prove to useful as homologs for HCV infection [152].

Another strategy involves the genetic humanisation of the mouse model, removing barriers to HCV infection and introducing human HCV entry factors. One study introduced human CD81 and OCLN to render the transgenic mice vulnerable to HCV infection, but found that the adaptive and innate immune responses restricted HCV persistence. In response, the gene producing cyclophilin A was removed from the murine systems, enabling the infection and replication of HCVcc particles [153]. Xenografts are a second method of making the murine system susceptible to HCV infection. The chimeric mice are generated by transplanting primary human hepatocytes to an immunodeficient mouse that has constitutive or introduced liver damage. Then, the human hepatocytes repopulate the damaged liver and can support HCV infection. The uPA-SCID mouse and the FRG mouse are the two dominant strains used for generating the chimeric mice, but TK-NOG and MUP-uPA have been shown to support HCV. However, both the humanised and chimeric murine systems need immunodeficiency, and might not produce an accurate image of the response to HCV infection [152].

1.6.2 In Vitro Systems

In vitro systems for studying HCV have been developed over the past two decades and have proven crucial for understanding the mechanisms of HCV infection and replication. These *in vitro* systems include sub-genomic replicons, the soluble E2 glycoprotein, HCV pseudoparticles, and HCVcc system. The first obstacle to the creation of *in vitro* systems for studying HCV was the identification of cellular systems that could support HCV infection. Primary hepatocytes from humans and chimpanzees could support HCV infection, but primary hepatocytes from baboons were not susceptible to infection. Peripheral blood mononuclear cells (PBMCs) were also permissive to HCV infection, but the PBMCs and primary hepatocytes did not display adequate efficiency and made for unsuitable systems. Additionally, primary hepatocytes could contain antibodies against HCV, reducing the infectious potential of HCV+ sera [154]. To avoid these problems, a number of immortalised systems were investigated, and, among other lineages, the

human hepatoma lineages Huh7, HepG2, and PH5CH were found to be very susceptible to HCV infection, – with Huh7 being particularly favourable for HCV replication [155].

Liver biopsies taken from HCV-infected patients were tested as a potential *in vitro* HCV system. The liver biopsies underwent collagenase digestion to isolate the hepatocytes before infection, and the presence of HCV RNA was used to verify successful HCV propagation. However, the awkwardness of obtaining biopsy samples from patients and the poor reproducibility of the system meant that this approach was not pursued further [154]. In 1999, Lohmann *et al.* defined the minimal set of HCV proteins needed for the initiation and maintenance of *in vitro* HCV replication – the NS3/4A protease, the NS5A protein, and the NS5B were found to be indispensable for replication, and this discovery led to the creation of the subgenomic replicon system [156]. The subgenomic replicon contains the HCV non-structural proteins expressed on an encephalomyocarditis virus (EMCV) construct, and an antibiotic resistance gene used for selection purposes. The replicon can then be introduced to various hepatoma lineages, enabling the identification of permissive lineages and adaptive mutations favourable for *in vitro* HCV replication. For instance, a clone harbouring one of the adaptive mutations identified using the subgenomic replicon system was found to be infectious to chimpanzees. The subgenomic replicon system has the advantage of not being restricted to specific HCV genotypes, and has been helpful for screening novel molecules for the presence of antiviral activity against HCV [157 158]. The production of soluble E2 glycoprotein ectodomain (sE2) is another method of studying HCV. The transmembrane region of the E2 glycoprotein is removed, with the remaining E2 ectodomain retaining proper folding properties and having the ability to interact with the HCV entry receptors CD81 and SR-BI [158]. Various culture systems have enabled the purification of the E2 ectodomain, with Tello *et al.* purifying 2 mg of the glycoprotein from 1 L of culture media [159]. The sE2 system has enabled the further characterisation of the E2 glycoprotein, such as elucidation of the glycosylation pattern, oligomerisation state, and secondary structure of E2 [160].

1.6.3 HCV Retroviral Pseudoparticles (HCVpp)

The HCV pseudoparticle (HCVpp) system has emerged as one of the most successful methods for studying the HCV glycoproteins' importance for virus entry and infection. The method involves incorporating the E1 and E2 HCV glycoproteins onto a lentiviral or retroviral core particle and the introduction of a reporter gene. This "three plasmid" system then forms a pseudoparticle which closely mimics the entry and infection capabilities of wildtype HCV [161]. Bartosch *et al.* are credited with pioneering the HCVpp system, creating a functional and infectious genotype 1a pseudoparticle around March 2003. The pseudoparticle used genotype 1a E1 and E2 glycoproteins derived from a pTM1p5E1E2(745) plasmid, and inserted those glycoproteins onto a CMV-Gag-Pol murine leukaemia virus (MLV) packaging vector, coupled with a MLV-GFP plasmid to act as the reporter gene. *Gag* and *pol* are two of the three major proteins encoded within the retroviral genome – *gag* (an acronym for group antigens) codes for the core structural proteins of a retrovirus, whereas *pol* acts as the reverse transcriptase needed for replication. Together, the *gag* and *pol* of the MLV (a retrovirus) plasmid package and form the functional pseudoparticle. The reporter plasmid consisted of a second MLV construct containing green fluorescent protein (GFP). 293T HEKs were transfected with the three plasmids to produce the pseudoparticles, and a range of infection assays were carried to measure the infectivity of the pseudoparticles. The assays showed that the pseudoparticles were highly infectious and showed a distinct preference for hepatocytes. Importantly, Bartosch *et al.* now had a system that could reliably produce infectious pseudoparticles [162]. Since 2003, the pseudoparticle system has been verified by numerous laboratories. MLV and HIV are both used the retroviral packaging vectors, and GFP and luciferase are used for the reporter gene. Importantly, the HCVpp system works across numerous genotypes and subtypes.

However, there are disadvantages to the HCVpp system. Generation of pseudoparticles using E1 and E2 glycoproteins from infectious/patient-derived HCV samples does not guarantee that the pseudoparticle produced is going to be infectious – for example, Urbanowicz *et al.* found 493 HCV

clones originating from 63 patients that contained complete E1E2 ORFs, with 118 clones found be infectious. Despite this, only 78 pseudoparticles were produced, and several of the clones were found to be incompatible with the pseudoparticle system. Additionally, it was found that the pseudoparticles were more vulnerable to neutralisation when compared to corresponding HCVcc chimeras, meaning that findings from the pseudoparticle assays would need to be validated using the HCVcc system [163]. Similarly, Lavillette *et al.* found that only 24 infectious pseudoparticles could be produced from 88 clones, and found that the infectious titres of genotype 3 pseudoparticles were over a log lower than pseudoparticles derived from other genotypes [164].

Despite these problems with the pseudoparticle system, Urbanowicz *et al.* has made progress with optimising the system and identifying the factors that prevent patient-derived clones from becoming functional pseudoparticles. 24 – 27% of glycoprotein clones had been previously reported to yield infectious pseudoparticles, but Urbanowicz *et al*. found that the packing vector used had a distinct influence on the infectivity of the pseudoparticles. It was found that the infectivity phenotype of HCVpp generated from MLV packaging vectors was not the same as the infectivity phenotype of HCVpp generated from HIV packaging vectors. The quantity used of the packaging vector was also important for the production of functional particles, with the greatest incorporation of the E1E2 glycoprotein observed at 1.6 µg and 8 µg of packing vector [165]. In 2019, Bailey *et al.* published an updated standardised protocol for producing HCVpp, with the latest recommended reagents and two specific transfection protocols for HIV Gag and MLV Gag packaging vectors [166]. Additionally, Soares *et al.* demonstrated that co-expression of a p7NS2 ORF alongside HCVpp could enhance the infectivity of HCVpp. The exact mechanism of the increased mechanism was not defined, but Soares *et al*. were able to eliminate increased virus particles and increased envelope density as the causes of the increased infectivity [167]. The HCVpp system remains one of the most important systems for studying HCV and has been used to investigate various aspects of HCV entry and host receptors, including the roles of CD-81, CLDN, CLDN, and glycosaminoglycans, and to

investigate the interaction between antibodies and the entry pathways of HCV.

1.6.4 Cell Cultured HCV (HCVcc)

The HCVcc system represents another model for studying HCV and originated from a genotype 2a isolate taken from a Japanese patient with fulminant hepatitis – this clone was named JFH-1. Kato *et al.* extracted the total RNA from the patient and produced 14 fragments of HCV cDNA covering the entire HCV genome. The non-structural proteins were selected and assembled to form subgenomic replicons (Figure 6), which were introduced to Huh7s. It was found that the JFH-1 subgenomic replicon (pSGR-JFH-1) was capable of replication and could transmit G418 resistance (G418 was the selective antibiotic used during this study) during replication. pSGR-JFH-1 displayed superior colony formation efficiency when compared to an adapted mutant construct, and was the first successful subgenomic replicon to be derived from a genotype other than HCV genotype 1 [168]. Studies have shown that the JFH1 HCVcc isolate has the ability to replicate *in vivo* using the chimpanzee model, with the chimpanzees developing symptoms that mimic the symptoms of chimpanzees infected with HCV derived from infected humans [169].

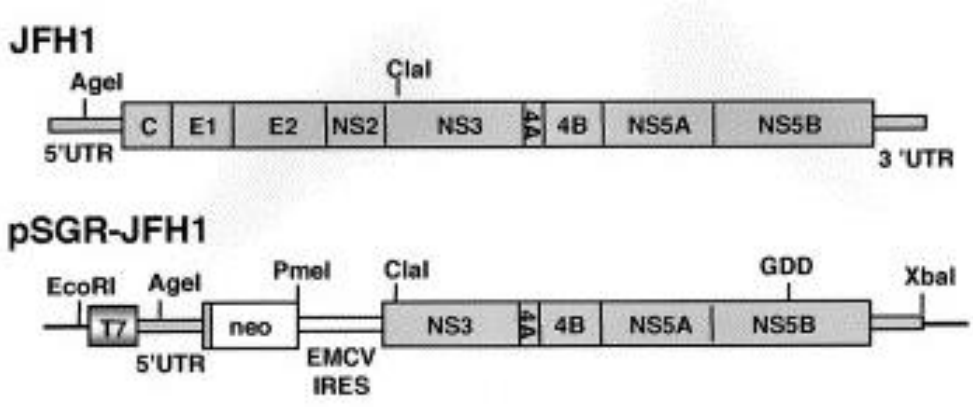


Figure 6 – A comparison of the complete JFH-1 genome and the sub-genomic pSGR-JFH-1 replicon assembled by Kato et al. [168]

Since Kato *et al.*, the HCVcc system has expanded to include other HCV genotypes. Gottwein *et al.* managed to produce an intergenotypic 3a/JFH1 recombinant with the core, E1, E2, p7, and NS2 proteins from the genotype 3a reference strain S52 but found that many of the recombinants contained

putative adaptive mutations, and only two of the nine recombinants with mutations were viable. However, the 3a/JFH1 recombinants proved useful for studying the dependence of HCV infection on CD81, and provided a means of studying therapeutics targeting the HCV genotype 3a core and NS2 proteins [170]. It has also been found that clones that were not viable when produced by the pseudoparticle system could become viable and infectious once transferred to the HCVcc system [171]. Biophysical studies of HCV and potential vaccine development requires the generation of concentrated and purified HCVcc stocks – however, culture conditions for HCVcc tend to include foetal bovine serum (FBS), which impedes the production of purified stocks. Mathiesen *et al*. have had success producing HCVcc particles using serum-free culture conditions, and found that the serum-free HCVcc (sf-HCVcc) were 0.6 – 2.1 $\log_{10}$ more infectious than HCVcc produced using media supplemented with FBS [172]. One disadvantage to the HCVcc system is the use of intergenotypic NS5B – NS5B influences the replication efficiency of the completed HCVcc construct, and it has been found that introducing intergenotypic NS5B can cause the HCVcc construct to have reduced replication efficiency or become completely nonviable [173].

1.7 Thesis Research Summary

The hypothesis for the thesis states that epitope mapping antibodies isolated from patient-derived antibody-associated virus (AAVs) complexes can be used to identify conserved HCV epitopes.

In chronic HCV infections, the body produces antibodies against HCV. However, the host Abs are inefficient and further drive the evolution of the circulating HCV quasispecies population. HCV variants resistant to the current iteration of host Abs become the dominant variant in the population through selective pressure, and the infection perpetuates. The host antibody response is not enough to clear a chronic HCV infection, but host Abs can and do bind to circulating HCV variants. Complexes of antibodies and virus particles are detectable in the blood of patients with chronic HCV infections and are known as AAVs (antibody-associated virus). Examination of these

AAVs – where a host antibody *has* bound to HCV – could identify conserved viral structures and reveal vulnerable HCV epitopes. Using techniques based on previous research carried out by Molecular Virology Research and Diagnostic Laboratory (MVDRL), Department of Medicine, University College Cork, the antibodies could be cleaved from the AAVs and examined. Twenty HCV genotype 3a serum and plasma samples were obtained from the National Virus Registry Laboratory (NVRL) and from the MVDRL. HCV genotype 3a was chosen for this research project given its prevalence (estimated to be 17.9% worldwide), resistance to treatment, and increased risk of causing severe steatosis and HCC [174]. IgG was cleaved and purified from the AAVs isolated from the patient samples. The antibodies were then used for a series of *in vitro* neutralisation experiments on HCV pseudoparticles generated using the HCV H77 isolate. H77 (GenBank: AAB67037.1) is an infectious genotype 1a isolate that has undergone complete genome sequencing. The antibodies that showed the greatest neutralisation potential against the H77 pseudoparticles were selected for epitope mapping. The epitope mapping tested the selected samples against a synthesised H77 E2 glycoprotein structure, and characterised the sites where the antibodies bound to the E2 structure. The epitope mapping procedure identified the conserved epitope $_{532}$NDTDVFLLNNTRPPPLGNWFGC$_{553}$. However, this sequence contains two glycosylation sites (E2N6 and E2N7), with a third glycosylation site (E2N8) positioned nearby. The presence of these glycosylation sites may prevent nAbs from accessing this conserved epitope. This characterisation provided insight into vulnerable epitopes on the HCV E2 glycoprotein and the role of glycosylation in HCV's evasion of the humoral immune response.

2 METHODOLOGY

2.1 MATERIALS

The reagents and materials used during this research project were stored and prepared according to the manufacturer's recommendations.

2.1.1 Reagents

Name	Producer
6X DNA Loading Dye	ThermoScientific
100 mM dNTP Set	Biosciences Ltd.
Ab Buffer Kit	GE Healthcare
Agarose	Sigma-Aldrich
Biotinylated Molecular Weight Marker	Sigma-Aldrich
Bovine Serum Albumin	Sigma-Aldrich
Bright-Glo Luciferase Assay System	Promega
Brilliant Blue	Sigma-Aldrich
Calcium Phosphate Transfection Reagent	Invitrogen
Cell Freezing Media	Sigma-Aldrich
Dulbecco's Modified Eagle Medium – High Glucose	Sigma-Aldrich
Dulbecco's Phosphate-Buffered Saline	Sigma-Aldrich
Ethanol – BioUltra, For Molecular Biology - ≥99.8%	Sigma-Aldrich
Foetal Bovine Serum	Sigma-Aldrich
Gel Loading Dye, Blue (6X)	NEB
GeneRuler 1kb DNA Ladder	ThermoScientific
GeneJET PCR Purification Kit	ThermoScientific
Glo Lysis Buffer (1X)	Promega
HEPES-Buffered Saline	Sigma-Aldrich
IgG (Total) Human Uncoated ELISA Kit	Invitrogen
illustra™ Ready-To-Go™ GenomiPhi™ V3 DNA Amplification Kit	GE Healthcare
Luria-Bertani Broth	Sigma-Aldrich
Luria-Bertani Agar (Lennox)	Sigma-Aldrich
Pfu DNA Polymerase – Recombinant	ThermoScientific

Phosphate Buffered Saline	Sigma-Aldrich
Protein G HP SpinTrap	GE Healthcare
One Shot Top10 Chemically Competent E. coli	Invitrogen
QIAamp Viral RNA Mini Kit	Qiagen
QuikChange Lightning Site-Directed Mutagenesis Kit	Agilent Tech
Random Hexamer Primers	ThermoScientific
SuperScript ® II Reverse Transcriptase	Invitrogen
SYBR ® Safe DNA Gel Stain	Invitrogen
Taq DNA Polymerase – Recombinant	Invitrogen
Trizma® (TRIS Base)	Sigma-Aldrich
Trypan Blue Solution 0.4%	Sigma-Aldrich
Trypsin-EDTA Solution	Sigma-Aldrich
Water – Molecular Grade	Sigma-Aldrich

2.1.2 Patient Samples

HCV genotype 3a samples were sourced from the National Virus Registry Laboratory (NVRL), Dublin and from the Molecular Virology Research and Development Laboratory (MVDRL), Cork. Patient samples were consented for research purposes. Patient information was appropriately anonymised before being transferred to this project and any information was stored securely.

Sample ID	Date of Collection	Source	Viral Load (Log10)	AAV Status
A	10/11/2016	NVDRL	5.84	-
B	11/01/2016	NVDRL	6.02	-
C	08/02/2016	NVDRL	5.12	+
D	05/12/2015	NVDRL	4.79	-
E	07/11/2016	NVDRL	5.59	-
F	08/02/2016	NVDRL	6.33	-
G	24/01/2016	NVDRL	4.19	+
H/J	13/09/2016	NVDRL	6.85	+
I/K	13/09/2016	NVDRL	5.96	-
L	24/08/2016	NVDRL	6.72	-

M	20/09/2016	NVDRL	6.24	+
N	20/09/2016	NVDRL	5.35	-
O	25/10/2016	NVDRL	5.44	-
246	31/03/2011	MVDRL	6.57	-
859	28/01/2010	MVDRL	4.70	-
425	01/07/2010	MVDRL	4.91	+
718	03/12/2009	MVDRL	6.28	-
690	06/11/2009	MVDRL	5.18	-
502	16/06/2011	MVDRL	6.74	+
052		MVDRL	~6.00	+

2.1.3 Oligonucleotide Primers

Primer Name	Tm (°C)	Sequence
OR(I-II)	53.7	5'-ATGGCATGGGATATGAT-3'
OF (I)	47.3	5'-AAGGCCGTCCTGTTGA-3'
IF(I)	49.9	5'-GCATGGGATATGATGATGAA-3'
IR(I-II)	52.4	5'-GTCCTGTTGATGTGCCA-3'
Core FWD	64.4	5'-CTTGTGGTACTGCCTGATAGGGTG-3'
Core REV		5'-CCARYTCATCATCATRTCCCANGCCA-3'
REV Spec2	59.0	5'-GGTTCTTGTCCCGGCCTGTGAGG-3'
HCV Mod FWD	66.7	5'-CACCATGGGTTGCTCYTTYTCTATCTTCC-3'
HCV pcDNA REV	61.4	5'-TTATATCATBAGCATCARCCARARRGC-3'
NSREV1	61.4	5'-AGYTCTRTBAGYCCYGCDGC-3'
NSREV2	58.0	5'-YTYGCDACRGCRAGYTCTRT-3'
E2REV1	54.2	5'-HTTCAGRYTGACBAGYTTCT-3'
E2REV2	59.3	5'-BAGYTTCTCCAHYGCRGCRT-3'

2.1.4 Plasmids

MLV gag-pol packaging vector and phCMV-ΔC/E1/E2 H77 constructs, used to produce HCV pseudoparticles, were acquired through a previous research project at the MVDRL, and originated from a material transfer agreement with Dr. Francois Louis-Cosset, INSERM (Institut National de la Santé et de

la Recherche Médicale), France. The MLV luciferase reporter vector was also acquired through a previous research project at the MVDRL, and was a kind gift from Dr. Arvind Patel, CVR (Centre for Virus Research), University of Glasgow, Scotland.

2.2 Agarose Gel Electrophoresis

Products from PCRs were analysed using the agarose gel electrophoresis technique. Agarose gel electrophoresis uses the principle that nucleic acid molecules possess a negative charge, and molecules with negative charges tend to migrate towards a positive charge. The agarose gel matrix means that the nucleic acid molecules are separated out by molecular weight during this migration to the positive charge. A 1% agarose gel was prepared by microwaving 0.5 g agarose with 50 ml 1X tris-acetate ethylene diamine tetra acetic (TAE) buffer. 50X TAE buffer was prepared using 40 mM Trizma® base (Sigma-Aldrich), 20 mM glacial acetic acid, and 1 mM ethylene diamine tetra acetic acid (EDTA), and then appropriately diluted using dH_2O to produce 1X TAE buffer. SYBR Safe DNA Gel Stain (ThermoScientific) was added prior to casting the 1% agarose gel at 2 µl per 50 ml – the SYBR Safe DNA Gel Stain is necessary for visualising the products under UV. The products were analysed next to either the GeneRuler™ 100bp DNA Ladder (100 – 3,000bp) or the GeneRuler™ 1kb DNA Ladder (250 – 10,000bp), depending on the expected size of the product. 10 µl of the PCR product were added to 2 µl 6X DNA Loading Dye (ThermoScientific), mixed thoroughly, and loaded onto the agarose gel. The power supply was set at 80 - 100 V and 80 - 100 mA for 35 – 45 minutes. The products were visualised and photographed using the AnalytikJena UV/White transilluminator and the UVP VisionWorks® software.

2.3 Protein G Column Affinity Chromatography

The Protein G HP SpinTrap™ (GE LifeSciences) columns were used to screen the samples for AAV and for the purification of VF-FAbs. The columns are based on the principle of affinity column chromatography - the

protein G columns have a high binding capacity for the Fc region of IgG from a variety of species. The column membrane captures the IgG, and the antibody can be eluted once the remaining components have been washed through the column. The column must be shaken vigorously to re-suspend the storage medium before starting. The end closure was removed, and the storage medium removed by centrifugation for 30 seconds at 70 – 100 x g. 600 µl binding buffer was added to the column, and removed by centrifugation. A maximum of 600 µl of the serum or plasma sample was added to the column and incubated at room temperature for 4 minutes on the end-over-end rotator. Centrifugation was used to remove the sample once the incubation period was complete. 5 – 8 column washes were carried out using 300 µl binding buffer. The final wash was saved to confirm that there was no remaining virus being washed through the column before the elution step. Once the washes were completed, 400 µl elution buffer was added to the column, and the column placed into 1.5 ml microcentrifuge tube containing 30 µl neutralising buffer. The neutralising buffer contained sodium azide to preserve the elute and prevent contamination. The column was incubated at room temperature for 1 minute and centrifuged for 1 minute at 70 – 100 x g to elute the desired sample. The elution step was repeated to extract the remaining antibody. The elutes were stored at -20 °C.

2.4 vRNA Extraction

vRNA isolation was carried out using the QIAamp Viral RNA Mini Kit (Qiagen). Carrier RNA was used to enhance the binding of target RNA to the QIAamp column, and to reduce vRNA degradation during the denaturing phase. First, the recommended volume of carrier RNA and buffer AVE was added to the recommended volume of buffer AVL. 140 µl of the sample was added to the mixture, mixed by pulse-vortexing, and incubated at room temperature for 10 minutes. This step denatured the sample and inactivated any potential RNases. 560 µl 96 – 100% ethanol was added, mixed by pulse-vortexing, and 630 µl of the mixture applied to the QIAamp column. The QIAamp column uses a silica-based membrane, which has selective binding properties. The column was centrifuged at 8,000 RPM for 1 minute and the

filtrate discarded. This step was repeated, and 500 µl of buffer AW1 was applied to the column. Buffer AW1 and AW2 were washes designed to remove proteins, nucleases, and contaminants from the product. The column was centrifuged at 8,000 RPM for 1 minute, the filtrate discarded, and 500 µl of buffer AW2 was added. The column was centrifuged at 14,000 RPM for 3 minutes, and an optional secondary centrifugation step at 14,000 RPM for 1 minute was included to remove any residual buffer AW2. The QIAamp column was placed into a 1.5 ml microcentrifuge tube, 40 µl of buffer AVE was added, and the column was incubated at room temperature for 1 minute. Buffer AVE contained 0.04% sodium azide to prevent microbial growth and contamination of the RNA product. The column was centrifuged at 8,000 RPM for 1 minute, and the elute from this stage contained the vRNA. A second 40 µl elution using the same conditions was performed to improve the RNA yield. The vRNA was immediately stored at -20°C once eluted.

2.5 cDNA Synthesis

cDNA synthesis from the isolated vRNA was carried out using a previously optimised RT-PCR protocol and using SuperScript ® II Reverse Transcriptase (Invitrogen). 11 µl of vRNA was mixed with 1 µl of random hexamer primer (ThermoScientific) and incubated at 75°C for 10 minutes. The tubes were quickly added to “slush ice” as 8 µl of the reaction mixture (Table 3) was added. Keeping the tubes cold ensured that the RT-PCR did not progress any further during the brief interruption. The samples were returned to the thermal cycler and incubated at 42°C for 60 minutes, and the reverse transcriptase was inactivated by incubation at 95°C for 3 minutes.

Table 3 – Components and Reaction Parameters for cDNA Synthesis - Per One Reaction

Reaction Component	Volume (µl)
5X First Stand Buffer	4.0
0.1 M DTT	2.0
10 mM dNTP	1.0
RNase OUT	0.25

SuperScript ® II Reverse Transcriptase	0.25
Nuclease-Free H_2O	To 20.0

Temperature (°C)	Duration (Minutes)
75	10:00
42	1:00:00
94	3:00

2.6 320BP REVERSE TRANSCRIPTION POLYMERASE CHAIN REACTION (RT-PCR)

The 320bp RT-PCR was used during the project to verify the presence of HCV and for generating the initial sequence information for the patient samples. RT-PCR involved synthesising cDNA from vRNA (per Section 2.5) and using a nested PCR approach to amplify a 320bp region which spanned the junction of the E1 glycoprotein and the E2 glycoprotein. The nested PCR approach was used to reduce the frequency of non-specific binding. This specific RT-PCR protocol had been previously optimised by the Molecular Virology Research and Diagnostics Laboratory, University College Cork.

Table 4 - Components and Reaction Parameters for the Primary RT-PCR – Per One Reaction

Reaction Component	Volume (µl)
10 pm/µl “OF” FWD Primer	1.5
10 pm/µl “OR” REV Primer	1.5
10 mM dNTP	1.0
10X *Pfu* Buffer	5.0
25 mM $MgSO_4$	1.5
Pfu DNA Polymerase	0.5
cDNA	5.0
Nuclease-Free H_2O	To 50.0

Temperature (°C)	Duration (Minutes)	Cycles

94	3:00	1
94	0:15	
51	0:30	35
72	0:45	
72	7:00	1

Table 5 - Components and Reaction Parameters for the Secondary RT-PCR – Per One Reaction

Reaction Component	Volume (µl)
10 pm/µl “IF” FWD Primer	1.5
10 pm/µl “IR” REV Primer	1.5
10 mM dNTP	1.0
10X *Pfu* Buffer	5.0
25 mM $MgSO_4$	1.5
Pfu DNA Polymerase	0.5
Primary PCR Product	4.0
Nuclease-Free H_2O	To 50.0

Temperature (°C)	Duration (Minutes)	Cycles
94	3:00	1
94	0:15	
53	0:30	35
72	0:45	
72	10:00	1

2.7 CORE PROTEIN PCR

The core protein PCR was used to amplify the HCV core protein for the purposes of phylogenetic analysis. A nested PCR approach was used to reduce the frequency of non-specific binding. The cDNA used as the template for the primary PCR was generated as per Section 2.5. The reagents and parameters for the reactions are given below.

Table 6 - Components and Reaction Parameters for the Core Protein PCR – Per One Reaction

Reaction Component	Volume (µl)
10 pm/µl “All Core” FWD Primer	1.5
10 pm/µl “All Core” REV Primer	1.5
10 mM dNTP	1.0
10X PCR Buffer	5.0
50 mM $MgCl_2$	1.5/1.0
Taq Polymerase	0.5
Target cDNA/Primary PCR Product	5.0
Nuclease-Free H_2O	To 50.0

Temperature (°C)	Duration (Minutes)	Cycles
94	3:00	1
94	0:15	35
60	0:30	
72	1:30	
72	10:00	1

The volume of $MgCl_2$ was 1.5 µl for the primary PCR, and 1.0 µl for the secondary PCR, with the volume of H_2O adjusted accordingly. The same reaction parameters were used for the primary and secondary reactions and can be viewed below.

2.8 SPECTROPHOTOMETRY

Spectrophotometry was used to assess to the quality and quantity of DNA produced from various procedures. The Eppendorf BioPhotometer Spectrophotometer UV/VIS was used for the spectrophotometry portion of this project. The spectrophotometer was first “blanked” using dH_2O, and the samples were diluted with dH_2O at a ratio of 1:10 and analysed. The 260/280nm and 260/230nm values were used to assess the quality of the generated DNA and the presence of protein contamination. A 260/280nm

value of ~1.7 and a 260/230nm value of ~2.2 indicated a clean preparation with minimal contamination.

Spectrophotometry was used on one occasion as an alternative to IgG ELISAs for IgG quantification. The concentration of IgG in solution was determined by substituting the molecular weight of IgG, the extinction coefficient (the measure of how strongly a substance absorbs light at a particular wavelength), and the λ_{max} (the wavelength along the absorption spectrum where a substance has its strongest photon absorption) value into a modified version of the Beer-Lambert Law. The formula is:

$$Concentration = \frac{Absorbance\ at\ \lambda_{max}}{Extinction\ Coefficent\ \times\ Pathlength} \times Molecular\ Weight$$

2.9 PURIFICATION OF PCR PRODUCTS

Before sending the PCR products for sequencing, the products were purified using GeneJET PCR Purification Kit (ThermoScientific) according to the manufacturer's recommended protocol. The binding buffer contained a chaotropic agent to denature proteins and facilitate the binding of DNA to the silica membrane of the GeneJET column. Equal volumes of binding buffer and the PCR product were added to the GeneJET column and centrifuged at 10,000 RPM for 30 seconds. If the PCR product was <500bp, 100 µl isopropanol was added to 100 µl binding buffer and combined with 100 µl of the PCR product before adding to the GeneJET column. Once the elute had been discarded, 700 µl wash buffer was added to remove unincorporated nucleotides, primers, and enzymes. The column was centrifuged at 10,000 RPM for 30 seconds, the elute was discarded, and the column centrifuged at 14,000 RPM for 1 minute to remove any trace amounts of wash buffer. The DNA was eluted using 30 µl elution buffer and centrifuged at 10,000 RPM for 1 minute. The DNA was quantified and qualified by spectrophotometry and stored at -20°C.

2.10 Sequencing and Bioinformatics

Purified PCR products and plasmids were sequenced by Eurofins Genomics (https://www.eurofinsgenomics.eu/), Germany. Sequencing reads were aligned with reference sequences from the HCV Sequence Database (https://hcv.lanl.gov) [175] and the NCBI Viral Genome RefSeq Database (https://www.ncbi.nlm.nih.gov/genome/viruses/), and analysed using the MEGA7 software [176]. The sequencing files were curated using the provided electropherogram from Eurofins Genomics, where needed.

2.11 IgG Enzyme-Linked Immunosorbent Assay (ELISA)

ELISA is an assay technique used for detecting and quantifying peptides, proteins, antibodies, and hormones. The antigen of interest is immobilised on the surface of the plate and complexed with a secondary antibody. This secondary antibody is conjugated with an enzyme that acts as a reporter – such as HRP (horseradish peroxidase) or biotin. The reaction between the antigen and antibody is measured using a microplate spectrophotometer. An IgG ELISA was carried out using the IgG (Total) Human Uncoated ELISA Kit (Invitrogen) to quantify the IgG purified from the genotype 3a patient samples. The Corning™ Costar™ 9018 ELISA plates were coated with 100 µL/well of capture antibody (pre-titrated and purified anti-human IgG mAb) diluted with coating buffer (PBS), sealed, and incubated overnight at 4 °C. The next day the Blocking Buffer was prepared by making a 1:10 dilution of Assay Buffer A Concentrate (20X) with deionised water. The plates were aspirated and washed twice with 400 µl/well Wash Buffer (1X PBS and 0.05% Tween™20). About 1 minute of soaking between washing was used to improve the effectiveness of the washes. Blocking Buffer (250 µl/well) was added and plates incubated for 2 hours at room temperature. The standard (recombinant human IgG – 200 ng/ml upon reconstitution with dH_2O) was prepared and the plates aspirated and washed with Wash Buffer as previously described. To prepare the standard curve, 100 µl Assay Buffer (1:10 dilution of Assay Buffer A Concentrate (20X) with deionised water) was added to the selected wells. 100 µl of the reconstituted standard was added

to A1, mixed by repeated aspiration, and 100 µl of the mixture transferred to B1. This was repeated five times to give a concentration gradient of 100 ng/ml, 50 ng/ml, 25 ng/ml, 12.5 ng/ml and 6.25 ng/ml. F1 contained just Assay Buffer to act as a “zero” for calculation purposes. The samples – including the standard curve – were done in duplicate. The experimental samples – which had been diluted down to 10^{-5} and 10^{-6} using Assay Buffer – were added to 100 µl/well Assay Buffer and the plate sealed and incubated for 2 hours at room temperature on a shaker. The plates were aspirated and washed four times using the previous wash protocol and 100 µl/well diluted detection antibody (pre-titrated HRP-conjugated anti-human IgG mAb, diluted 1:250 with 1X Assay Buffer A) was added. The plates were sealed and incubated for 1 hour at room temperature on a shaker. The plates were aspirated and washed four times using the previous wash protocol and 100 µl/well Substrate Solution (tetramethylbenzidine (TMB)) was added. The plates were incubated for 15 minutes at room temperature and 100 µl/well of Stop Solution (2 N H_2SO_4) was added. The plates were read using a microplate spectrophotometer (Promega GloMax System) to read the plates at 570 nm and 450nm.

Analysis of the procedure was done using Microsoft Excel using the standard antibody to generate a standard curve of absorption against concentration (ng/ml), which enabled the quantification of the samples using linear regression.

2.12 Sodium Dodecyl Sulphate – Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE (sodium dodecyl sulphate – polyacrylamide gel electrophoresis) is a technique used for the separation of proteins based on their molecular weight. SDS is a detergent used to disrupt the tertiary structure of the proteins and reduce them to linear molecules - by applying an electrical current, this enables the proteins to migrate through a polyacrylamide gel based solely on their molecular weight. Without the introduction of SDS, the tertiary structure and charge of the proteins would change the proteins’ rate of migration through the polyacrylamide gel.

The polyacrylamide gel is prepared by making a separating gel and a stacking gel. The stacking gel is used to load the protein samples and ensure that the proteins enter the separating gel simultaneously. The stacking gel (4.6%) is made using 950 µl stacking buffer (0.5M Tris (4X) and 0.4% SDS), 550 µl 30% acrylamide, 50 µl 10% SDS, 27.5 µl 10% APS (ammonium persulphate), 3.75 µl TEMED (tetramethylethylenediamine), and 2.25 ml H_2O. The separation gel (10%) is made using 1.75 ml separating buffer (1.5M Tris (4X) and 0.4% SDS at Ph 8.8), 2.33 ml 30% acrylamide, 75 µl 10% SDS, 41 µl 10% APS, 3.9 µl TEMED, and 2.92 ml H_2O. APS should be added just before pouring the gels, as this starts the polymerisation reaction.

Using the glass plates and the casting apparatus, pour the separating gel to ~2 cm under the top of the shorter glass plate. Place a layer of dH_2O on the top of the separating gel to prevent meniscus formation. The separating gel should take about 30 minutes to solidify at room temperature. Remove the dH_2O and pour the stacking gel to the top of the shorter glass plate. Insert the comb. The stacking gel should take about 1 hour to solidify at room temperature. The comb was removed once the gel had completely solidified.

The plates were transferred to the SDS-PAGE electrophoresis chamber (Bio-Rad PowerPac Basic) and the space between the two plates and approximately a third of the space outside the plates was topped up using running buffer. The running buffer was prepared using 400mM Tris, 190 mM glycine, 0.1% SDS, and dH_2O. To prepare the PAGE for loading, the sample areas were washed out twice using a syringe and running buffer. The samples were prepared using PBS to a quantity of 50 µg protein and a volume of 20 µl. Before loading, the samples were heated to 95°C for 5 minutes to linearise them and enable them to become properly coated with SDS. 4 µl of Sigma Biotinylated Molecular Weight Marker (6,500 – 180,000 Da) was added to the first sample chamber. 15 µl of the samples were loaded with 5 µl Gel Loading Dye, Blue (6X) (New England BioLabs). The loading dye contained bromophenol blue as a standard tracker and added SDS to improve band sharpness. The electrophoresis was run at 50 mA for approximately 45 minutes – however, the time needed depends on the size of the expected product. Larger products take longer to migrate through the

gel. The blue dye should reach the end of the gel before removing the gel from the electrophoresis apparatus. The polyacrylamide gel should be cooled for the duration of the electrophoresis, as excess heat can cause warping throughout the gel, particularly towards the centre where heat dispersion is poorer. This manifests as the typical “frowning” bend. Icepacks designed for the chamber can be used, switching them out as they thaw.

Once the gel was removed from the apparatus, the gel was stained using the Coomassie Blue (Brilliant Blue R, Sigma-Aldrich) staining method. The Coomassie Blue was prepared by adding 0.25 g Brilliant Blue R-250 to 90 ml MeOH and H_2O (50/50 mixture by volume) and 10 ml glacial acetic acid. The solution was passed through Whatman paper to remove any clumps. The prepared Coomassie Blue can be stored at room temperature and reused. The polyacrylamide gel was rinsed with dH_2O and transferred to a container and covered with the prepared Coomassie Blue. The gel was rocked overnight at room temperature to complete the staining. The Coomassie Blue was removed from the container and the gel rinsed using dH_2O. A 10% acetic acid solution was used to de-stain the gel. The gel was covered with the 10% acetic acid solution and rocked at room temperature for 2 – 4 hours. The 10% acetic acid solution was replaced as it became saturated with Coomassie Blue. The de-stained transferred to the UV Trans-Illuminator and photographed.

2.13 Site-Directed Mutagenesis (SDM) and Transformation into XL10-Gold UltraCompetent Cells

SDM is a procedure used to make specific and intentional changes to a DNA sequence by using short oligonucleotide primers to introduce the mutation of choice. The QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Tech) was used to perform SDM during this project. The primers used for SDM were previously designed by Dr. Amruta Naik using the Agilent primer design program: https://www.genomics.agilent.com/primerDesignProgram.jsp. The SDM and transformation using XL10-Gold UltraCompetent Cells were performed according to manufacturer’s instructions. The reaction components and the

thermal cycler conditions can be viewed in Table 7. Once the reaction was complete, 2 µl Dpn I restriction enzyme was applied to the amplification reaction and incubated at 37°C for 5 minutes to digest the parental dsDNA. The XL10-Gold UltraCompetent Cells were defrosted and kept cold, and 45 µl transferred to a cold 14 ml BD Falcon polypropylene round bottomed tube. 2 µl β-mercaptoethanol mix (β-ME) was added and swirled. The tube was incubated on ice for 2 minutes before 2 µl of reaction solution was added. The solution was incubated on ice for a further 30 minutes before being heat shocked at 42°C for 30 seconds. The tube was incubated on ice for 2 minutes before 0.5 ml of preheated NZY^+ broth was added, and the mixture incubated at 37°C for 1 hour, shaking at 225 RPM. 10 – 250 µl of the transformation reaction was plated on LB agar plates containing 50 µg/ml ampicillin, 80 µg/ml X-gal, and 20 mM IPTG, and incubated at 37°C for >16 hours. These plates were designed for colour screening of transformed colonies.

Table 7 - Components and Reaction Parameters for QuikChange Lightning Site-Directed Mutagenesis Reaction

Reaction Component	Volume (µl)
10X Reaction Buffer	5
dsDNA Template	*X* (10 – 100 ng)
Oligonucleotide Primer #1	*X* (125 ng)
Oligonucleotide Primer #2	*X* (125 ng)
dNTP Mixture	1
QuikSolution Reagent	1.5
dd H_2O	To 50.0
QuikChange Lightning Enzyme	1

Temperature (°C)	Duration (Minutes)	Cycles
95	2:00	1
95	0:20	
60	0:10	18
68	0:30/kb Plasmid Length	
68	5:00	1

2.14 MiniPrep

Bacterial colonies grown overnight on LB agar at 37°C were inoculated into 5 ml of LB broth with 50 µg/ml ampicillin. The inoculated cultures were grown overnight at 37°C and shaking at 225 RPM. The bacteria were harvested by centrifugation of the inoculum at 6,000 RPM for 15 minutes at 4°C. The supernatant was discarded and the pellet re-suspended using 250 µl re-suspension solution, which contained RNase A. The re-suspension solution was kept cold before use. The re-suspension was transferred to a 1.5 ml microcentrifuge tube. The bacteria were subjected to lysis by adding 250 µl of lysis solution and were mixed by 4 – 5 inversions. The solution was neutralised, and the lysis stopped by adding 350 µl neutralisation solution and mixed thoroughly by inverting. Debris and chromosomal DNA was removed by centrifugation at 10,000 RPM for 5 minutes, and the supernatant added to a GeneJET spin column. The GeneJET column was centrifuged at 10,000 RPM for 1 minute, washed twice with 500 µl wash solution, and centrifuged again. A further centrifugation at 10,000 RPM for 1 minute was used to remove any residual wash solution. The elution buffer was brought to room temperature and 50 µl was added to the GeneJET column to elute the plasmid DNA. The yield and quality of the plasmid DNA was determined using spectrophotometry. The plasmid preparation was stored at -20°C.

2.15 MaxiPrep

An *E. coli* colony containing the plasmid of choice was grown overnight on a selective agar plate at 37°C. A suitable colony was picked and inoculated into a starter culture of 5 ml LB broth containing 50 µg/ml ampicillin. The starter culture was incubated for 8 hours at 37°C and shaking at 225 RPM, before being used to inoculate a 200 ml volume of LB broth containing 50 µg/ml ampicillin – this was to ensure a sufficient amount of plasmid was obtained for downstream procedures. The 200 ml volume was incubated for 16 – 21 hours at 37°C and shaking at 225 RPM. The MaxiPrep system used

was the PureYield™ Plasmid Maxiprep System (Promega). The inoculum was centrifuged ay 5,000 x *g* for 10 minutes and the supernatant discarded. The bacteria were re-suspended using 12 ml of cell re-suspension solution and mixed vigorously by vortexing. 12 ml of cell lysis solution was added, and the mixture inverted 3 – 5 times. The mixture was incubated for 3 minutes at room temperature before 12 ml of neutralisation solution was added to stop the lysis process. The lysate was then centrifuged at 14,000 x *g* for 20 minutes at room temperature, and the supernatant added to the blue PureYield™ Clearing Column. A vacuum manifold and the two PureYield™ columns (the clearing column stacked on the binding column) were used to draw the lysate through the clearing column. Once the lysate was passed through the clearing column, the clearing column was discarded. 5 ml of endotoxin removal wash was added to the PureYield™ Maxi Binding Column and the vacuum applied. Contaminants were removed by adding 20 ml of column wash to the binding column, and the binding membrane dried by applying the vacuum for 5 minutes once the column wash solution had completely passed through the binding column. Once the membrane was completely dried, the binding column was placed onto a 50 ml Falcon tube and 1.5 ml of nuclease-free H_2O was added to the binding membrane. The column and 50 ml tube were centrifuged using a swinging bucket rotor at room temperature at 2,000 x *g* for 5 minutes. The filtrate was transferred to a 1.5 ml microcentrifuge tube and analysed using spectrophotometry. The plasmid preparation was stored at -20°C.

2.16 illustra™ Ready-To-Go™ GenomiPhi™ V3 DNA Amplification Kit

The illustra™ Ready-To-Go™ GenomiPhi™ V3 DNA Amplification Kit was used during the E1E2 glycoprotein PCR to increase the amount of primary PCR product for the secondary PCR. The GenomiPhi™ kit uses Phi29 DNA polymerase, random hexamer primers, and dNTPs to amplify the template DNA using the isothermal strand displacement methodology. The kit can amplify 10 ng input DNA to ~18 µg DNA. 10 µl 2X denaturation buffer was added to 1 µl of 10 ng template DNA. 9 µl of molecular H_2O was added. The mixture was heated to 95°C for 3 minutes and cooled to 4°C. The Ready-To-

Go GenomiPhi V3 cake (which contains the Phi29 DNA polymerase, random hexamers, and dNTPs) was reconstituted with the denatured template DNA. 20 µl of the template DNA was added to the cake, and the mixture incubated at 30°C for 1.5 hours. The Phi29 DNA polymerase was inactivated by heating the samples to 65°C for 10 minutes. The amplified DNA was cooled to 4°C and stored at -20°C.

2.17 Cell Culture

The Human Embryonic Kidney 293T (HEK 293T) and Human Hepatoma 7 (Huh7) lineages were chosen for the preparation of HCV pseudoparticles. HEK 293T and Huh7 were grown using Dulbecco's Modified Eagle Medium High Glucose (DMEM) (Sigma-Aldrich), supplemented with 10% foetal bovine serum (FBS) and 1% PenStrep – a mixture of 10,000 units of penicillin and 10 mg/ml streptomycin suspended with 0.9% NaCl. The HEK 293Ts were also supplemented with 1% non-essential amino acids (NEAA). The HEKs and Huh7s were grown using T-75 ml flasks at 37°C and 5% CO_2 and passaged once confluency reached 80 – 90%. Once the media had been removed and the internal surface of the flasks washed twice with 10 ml PBS, 2 ml trypsin-EDTA was used to trypsinise the HEKs/Huh7s from the surface of the flask. After 5 minutes of trypsinisation, 8 ml DMEM containing FBS was added to neutralise the trypsin-EDTA, and the solution was centrifuged at 1,500 RPM for 5 minutes. The supernatant was discarded, and the re-suspension used to seed fresh flasks. The passage numbers were tracked and kept under p35 during the project.

Stocks were prepared using 1ml of cell freezing media (Sigma-Aldrich). The HEKs/Huh7s were grown to a confluency of ~90%. The flask was washed, trypsinised, and centrifuged as described, and resuspended using 1 ml of freezing media at the desired density. The stocks were slowly frozen to -80°C overnight using an isopropanol freezing container to prevent damage and the formation of ice crystals, before being transferred to nitrogen storage.

2.18 GENERATION OF HCV PSEUDOPARTICLES

The HCV pseudoparticle (HCVpp) system is a retrovirus model system developed by Bartosch *et al.* and involves transfecting HEK 293Ts with three plasmids - a plasmid containing HCV E1E2 glycoproteins, a gag-pol packaging plasmid, and reporter plasmid, which tends to have a luciferase or GFP gene. The HCV E1E2 glycoproteins enable the pseudoparticle to enter and infect Huh7s – a type of hepatocyte – but the pseudoparticles are unable to replicate.

The transfection of HEK 293Ts with the three plasmids is based upon the $CaPO_4$ transfection method. This technique was pioneered during the 1970s and remains popular because of the simple and inexpensive reagents used. The DNA/plasmid of interest is combined with calcium chloride to generate a $CaPO_4$-DNA/plasmid precipitate. $CaPO_4$ enables the precipitate to bind to the surface of the HEKs, and the HEKs can internalise the DNA/plasmid through endocytosis.

1. **Day One**

   HEKs were seeded at a 1 - 1.5 x 10^6 density using a 10 cm diameter tissue culture plate, with ~15 ml supplemented DMEM (see Section 2.17). The HEKs were incubated for ~16 hours at 37°C at 5% CO_2.

2. **Day Two**

   The HEKs were a minimum of 40% confluency before starting the transfection procedure. The HEKs were transfected using a $CaPO_4$ transfection kit (Invitrogen). The reagents must be kept cold during the transfection process. A “no envelope” control pseudoparticle should be prepared using dH_2O instead of the E1E2 glycoprotein plasmid. For a single pseudoparticle, the components from Table 8 were added to a 1.5 ml microcentrifuge tube. The reagents should be added as per the order of Table 8. 500 µl of 2X HBS (HEPES-Buffered Saline) was added to an empty 50 ml Falcon tube, and the transfection components were added dropwise to the 2X HBS. Using a PipetteBoy, the 2X HBS was bubbled as the transfection components

were added, as the aeration generates a finer $CaPO_4$ precipitate – finer precipitates are associated with improved rates of pseudoparticle uptake by the HEKs. The mixture was vortexed and incubated at room temperature for 20 minutes. The 1 ml of the mixture was then distributed dropwise over the prepared HEK plate and incubated overnight at 37°C at 5% CO_2.

3. **Day Three**

   The medium from the transfection plate was discarded and replaced with 6 ml DMEM (supplemented for HEKs are described during Section 2.17) containing 10 mM HEPES buffer and incubated overnight at 37°C at 5% CO_2.

4. **Day Four**

   The first round of HCVpp was harvested. The media from the plate was taken up using a syringe and passed through a 0.45 µm filter to an empty 15 ml Falcon tube. The HCVpp harvest can be stored at 4°C for two weeks. The HCVpp can also be stored at -80°C, but storage at this temperature causes a 50% reduction of infectivity. The HEK plate was replenished with 6 ml supplemented DMEM containing 10 mM HEPES buffer and incubated overnight at 37°C at 5% CO_2.

5. **Day Five**

   The second harvest of HCVpp are more infectious than the first harvest of HCVpp, and were the pseudoparticles used for the experiments during this project. The second harvest was harvested as stated above and stored at 4°C.

Table 8 - Components required for transfection of HEK 293T cells.

Tissue Culture Dish	**10 cm**	**6 cm**	**3.5 cm**
Area	48 cm^2	21 cm^2	8.5 cm^2
HEK Seeding Density	1 x 10^6	4 x 10^5	2 x 10^5
Luciferase Plasmid	8 µg	3 µg	1.5 µg
pMLV gag-pol	8 µg	3 µg	1.5 µg
E1E2 Glycoprotein	3 µg	1.2 µg	0.5 µg
dH_2O (to)	400 µl	160 µl	80 µl
2M $CaCl_2$	100 µl	40 µl	20 µl
2X HBS	500 µl	200 µl	100 µl
Total Volume	1000 µl	400 µl	200 µl

2.18.1 Infectivity Assay Using HCVpp

Huh7s are seeded the day before the planned infection assay at the densities stated on Table 9. The next day, the volume of harvested HCVpp inoculum stated on Table 9 is added to the Huh7s. The Huh7s are incubated at 37°C for 3 – 4 hours and tapped every 30 minutes during the incubation period to ensure an even spread of pseudoparticles. Once the incubation period has ended, the HCVpp inoculum is removed and replaced with fresh DMEM, supplemented for Huh7s. The amount of DMEM added depends on the type of plate used, but the standard working volumes of the plates should be used. The plates are incubated for 72 - 96 hours at 37°C at 5% CO_2.

2.18.2 Neutralisation Assay Using HCVpp

The Huh7s are seeded as described in Section 2.18.2. Before the HCVpp inoculum is added to the Huh7s, the chosen antibodies are incubated with HCVpp for 1 hour at 37°C – this can be done using an incubator or water bath. The antibodies used for the neutralisation assay should be at a consistent concentration (mg/ml) and equal volumes of antibody and HCVpp should be used. Once the incubation period has been completed, the HCVpp and antibody inoculum should be added to the Huh7s and the Huh7s incubated as described for Section 2.18.1.

2.18.3 Luciferase Assay Using HCVpp

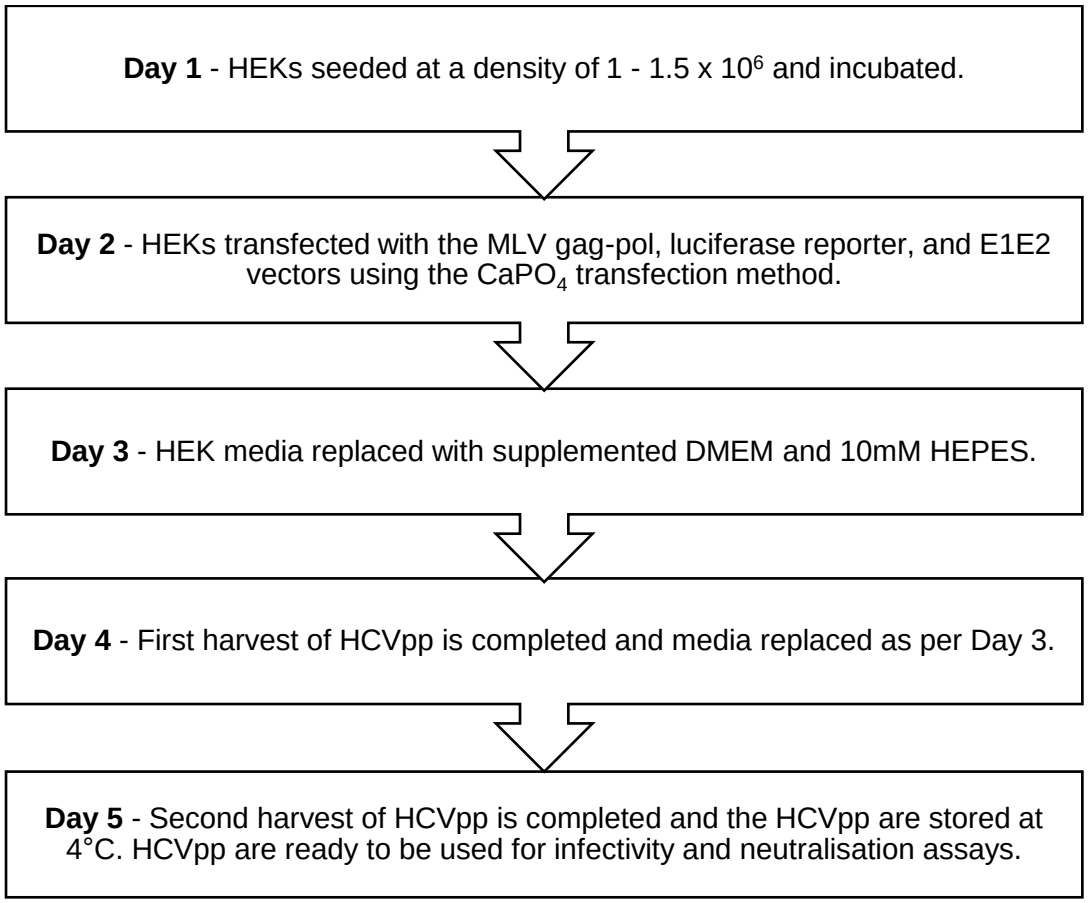
Once the incubation period for the infection/neutralisation HCVpp assay has been completed, the Huh7s were lysed using 1X Lysis Buffer (Promega) for

~20 minutes at room temperature. The plate can also be rocked during the lysis stage. The volume of 1X Lysis Buffer used should match the HCVpp inoculum volume from Table 9. The lysate was transferred to a black luminometer plate and 100 µl BrightGlo Luciferase Substrate (Promega) was added. Relative light unit (RLU) emission was measured using a luminometer (Promega GloMax System).

Table 9 - Infection parameters for infectivity assays using HCVpp.

Tissue Culture Dish	6-Well	12-Well	24-Well	48-Well	96-Well
Area	8.5 cm^2	3.8 cm^2	2.0 cm^2	0.8 cm^2	0.25 cm^2
Huh7 Density	1 x 10^5	4.5 x 10^4	2.5 x 10^4	1.0 x 10^4	4.0 x 10^3
HCVpp Inoculum	600 µl	300 µl	150 µl	90 µl	50 µl

Figure 7 - Summary of process for generating HCV pseudoparticles.



2.19 EPITOPE MAPPING

Epitope mapping is the process of identifying the binding sites (epitopes) of antibodies to their corresponding antigens. Epitopes can be categorised as linear epitopes or conformational epitopes. Linear epitopes are formed by continuous sequences of amino acids, whereas conformational epitopes are composed of discontinuous sequences that are brought together by the tertiary structure of the protein. For this project, the epitope mapping was carried out by Pepscan Presto B. V. using the existing H77 E2 glycoprotein chemically-linked peptides on scaffolds (CLIPS) technology platform from Dr. Amruta Naik's epitope mapping studies.

Sequence Used

#H77_E2

THVTGGSAGRTTAGLVGLLTPGAKQNIQLINTNGSWHINSTALNCNESLNT
GWLAGLFYQHKFNSSGCPERLASCRRLTDFAQGWGPISYANGSGLDERP
YCWHYPPRPCGIVPAKSVCGPVYCFTPSPVVVGTTDRSGAPTYSWGAND
TDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGGVGNNTLLCPT
DCFRKHPEATYSRCGSGPWITPRCMVDYPYRLWHYPCTINYTIFKVRMYV
GGVEHRLEAACNWTRGERCDLEDRDRSELSPLLLSTTQWQVLPCSFTTLP
ALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEYVVLLFLLLADARVCSCL
WMMLLISQAEA

The linear and CLIPS peptides were synthesized based on the amino acid sequence of H77 E2 glycoprotein using Fmoc (fluorenylmethyloxycarbonyl protecting group) chemistry and de-protected using trifluoric acid with scavengers. The constrained peptides were synthesized using CLIPS technology to reconstruct the conformational epitopes. The side-chains of the multiple cysteines in the peptides are coupled to CLIPS templates by reacting onto the polypropylene PEPSCAN cards with a 0.5 mM solution of CLIPS template such as 1,3-bis (bromomethyl) benzene in ammonium bicarbonate (20 mM, pH 7.9)/acetonitrile (1:1(v/v)). The cards were shaken for 30 - 60 minutes while completely covered by the solution. The cards were washed thoroughly with H_2O and sonicated in disrupt-buffer containing 1% SDS/0.1% betamercaptoethanol in PBS (pH 7.2) at 70°C for 30 minutes,

followed by sonication in H_2O for another 45 minutes. The binding of antibody to each peptide was tested using a PEPSCAN-based ELISA. The polypropylene cards containing the peptides were incubated with primary antibody solution consisting of 1 µg/ml diluted with a blocking solution (4% horse serum, 5% ovalbumin (w/v) in PBS/1% Tween, for example). The card was washed and the peptides were incubated with a 1/1000 dilution of antibody peroxidase conjugate for 1 hour at 25°C. The card was washed, and the peroxidase substrate 2,2'-azino-di-3-ethylbenzthiazoline sulfonate (ABTS) and 2 µl of 3% H_2O_2 are added, and the card incubated for 1 hour. The colour development was quantified with a charge coupled device (CCD) camera and an image processing system.

3 RESULTS

3.1 WORK PERFORMED ON NAIK *ET AL.* PAPER

3.1.1 Hypothesis

The hypothesis, proposed by Dr. Amruta Naik, was that a single amino acid change to the E1E2 glycoprotein would be enough to increase or decrease the infectivity of a HCV pseudoparticle. Two plasmids (1b-1-2 and 1b-1-3) coding for the E1E2 glycoprotein were generated from two serum samples from a single patient, separated by six months. However, the second plasmid (1b-1-3) was found to have amino acid changes at position 292 at the E1 glycoprotein, and positions 388 and 395 at the HVR1 domain. When HCV pseudoparticles were generated using the two plasmids, there was found to be significant change to the infectivity, with 1b-1-2 being highly infectious and 1b-1-3 being only moderately infectious. It was proposed that the amino acid changes could be the cause of this.

This work contributed to the paper:

"*Reverse epitope mapping of the E2 glycoprotein in antibody associated hepatitis C virus*" Amruta S. Naik, Ania Owsianka, Brendan A. Palmer, Ciaran J. O'Halloran, Nicole Walsh, Orla Crosbie, Elizabeth Kenny-Walsh, Arvind H. Patel, Liam J. Fanning – PLOS ONE, May 30th, 2017.

3.1.2 Methodology

The following methodologies were used during the course of this experiment. Please refer to the named sections for further information on the protocols used.

- Site-Directed Mutagenesis and Transformation (Section 2.13)
- MiniPrep (Section 2.14)
- MaxiPrep (Section 2.15)
- Generation of HCV Pseudoparticles (Section 2.18)
- Infectivity Assay Using HCV Pseudoparticles (Section 2.18.1)

For the purpose of this study, a number of mutant HCVpp were generated, using single amino acid changes to verify whether that particular amino acid change could influence the infectivity of the HCV pseudoparticle. The

mutations were introduced using site-directed mutagenesis. Once the site-directed mutagenesis was carried out, the plasmids underwent transformation, MiniPrep, and MaxiPrep to prepare them for the generation of the HCVpp. Once the HCVpp were generated, infectivity assays were carried out as quadruplets and using three technical replicates to verify the data.

The mutants generated were:

1b-1-2$_{\mathbf{I388V}}$

1b-1-2$_{\mathbf{V395A}}$

1b-1-2$_{\mathbf{T292A+I388V}}$

1b-1-2$_{\mathbf{T292A+V395A}}$

1b-1-2$_{\mathbf{DM}}$

1b-1-2$_{\mathbf{T292A+DM}}$

Regarding the notation of “I388V”, “I” indicates that the original amino acid at position 388 on plasmid 1b-1-2 was changed from isoleucine to valine – “V”. “DM” refers to the double mutation of 1389V and V396A. The generation of the mutants and the infectivity assays were carried out by Amruta Naik. Ciaran O’Halloran, Nicole Walsh, and Kian Hearty.

3.1.3 Results

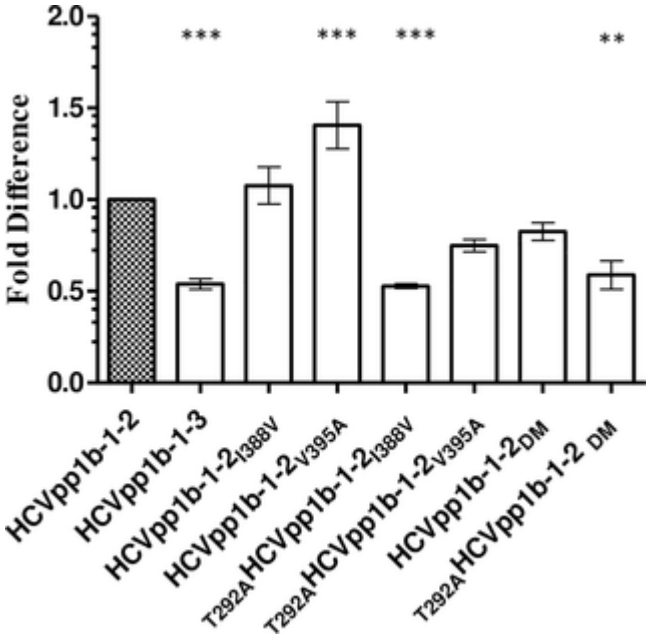


Figure 8 - Changes to three single amino acids were found to influence the infectivity of HCVpp. The X axis represents the infectious HCVpp generated for the study, and the Y axis

represents the percentage of infectivity. The unaltered 1b-1-2 and 1b-1-3 plasmids serve as comparisons for the mutant HCVpp. [177]

3.1.4 Discussion

It had been observed that HCVpp 1b-1-2 was, on average, twice as infectious as HCVpp 1b-1-3, despite both the E1E2 glycoproteins for these pseudoparticles being sourced from the same patient. The two samples were taken six months apart, and 1b-1-3 was shown to have three amino acid changes from 1b-1-2. It was hypothesised that these amino acid changes were responsible for the reduced infectivity of 1b-1-3.

It was found that the amino acid changes did contribute to the infectivity of the two plasmids, as demonstrated by Figure 8. One of the mutants was generated to copy two of the amino acid changes seen throughout the original 1b-1-3, with the hypothesis being that plasmid 1b-1-2$_{\mathbf{T292A+I388V}}$ and the unaltered 1b-1-3 should have the same percentage of infectivity, given that the amino acid sequences are identical – this was shown to be the case, and confirmed that the threonine residue at position 292 and the isoleucine residue at position 388 were important for infectivity.

The statistical significance of the HCVpp infectivity when compared to the original HCVpp 1b-1-2 was determined by using one-way analysis of variance (ANOVA) and Dunnet's Test. The statistical analysis showed that the amino acid at position 292 (the E1 glycoprotein) appears to have some impact on the degree of infectivity *in vitro*, and that the T292A 1b-1-2 DM clone had the closest infectivity profile to the HCVpp 1b-1-3 wildtype.

The information generated from the study showed the importance of single amino acid residues and positions for HCV infectivity, and demonstrated the degree of change an HCV quasispecies population can undergo.

The main work carried out by Nicole Walsh was part of the SDM performed on the mutants to determine whether these amino acid substitutions were related to the increased infectious capacity of 1b-1-2. However, the epitope mapping portion of this paper ("*Reverse epitope mapping of the E2 glycoprotein in antibody associated hepatitis C virus*") was the basis for the epitope mapping of this project. Amruta Naik isolated antibodies from AAV+ samples using a proteinase K digestion and purified these samples using

Protein G columns and created the protocol for the pseudoparticle assays carried out during this project. During Section 3.5, the epitopes found by Amruta Naik and mentioned during this paper are compared to the putative epitopes that were discovered during this paper.

3.2 Isolation and Amplification of the HCV E1E2 Glycoprotein

3.2.1 Hypothesis

For the generation of HCV pseudoparticles, a vector containing the HCV E1E2 glycoprotein is needed – this is what enables the pseudoparticle to enter the target. The goal of this experiment was to isolate the E1E2 glycoprotein from one of the 3a HCV serum or plasma samples obtained from the NVRL using PCR techniques. Once the glycoprotein had been isolated, the glycoprotein would be added to an appropriate vector and used for the pseudoparticle experiments – see Section 3.4.

3.2.2 Methodology

The following methodologies were used during the course of this experiment. Please refer to the named sections for further information on the protocols used.

- vRNA Extraction (Section 2.4)
- 320bp RT-PCR (Section 2.6)
- Agarose Gel Electrophoresis (Section 2.2)

The primers below were used during this experiment. These primers were obtained from Eurofins Genomics:

Primer Name	Tm (°C)	Sequence
OR(I-II)	53.7	5'-ATGGCATGGGATATGAT-3'
OF (I)	47.3	5'-AAGGCCGTCCTGTTGA-3'
IF(I)	49.9	5'-GCATGGGATATGATGATGAA-3'
IR(I-II)	52.4	5'-GTCCTGTTGATGTGCCA-3'
Core FWD	64.4	5'-CTTGTGGTACTGCCTGATAGGGTG-3'
Core REV		5'-CCARYTCATCATCATRTCCCANGCCA-3'
REV Spec2	59.0	5'-GGTTCTTGTCCCGGCCTGTGAGG-3'
HCV Mod FWD	66.7	5'-CACCATGGGTTGCTCYTTYTCTATCTTCC-3'
HCV pcDNA REV	61.4	5'-TTATATCATBAGCATCARCCARARRGC-3'
NSREV1	61.4	5'-AGYTCTRTBAGYCCYGCDGC-3'
NSREV2	58.0	5'-YTYGCDACRGCRAGYTCTRT-3'
E2REV1	54.2	5'-HTTCAGRYTGACBAGYTTCT-3'
E2REV2	59.3	5'-BAGYTTCTCCAHYGCRGCRT-3'

3.2.3 Results

At the start of this experiment, the primers generated by Dr. Amruta Naik were used, as these had been used to isolate the E1E2 glycoprotein from a number of HCV genotypes. A nested PCR approach was used, and the outer primers were Core FWD and REV Spec2. The inner primers were HCV Mod FWD and HCV pcDNA REV, which was a designed for genotype 3a. The start of the HCV Mod FWD primer included a CACC sequence at the 5’ end to support directional cloning using the TOPO group of cloning vectors from Invitrogen. TOPO cloning refers to the technique of cloning DNA fragments to vectors without the need for DNA ligases – instead, the inherent properties of DNA topoisomerase I is used to cleave and ligate DNA ends.

Despite Dr. Naik’s previous success using the above primers, we were unable to isolate the E1E2 glycoprotein using the protocol. Instead, using a genotype 3a reference sequence, the primers NSREV1, NSREV2, E2REV1, and E2REV2 were designed. A number of experiments with various combinations of primers and PCR conditions were carried out – these can be viewed below.

Reaction Component	Volume (µl)
10 pm/µl Core FWD Primer	1.5
10 pm/µl NSREV1/NSREV2 Primer	1.5
10 mM dNTP	1.0
10X *Pfu* Buffer + $MgSO_4$	5.0
Pfu DNA Polymerase	0.5
cDNA	5.0
Nuclease-Free H_2O	To 50.0

Temperature (°C)	Duration (Minutes)	Cycles
95	3:00	1
95	0:30	
58	0:30	35
72	3:00	
72	10:00	1

Table 10 - The conditions for the primary PCR.

Reaction Component	Volume (µl)
10 pm/µl HCV Mod FWD Primer	1.5
10 pm/µl E2REV1/E2REV2 Primer	1.5
10 mM dNTP	1.0
10X *Pfu* Buffer + $MgSO_4$	5.0
Pfu DNA Polymerase	0.5
Primary PCR Product	5.0
Nuclease-Free H_2O	To 50.0

Temperature (°C)	Duration (Minutes)	Cycles
95	3:00	1
95	0:30	
58	0:30	35
72	2:00	
72	10:00	1

Table 11 - The conditions for the secondary PCR.

The primary PCR was tested using the secondary PCR reagents and conditions from the 320bp RT-PCR procedure and agarose gel electrophoresis to verify that there was product present. With the combinations of the Core FWD and NSREV1 primers, the 320bp secondary PCR produced a clear band, implying that product had been isolated during the primary PCR. However, various combinations of the secondary PCR primers could not produce a product. Hoping to increase the amount of primary PCR product present – to improve the odds of the secondary PCR product being produced – the illustra Ready-To-Go GenomiPhi V3 DNA Amplification Kit (GE LifeSciences) was used on the primary PCR product. The protocol for the GenomiPhi DNA Amplification Kit can be found at Section 2.16. Unfortunately, the amplification of the primary PCR product did not enable the isolation of the secondary PCR product.

3.2.4 Discussion

We were unable to isolate the E1E2 glycoprotein through PCR. There are a few explanations for why the E1E2 glycoprotein was hard to isolate. First and foremost, the HCV E1E2 glycoprotein contains three hypervariable regions (HVR1, HVR2, and IgVR). The extreme genetic diversity that the E2 glycoprotein demonstrates complicates designing functional primers for PCR. The primary PCR primers were designed for the core protein and NS2 protein – both of which are more conserved across HCV genotypes and subtypes – and were able to generate a primary PCR product. The stable nature of the HCV core protein meant that the Core FWD primer did not need any degenerate (“wobble”) bases. However, the NSREV1 and NS2REV2

primers did include degenerate bases. The primers were designed using the Eurofins Genomics primer design program (*https://www.eurofinsgenomics.eu/en/ecom/tools/pcr-primer-design/*) and a reference HCV genotype 3a sequence. The HCV Mod FWD primer was used as the forward primer because it contained the CACC sequence for TOPO cloning, and the Eurofins Genomics primer design program suggested the E2REV1 and E2REV2 reverse primers. Both E2REV1 and E2REV2 contained many degenerate bases because the N-terminal end of the E2 glycoprotein contains the HVR1. The large number of degenerate bases, and the presence of the hypervariable regions, could be one explanation for why the secondary PCR did not work.

A future option for the isolation of the E1E2 glycoprotein for HCV pseudoparticle production could be using a MinION DNA sequencer (Oxford Nanopore Technology) to determine the sequence of the E1E2 glycoprotein. Once the sequence is determined, a plasmid construct carrying the E1E2 glycoprotein could be synthesised. A MinION unit was obtained by the University College Cork Department of Medicine and is a portable DNA and RNA sequencer capable of sequencing 10 – 20 Gb DNA/RNA within 48 hours. Synthesising the plasmid with the E1E2 glycoprotein would remove the obstacle of cloning experiments and would enable a quicker progression to pseudoparticle production. However, depending on how the glycoprotein insert and the plasmid are synthesised, there would be the risk that some of the tertiary protein structures would not be the same as the original virus.

3.3 Phylogenetic Analysis of the 320bp Region and Core Protein

3.3.1 Hypothesis

The 320bp RT-PCR was optimised by the MVDRL and served as a useful measure for verifying the presence or absence of HCV for a variety of procedures. However, the 320bp region also covers an area of interest regarding the HCV genome – the junction between the E1 and E2 glycoproteins. As previously mentioned, the E2 glycoprotein contain three HVR regions – HVR1, HVR2, and the intergenotypic variable region. Because of the high degree of genetic diversity displayed by the E2 glycoprotein, any sequences including the E2 glycoprotein could be used to identify a given sample. This would help confirm the genotype of the sample – by using BLAST to match the sample against other HCV 3a genomes – and could be used to start the process of phylogenetic analysis. Despite the fact the samples came from the NVRL, it cannot be assumed that the samples came from a single population or group of patients. The phylogenetic analysis would serve to check the relatedness of the samples and how the samples relate to groups of reference samples, both HCV 3a references, and references from other genotypes and subtypes.

The HCV core protein forms the viral capsid, which protects the vRNA as the virus moves throughout the host. The immature form of the core protein requires cleavage by several host signal peptidases. Polymorphisms of the HCV core protein have been associated with poor treatment responses. This experiment – conducted by Mr. Jack Kenny and Dr. Brendan Palmer – involved sequencing the core protein of the NVRL samples and carrying out phylogenetic analysis to examine the genetic variations between those samples. The core protein from the Pyro2 project samples had already been sequenced by Dr. Brendan Palmer.

3.3.2 Methodology

The following methodologies were used during the course of this experiment. Please refer to the named sections for further information on the protocols used.

- vRNA Extraction (Section 2.4)
- 320bp RT-PCR (Section 2.6)

- Core Protein PCR (Section 2.7)
- Agarose Gel Electrophoresis (Section 2.2)
- Purification of PCR Products (Section 2.9)
- Spectrophotometry (Section 2.8)
- Sequencing and Bioinformatics (Section 2.10)

RNA was extracted from the serum samples obtained from the NVRL and MVDRL. The 320bp RT-PCR procedure was carried out and the PCR products purified and sent for sequencing once the agarose gel electrophoresis confirmed that the 320bp region has been properly amplified. The samples underwent sequencing at Eurofins Genomics and the sequences curated using the provided electropherogram. MEGA 7 was the bioinformatics software used. Based on the secondary PCR primers and the HCV H77 1a genome on which they were based, the 320bp region covers from nucleotide position 1316 to position 1749, with most of the sequences ending up being 320 – 300bp once primer residues were trimmed from the sequences.

Primer Name	Tm (°C)	Sequence
OR(I-II)	53.7	5'-ATGGCATGGGATATGAT-3'
OF (I)	47.3	5'-AAGGCCGTCCTGTTGA-3'
IF(I)	49.9	5'-GCATGGGATATGATGATGAA-3'
IR(I-II)	52.4	5'-GTCCTGTTGATGTGCCA-3'

Table 12 - Oligonucleotide primers used for the 320bp RT-PCR procedure. The primers were created based on the HCV "H77" genotype 1a genome.

Based on the core protein PCR primers and the HCV H77 1a genome on which they were based, the core protein primers covered from nucleotide position 304 to position 1293, with most of the sequences ending up being ~500bp once the primers and degraded residues were trimmed from the sequence files.

Primer Name	Tm (°C)	Sequence
Core FWD	64.4	5'-CTTGTGGTACTGCCTGATAGGGTG-3'
Core REV		5'-CCARYTCATCATCATRTCCCANGCCA-3'

Table 13 - Oligonucleotide primers used for the core protein PCR procedure. The primers were created based on the curated HCV "H77" genotype 1a genome.

3.3.3 Results

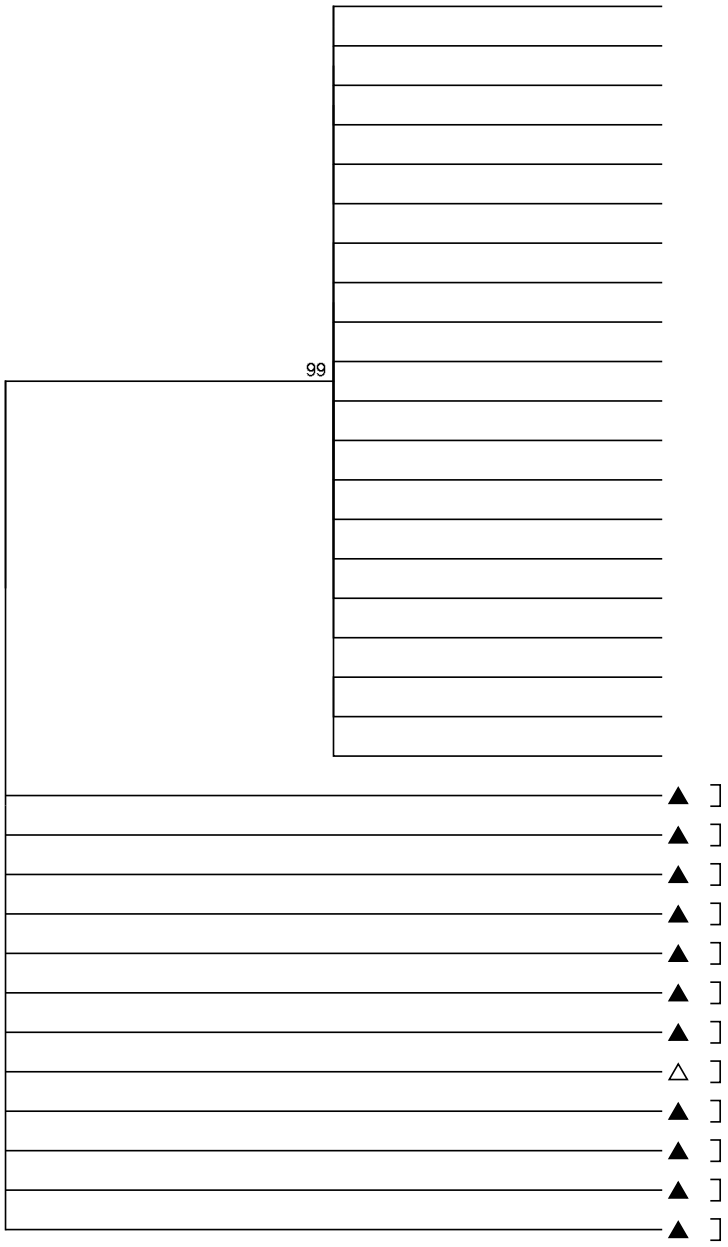


Figure 9 - Phylogenetic analysis of the 320bp samples and using 20 genotype 3a reference sequences submitted by the MVDRL. ▲ indicates samples from this study, and △ indicates the reference HCV 3a 320bp sample. The evolutionary history was from 500 replicates is taken to represent the evolutionary history of the taxa analysed. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The analysis involved 32 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 304 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. The sequences were aligned using ClustalW.

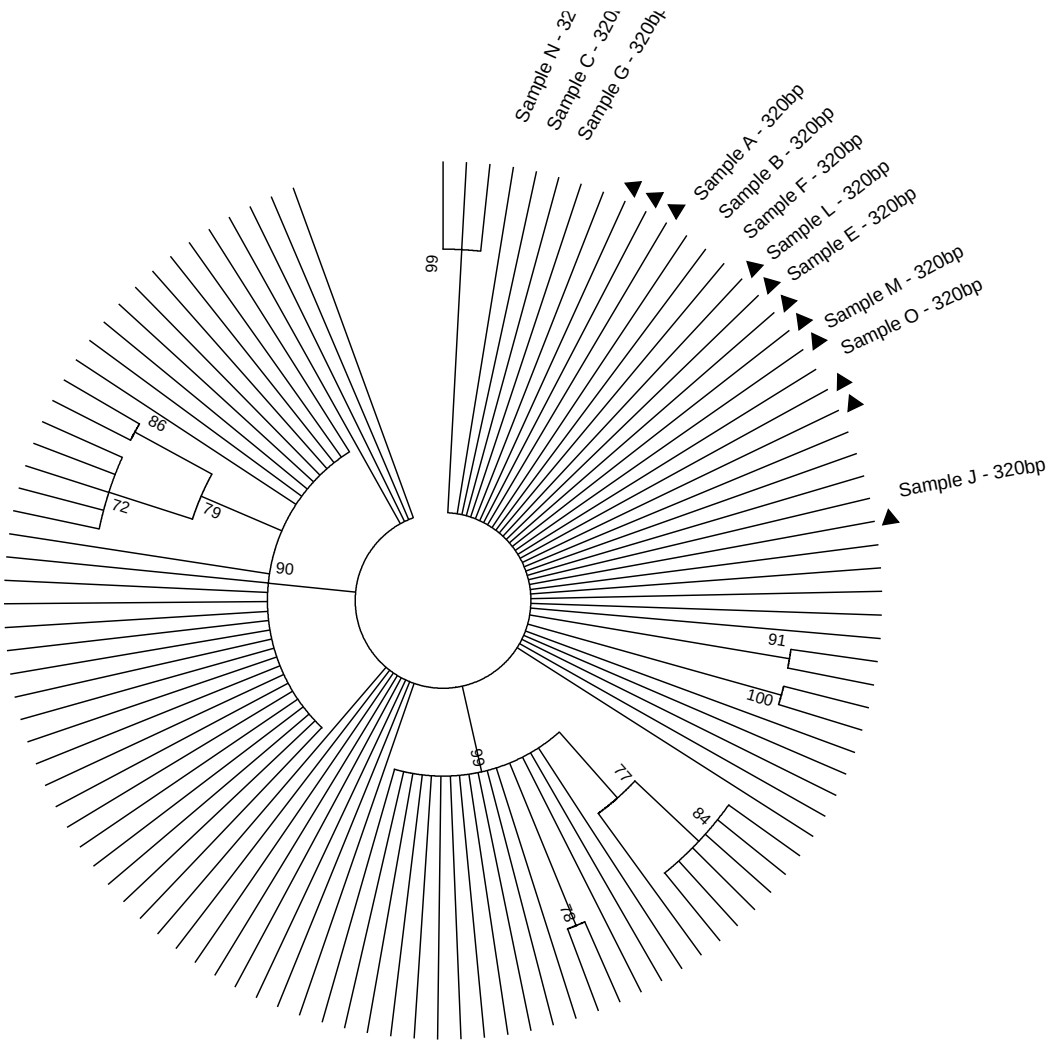


Figure 10 - Phylogenetic analysis of the 320bp samples when aligned with a number of HCV reference sequences across various genotypes. ▲ indicates samples from this study. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The analysis involved 112 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 291 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. The sequences were aligned by ClustalW.

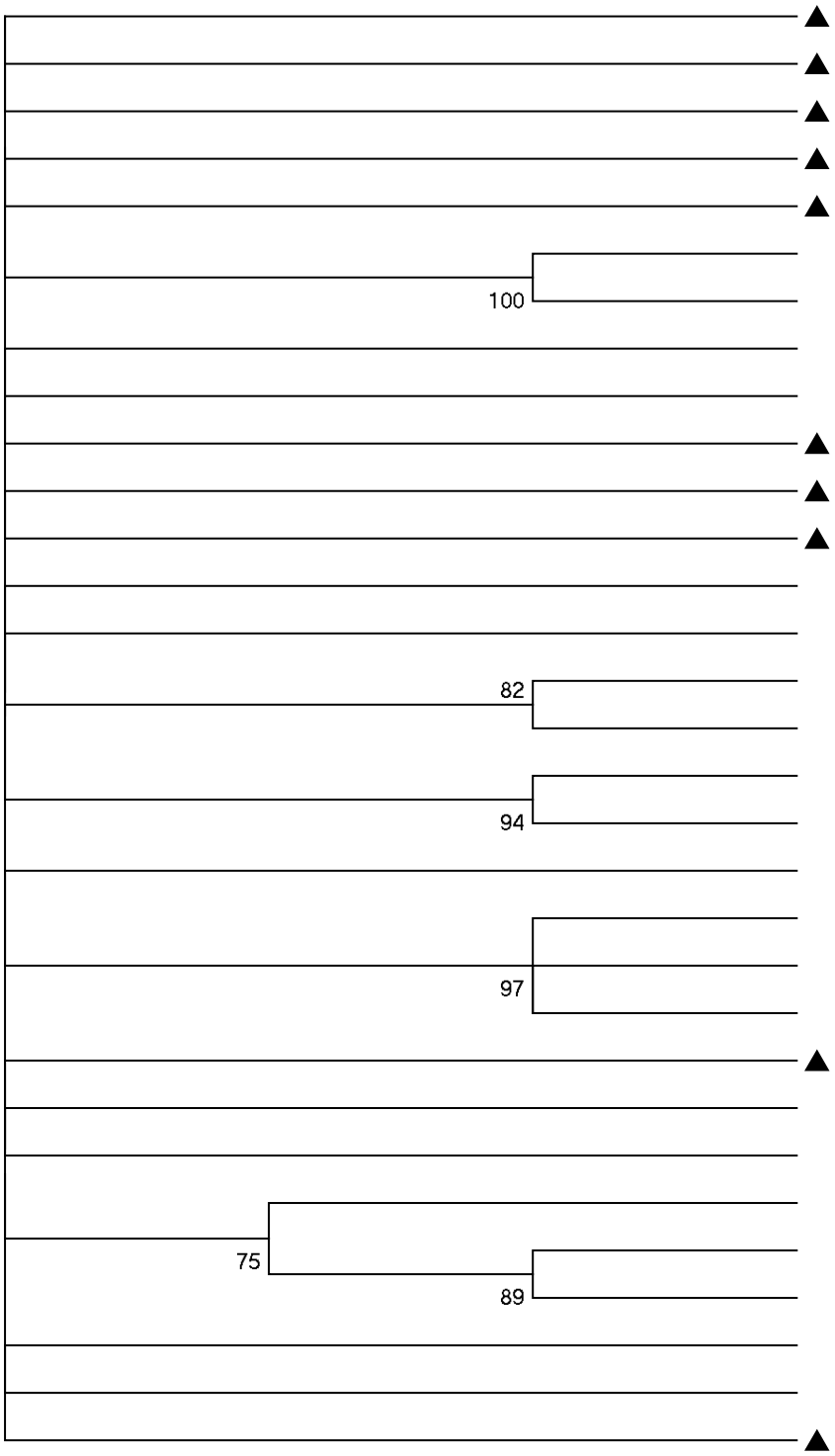


Figure 11 - Phylogenetic analysis of the core samples and 20 genotype 3a core sequences. ▲ indicates samples from this study. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The bootstrap consensus tree inferred from 500 replicates [3] is taken to represent the evolutionary history of the taxa analysed]. Branches corresponding to partitions reproduced in less than 70% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log

likelihood value. The analysis involved 31 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 521 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. The sequences were aligned by ClustalW.

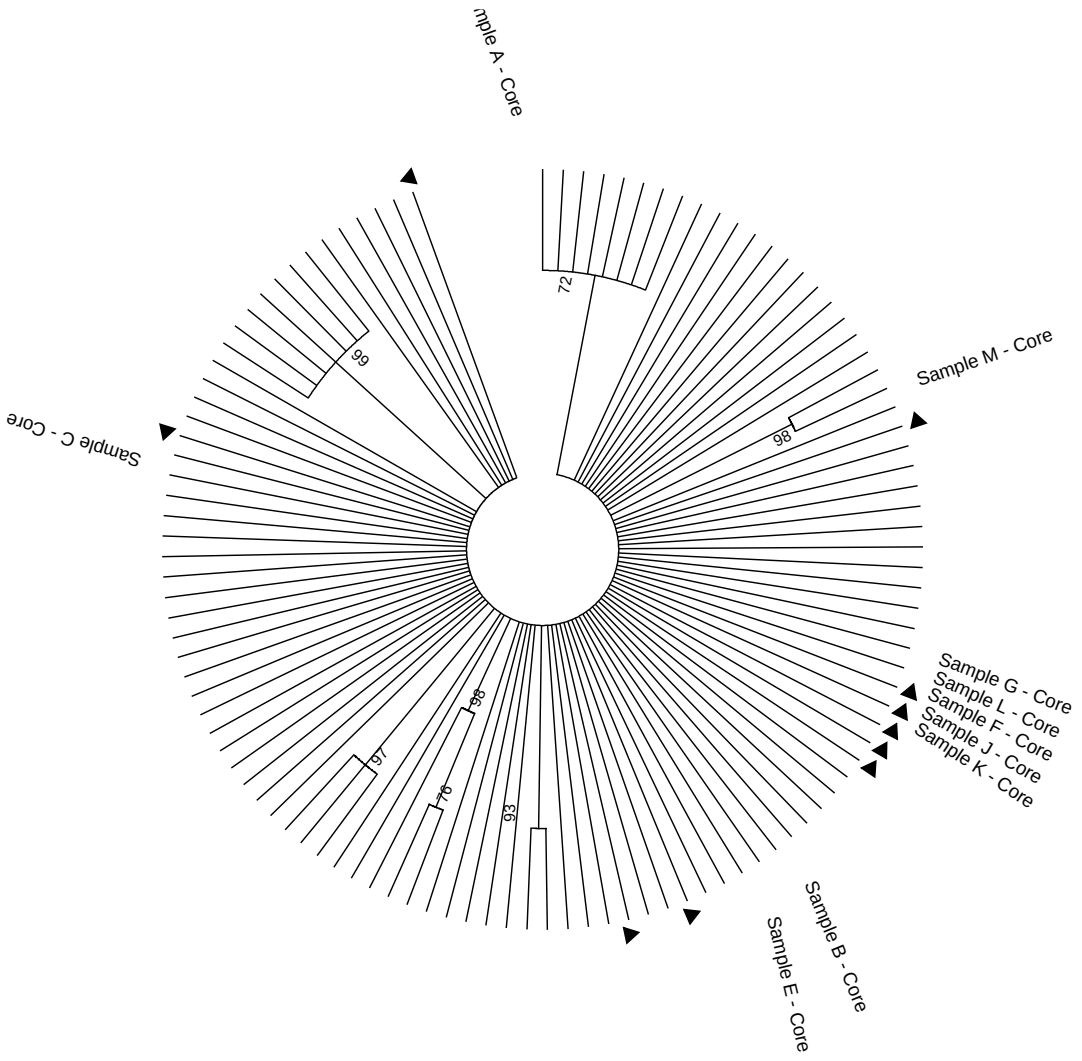


Figure 12 - Phylogenetic analysis of the core proteins samples when aligned with a number of HCV core protein sequences. ▲ indicates samples from this study. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The analysis involved 111 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 509 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. The sequences were aligned by ClustalW.

3.3.4 Discussion

See Appendix I – Sequences Used for the accession numbers of the sequences used. Bootstrap values are a method of assessing the accuracy

of a statistical estimate, and the method was introduced by Felsenstein around 1985 [178]. Regarding phylogeny, bootstrapping can be considered a re-sampling method - the data is re-sampled several times (for example, a bootstrap 1,000 on MEGA 7 would mean that the data is re-sampled 1,000 times), and a confidence value for each node of the phylogenetic tree is given, based on the number of times that that particular clade appeared during the re-sampling process. For example, a confidence value of 0.75 for a particular node from a bootstrap of 1,000 can be interpreted as that particular clade appearing 750 times during the re-sampling process, or 75% of the total re-samples. Bootstrapping has a number of advantages and disadvantages. The major advantage of bootstrapping is the simplicity of the process, and it also tends to be more accurate than the standard intervals obtained using sample variance and assumptions of normality. However, the bootstrap values do not indicate whether the clades would be present on the true phylogenetic tree - only the tree generated from data used. Short sequences can produce low bootstrap values. A clade with a bootstrap value of 0.70 (displayed as 70 in the above figures) and above is considered to have strong support for that clade appearing on the true tree - provided the data used to generate the tree was appropriate. The values produced for the above figures were low – this is likely due to the shortness of the sequences used and the genetic variability of the HCV genome.

The 320bp section of the HCV genome chosen for sequencing covers the junction between the E1 and E2 glycoproteins. The E2 glycoprotein contains three hypervariable regions and displays the highest genetic diversity of the HCV proteins. For the 320bp sequence phylogenetic analysis, the bootstrap values were below the acceptable threshold. This was probably caused by the short sequences used, and the fact that many of reference sequences and experimental sequences were not varied enough to produce a phylogenetic tree that properly represented the diversity of the samples. The 320bp region does not cover the HVR1 - a region diverse enough that it could have provided a clearer picture of the evolution of the experimental samples when compared to reference sequences from a variety of genotypes and subtypes. The samples from this study were compared to 20

genotype 3a 320bp samples obtained from patients in Ireland and added to GenBank by the MVDRL. The comparison samples appear to show grouping in Figure 9 but do not appear to be closely related to any of the samples from this study. The HCV genotype 3a reference sequence was trimmed to the length of the 320bp region and searched against the BLASTn database. The first 100 sequences matched to the 3a reference sequence were used to generate Figure 10. A condensed tree was generated, which hid bootstrap values for nodes where the bootstrap value was <70. The samples from the study appeared to group together, with two sequences interspersed between the samples. KY620474.1 is a genotype 3 (no subtype specified) sample submitted from the United Kingdom and collected in 2013. AM423059 is a genotype 3a sample that was submitted from France in 2007. Several HCV genotypes were represented in the BLASTn matches to the HVC 3a 320bp reference. The large clades with bootstrap values of 99 and 90 (DQ50XXXX – see Appendix I – Sequences Used) represent sequences obtained from perinatal infections in Italy, with genotypes ranging from 1a to 4d.

The mature form of the HCV core protein is a 21 kDa protein and has two domains – D1 and D2. These domains are separated by their amino acid compositions and hydrophobicity, and cleavage of the HCV core protein by host signal peptidases is needed for proper HCV assembly. The core protein is the most conserved of the HCV proteins, and sequence variations have been correlated with pathologies associated with HCV infection and disease progression – for instance, changes to D1 have been found to be associated with resistance to interferon therapy, and particular amino acid substitutions have been correlated with an increased risk of HCC [179] [180]. The core protein can be used to distinguish between genotypes 1a and 1b, but many diagnostic laboratories rely on the 5'-UTR for distinguishing between genotypes.

The bootstrap values for the core protein phylogenetic analysis gave poor support for the Maximum-Likelihood tree produced. Again, as stated during the discussion of the 320bp phylogenetic analysis, there are a number of factors that could have caused the low bootstrap values. The sequences used were short (~500 aa) and the sequences might not have been diverse

enough to generate a proper Maximum-Likelihood tree. However, the bootstrap values were an improvement on the 320bp phylogenetic analysis bootstrap values - probably because of the increased sequence lengths used. Figure 11 was generated by comparing the core samples to 20 BLASTn matches to the HCV 3a reference sequence – this was increased to 100 matches for Figure 12. Compared to the 320bp sequences, greater distancing between the core sequences from this study can be seen. In Figure 12 samples G, L, F, J, and K have clustered together, whereas samples M, C, and A are distanced from the cluster. Examining the uncondensed tree, sample A shows a weak relationship (bootstrap value = 25) with KY620456, which is genotype 3 sample from the United Kingdom, and was included in the same group of submissions as KY620474 – which was interspersed among the 320bp samples. Sample E showed a weak relationship (bootstrap value = 21) to KC143935, a genotype 3a sample collected in 2009 from Saudi Arabia.

Ultimately, the bootstrap values from the 320bp and core protein sequences were not high enough to support definitive phylogenetic relationships between the samples from the study and the samples from the GenBank database. Longer sequences that include conserved regions of the HCV genome would need to be obtained to improve the reliability of the phylogenetic analysis.

3.4 Preparatory Work for HCV Pseudoparticle Generation and Epitope Mapping

3.4.1 Introduction

This section covers the preparatory experiments carried out for the reverse epitope mapping procedure. As covered by Section 3.2, a genotype 3a E1E2 glycoprotein was unable to be isolated from the HCV samples received from the NVRL. Instead, a H77 E1E2 glycoprotein construct was used to create the pseudoparticles. H77 is a genotype 1a HCV isolate that has known infectious properties when used to create pseudoparticles, and has the complete genome sequenced (see

Appendix I – Sequences Used). This H77 E1E2 plasmid (phCMV-ΔC/E1/E2) was already stocked by the MVDRL, alongside the luciferase plasmid and packing vector (MLV gag-pol).

The first step for the preparatory experiments was determining which of the samples contained to the presence of antibody associated virus (AAV) complexes.

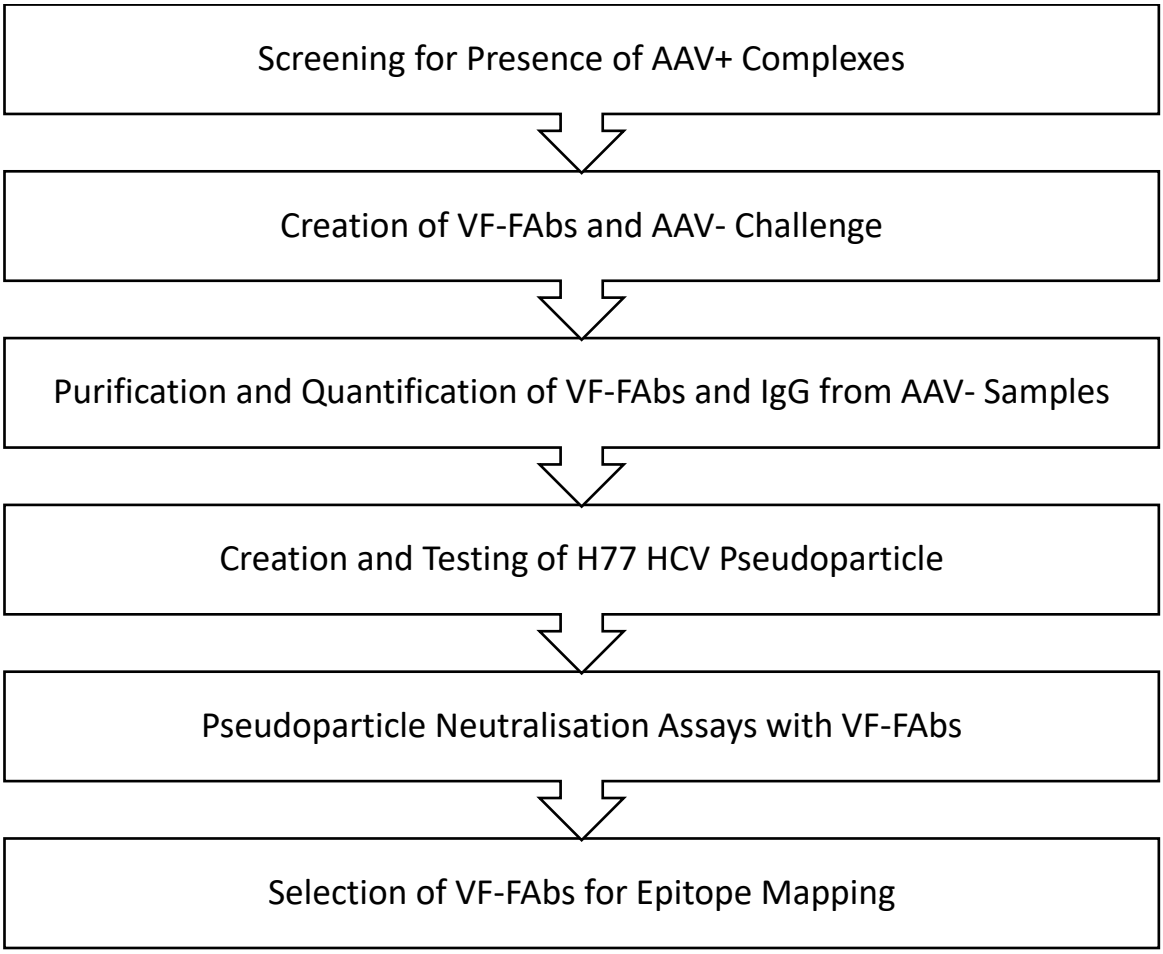


Figure 13 - Graphical summary of the preparatory process for selecting and sending samples for epitope mapping.

3.4.2 Screening of Serum/Plasma Samples for the Presence of Antibody Associated Virus (AAV) Complexes

3.4.2.1 Hypothesis

The first stage of the process was determining which of the samples received from the NVRL and from the Pyro2 project were AAV+. Because it was known that the antibodies from the AAV complexes could bind to HCV, the antibodies could be separated from the complex and later used to perform neutralisation assays using the HCV pseudoparticle system.

3.4.2.2 Methodology

Serum and plasma samples were obtained from the National Virus Reference Laboratory (NVRL) for 13 unrelated patients diagnosed with chronic HCV 3a infections. 6 HCV 3a samples were obtained from the Pyro2 project that the MVDRL had undertaken, and a single HCV 3a sample was

taken from the MVDRL stocks that had been approved for research. The serum and plasma samples were stored at -80°C. The following methodologies were used during the course of this experiment. Please refer to the named sections for further information on the protocols used.

- Protein G Column Chromatography (Section 2.3)
- vRNA Extraction (Section 2.4)
- 320bp RT-PCR (Section 2.6)
- Agarose Gel Electrophoresis (Section 2.2)

The samples underwent Protein G column chromatography to segregate the AAV+ samples from the AAV- samples – if the samples contained AAV, the complex would be captured by the Protein G membrane and eluted. However, if the same did not contain AAV, the virus particles would be washed through the column. The first and eighth washes (W0 and W8) from the Protein G columns also underwent the subsequent vRNA extraction and 320bp RT-PCR procedure to check that any unbound virus was properly washed through the Protein G column – W0 should have some unbound virus and produce a 320bp band, whereas W8 should be free of unbound virus and not produce a band. To verify whether any of the samples were AAV, a vRNA extraction and 320bp RT-PCR was used to check for the presence of HCV genetic material. If the AAV had been captured by the Protein G membrane, the subsequent vRNA extraction and PCR would amplify the genetic material from the virus portion of the complex. Negative, H_2O, and positive controls were used throughout the process to verify the protocols worked and to demonstrate that the preparations were not contaminated. The primary PCR and secondary PCR products were not processed at the same location, to reduce to chance of contamination between reactions. Completed samples were stored at -20°C.

3.4.2.3 Results

AAV-	AAV+
A	C

B	G
D	J
E	M
F	052
K	502
L	425
N	
O	
246	
859	
718	
690	

Table 14 - The AAV- and AAV+ samples, as determined by the screening process. The rate of AAV+ samples was 35%, which was characteristic of previous AAV screening procedures.

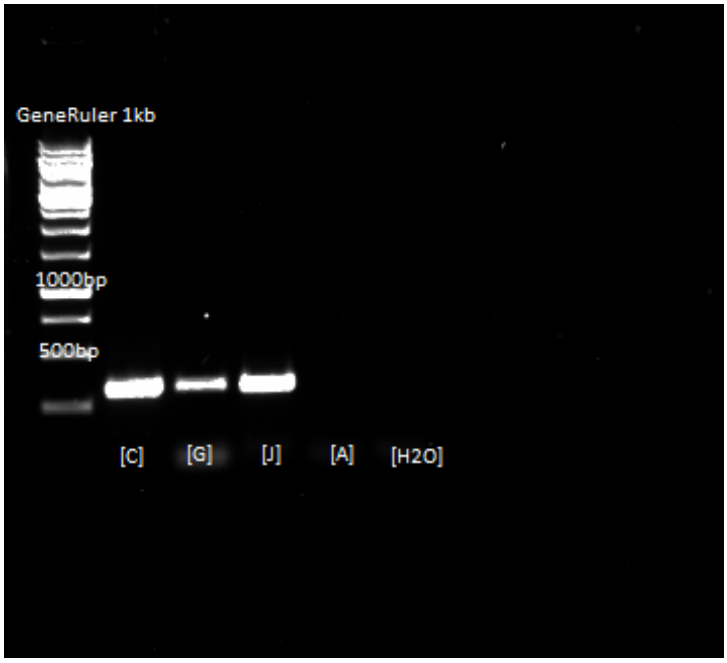


Figure 14 - Gel image showing positive bands for AAV+ samples at the 320bp molecular weight. Samples C, G, and J are AAV+, whereas Sample A did not show a 320bp band, and can be considered AAV-. H_2O was used as a control and to show the preparations were not contaminated by external DNA.

3.4.2.4 Discussion

7 of the 20 samples screened tested positive for the presence of AAV. This frequency of AAV+ samples was 35%, which matched previous AAV screening procedures carried out by the MVDRL. The antibodies isolated from the AAV- samples (using the Protein G columns) during the categorisation of the samples as AAV- or AAV+ was saved for the future pseudoparticle experiments - despite the fact there was no AAV complexes present for these genotype 3a samples, it was hypothesised that the antibodies could display some activity towards the genotype 1a H77 pseudoparticle.

Once the AAV+ samples had been identified, the antibodies needed to be separated from the AAV complex using proteinase K - these antibodies would be known as VF-FAbs (Virus-Free FAbs). The VF status of the antibodies was verified by using the 320bp PCR protocol (see Section 2.6).

3.4.3 Dissociation of AAV Complexes for the Generation of Virus-Free Antigen Binding Fragments (VF-FAbs) and AAV- Sample Challenge

3.4.3.1 Hypothesis

The reasoning behind the generation of VF-FAbs was based around the idea that an antibody from one AAV+ sample could be used to identify potential conserved epitopes among the AAV- samples. VF-FAbs originate from AAV+ samples which have undergone a proteinase K digestion to remove the virus portion of the AAV complex. The generated VF-Fab retains the antigen binding properties of the original antibody, but the Fc region of the antibody is degraded during the procedure.

3.4.3.2 Methodology

The following methodologies were used during the course of this experiment. Please refer to the named sections for further information on the protocols used.

- Protein G Column Chromatography (Section 2.3)

- vRNA Extraction (Section 2.4)
- 320bp RT-PCR (Section 2.6)
- Agarose Gel Electrophoresis (Section 2.2)

Proteinase K (5 mg/ml) was used to remove the virus from the AAV complex, generating the VF-FAbs for further processing. Proteinase K is a serine protease commonly used to digest proteins and remove contamination from nucleic acid preparations. The primary site of cleavage is the peptide bonds adjacent to the carboxyl group of aliphatic and aromatic amino acids with blocked alpha amino groups. The decision to use proteinase K to separate the antibodies from the AAV complex was based on previous work performed by Dr. Amruta Naik – Dr. Naik found that a 1-hour incubation at 37°C with a 50/50 mixture of the sample and proteinase K was enough to separate the antibodies.

A 1:1 volume of proteinase K and the AAV+ sample was added to a 1.5 ml microcentrifuge tube and incubated at 37°C for two hours. Once the incubation was completed, the mixture was added to a Protein G HP SpinTrap™ Column and washed according the instructions. A second wash through the Protein G HP SpinTrap™ was performed to remove any residual Proteinase K. A 320bp RT-PCR and agarose gel electrophoresis was performed to verify that the sample did not contain any virus. The VF-FAb was stored at -20°C. The isolated and purified VF-FAbs were used to challenge the serum samples that had not tested positive for the presence of antibody-associated virus complexes (AAV-). A 1:5 volume of the VF-FAb and the AAV- serum were added to a 1.5 ml microcentrifuge tube and incubated at 37°C for two hours. Once the incubation period was completed, the mixture was added to a Protein G HP SpinTrap™ Column and washed according the instructions. A 320bp RT-PCR and agarose gel electrophoresis was performed to check whether the VF-Fab had bound the AAV- sample.

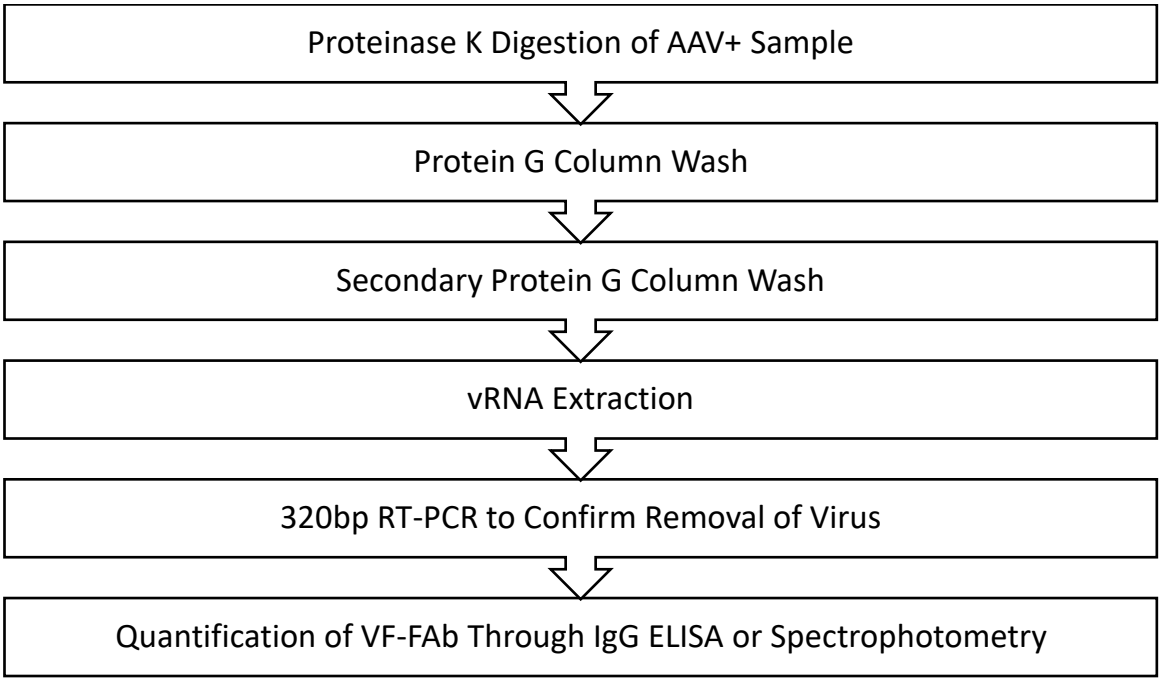


Figure 15 - Procedure for generation of VF-FAbs. A secondary protein G column wash was included to ensure complete removal of any residual proteinase K.

3.4.3.3 Results

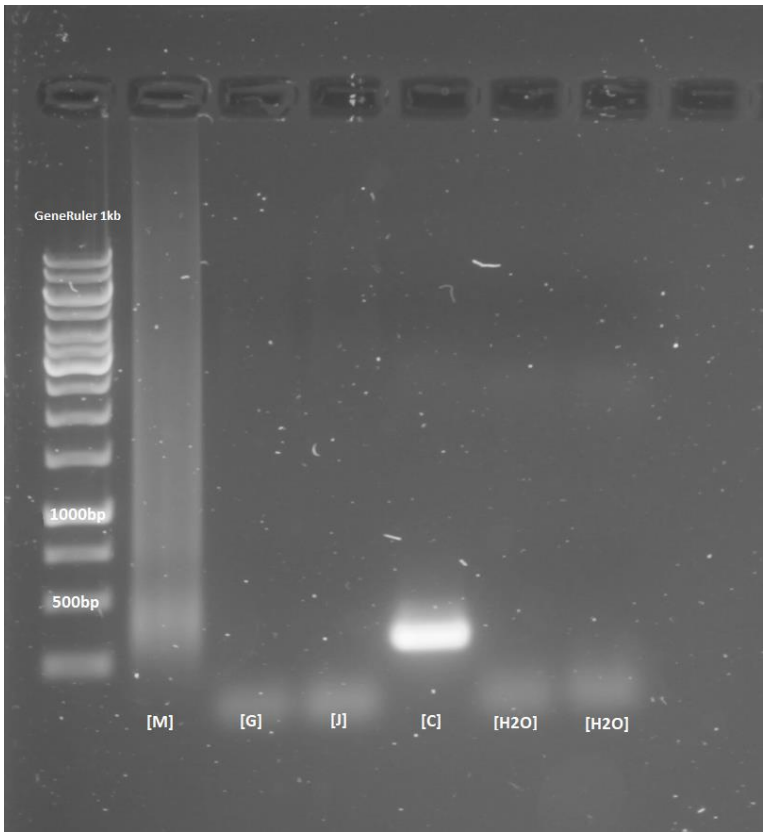


Figure 16 - Agarose gel electrophoresis on the VF-FAbs after proteinase K digestion and column washes. Sample M appeared to have some residual DNA remaining, and the HCV was not dissociated from the Sample C AAV complex. However, Sample G and J were

suitable for use as purified VF-FAbs and were confirmed to be free of HCV.

3.4.3.4 Discussion

The aim of this experiment was to produce VF-FAbs from the four AAV+ samples discovered during the AAV screening process. There were four AAV+ samples, and from that, three purified VF-FAbs were produced. Figure 16 shows that Sample C, despite the proteinase K digestion, showed evidence of HCV genetic material, and was not used for downstream experiments, as the residual DNA could produce false positives during the AAV- experiments. Sample G and J had the virus completely digested and were suitable for use against the AAV- samples. Sample M had some remaining genetic material, but subsequent washes and a longer proteinase K digestion were able to remove the smearing seen on Figure 16. The VF-FAbs and the AAV- samples were pooled for the experiments to preserve the stocks of HCV samples. Unfortunately, none of the VF-FAbs were able to capture the AAV- samples. This could mean that the antibodies from the G, J, and M AAV+ complexes bound to epitopes that were not found on the AAV- samples. The antibodies were investigated further using pseudoparticle experiments and epitope mapping.

It was also examined whether the VF-FAbs produced could “recapture” the original AAV+ samples that they were isolated from, but the VF-FAbs were unable to recapture the AAV+ samples. This could be because the proteinase K digestion damaged or changed the structure of the antibodies, or because the original samples already had antibodies blocking the VF-FAbs binding sites.

3.4.4 Isolation and Purification of IgG from Patient Samples for HCVpp Neutralisation Assays

3.4.4.1 Hypothesis

To use the IgG from the HCV 3a patients for the HCVpp assays, the IgG first needed to be isolated and purified from serum/plasma.

3.4.4.2 Methodology

The following methodologies were used during the course of this experiment. Please refer to the named sections for further information on the protocols used.

- Protein G Column Chromatography (Section 2.3)
- Proteinase K Digestion (Section 3.4.3.2)
- IgG ELISA (Section 2.11)
- SDS-PAGE (Section 2.12)

For consistency and accuracy when carrying out the antibody experiments, the IgG content of the samples was quantified using an IgG ELISA. During the SDS-PAGE procedure, the IgG isolated using Protein G column chromatography underwent heating at 95°C for 5 minutes. This was done to separate the disulphide bonds between the heavy (H) and light (L) IgG chains. The H chain is estimated to have a molecular weight of 50 kDa and the L chain around 25 kDa, meaning that two bands would separate out during the SDS-PAGE procedure. Observing the H and L chains would confirm that the IgG remained undamaged throughout the column chromatography procedures.

3.4.4.3 Results

The IgG ELISA was used to quantify the amount of protein isolated from the column chromatography.

Sample	Concentration (mg/ml)
A	2.80
B	1.48
C	2.07

D	1.44
E	3.72
F	3.61
G	2.33
H	2.00
I	2.94
L	2.92
M	0.82
N	2.46
O	2.84
VF-G	0.01
VF-J	0.26
VF-M	0
Average Yield	1.98

Table 15 - Quantities of IgG observed. As previously noted, AAV+ samples underwent proteinase K digestions to remove the virus from the AAV complex. However, the proteinase K also had an impact on the amount of IgG detected during the ELISA procedure. The R^2 value for the standard curve used was 0.9739.

The samples underwent the aforementioned heat treatment and were processed using the SDS-PAGE technique and Coomassie Blue staining.

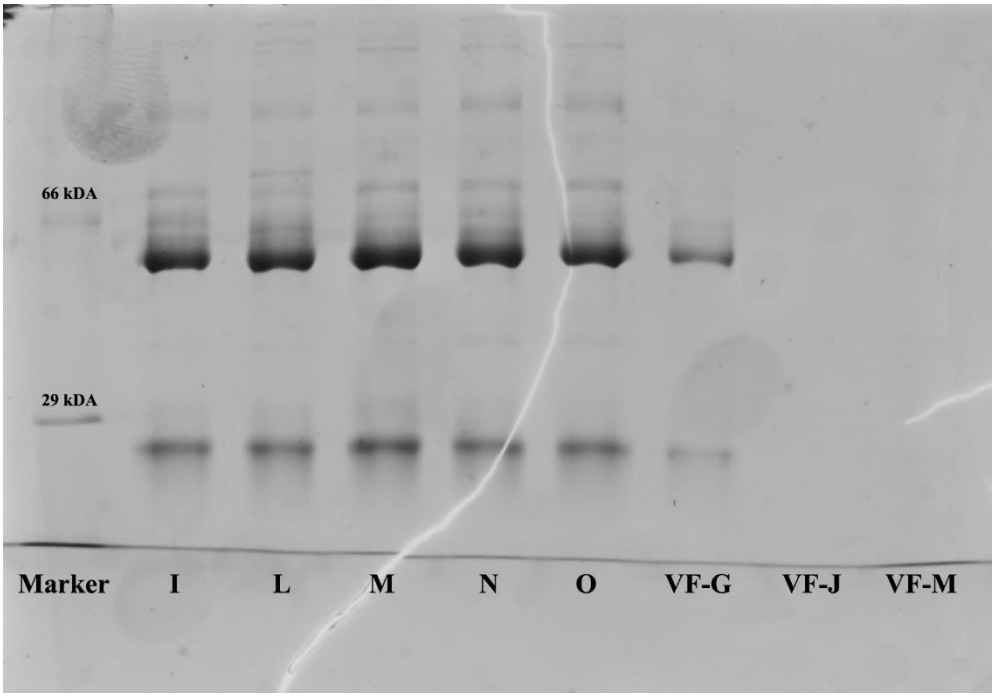


Figure 17 - A gel generated using the isolated IgG and the SDS-PAGE technique. Samples I, L, M, N, O, and VF-G showed clear separation of the IgG chains. Nothing was visible for sample VF-J or VF-M. Degradation of the tenth lane prevented the use of a negative H_2O control.

3.4.5 Generation of HCV Pseudoparticles and IgG Neutralisation Assays

3.4.5.1 Hypothesis

As discussed during Section 2.18, HCV pseudoparticles are generated using a variation of $CaPO_4$ transfection, combining a plasmid containing HCV E1E2 glycoprotein, a packaging vector, and a reporter plasmid. HCVpp were generated using the H77 E1E2 glycoprotein – a genotype 1a isolate that is known to produce infectious HCVpp. Once infectious HCVpp were produced, the pseudoparticles would be used to determine which of the isolated antibodies would be sent to Pepscan Presto B.V for epitope mapping. A number of scoping assays were performed to determine the best candidates for epitope mapping. Firstly, the IgG concentration of the antibodies was standardised. The antibodies were used to neutralise infectious H77 pseudoparticles. The antibodies that demonstrated the most neutralisation underwent further experiments using varying IgG concentrations. For the last experiment, the chosen antibodies were combined to examine whether combined antibodies would have increased neutralisation against H77 pseudoparticles when compared to single antibodies.

3.4.5.2 Methodology

The following methodologies were used during the course of this experiment. Please refer to the named sections for further information on the protocols used.

- Generation of HCV Pseudoparticles (Section 2.18)
- Infectivity Assay using HCVpp (Section 2.18.1)
- Neutralisation Assay using HCVpp (Section 2.18.2)
- Luciferase Assay using HCVpp (Section 2.18.3)

It was shown from the infectivity assays using Huh7s that the H77 pseudoparticles were functional and would be suitable for working with the VF-FAbs – this was validated by comparing the H77 pseudoparticles to pseudoparticles that did not contain the H77 plasmid (“empty”

pseudoparticles) and a H_2O control, showing that the H77 pseudoparticles were infectious compared to the empty pseudoparticles.

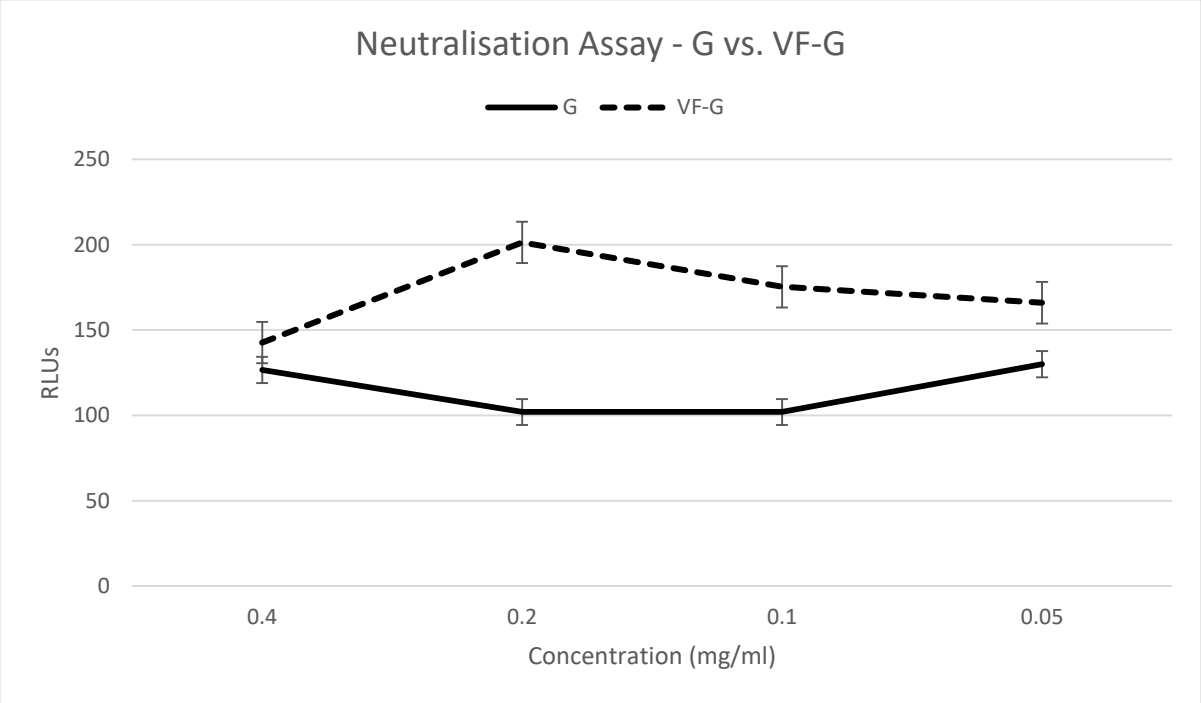
3.4.5.3 Results

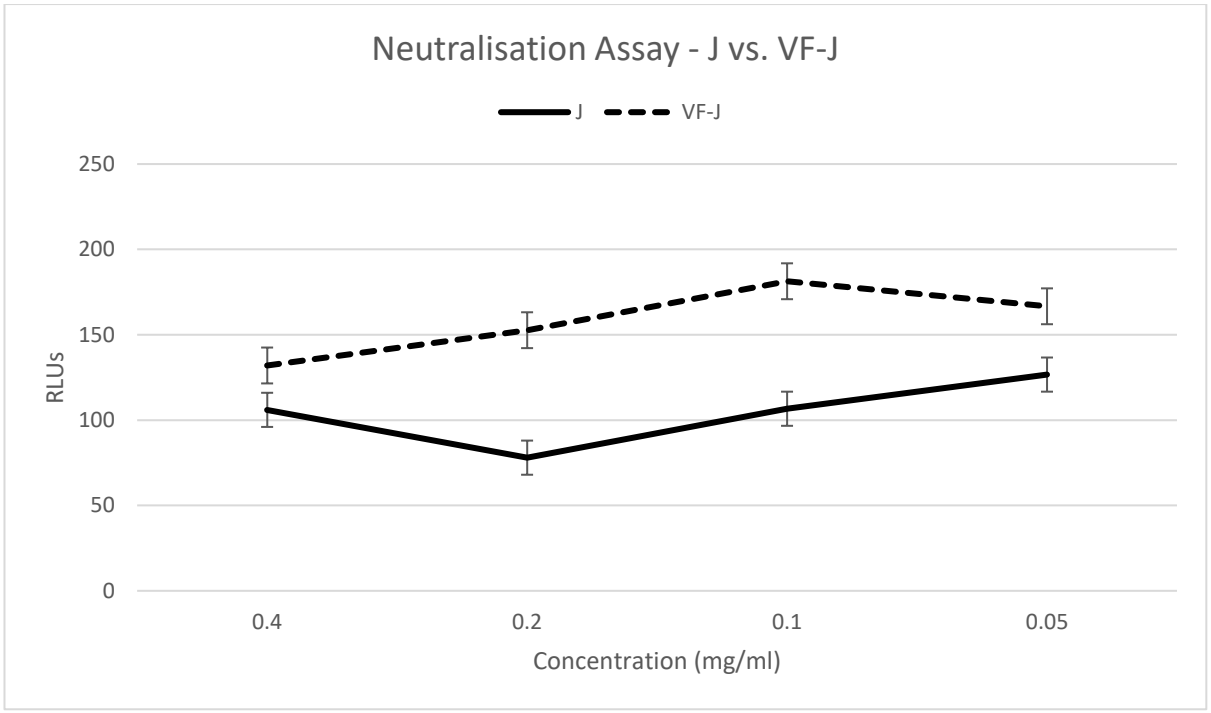
The purified antibodies were brought to the same concentration and tested against the H77 pseudoparticles to get a rough estimate of which antibodies would be the most suitable for further experimentation.

Sample	Relative Light Units (RLUs)	Reduction of Infectivity (%)
H77	8777	-
A	71	99.19
B	97	98.89
C	269	96.94
D	43	99.51
E	179	97.96
F	1147	86.93
G	29	99.67
H	197	97.76
I	469	94.66
L	265	96.98
M	617	92.97
N	649	92.61
O	499	94.31
VF-G	291	96.68
VF-J	661	92.47
VF-M	491	94.41
052	261	97.03
859	371	95.77
246	709	91.92
690	275	96.87
718	131	98.51
502	539	93.86
425	461	94.75
Average	**379**	**95.68**

Table 16 - Initial neutralisation assay on the IgG antibodies purified from the HCV 3a patient samples. Every antibody sample was able to neutralise the H77 pseudoparticle. The average reduction of infectivity was 95.68%.

From Table 16, samples G, J, M, VF-G, VF-J, and VF-M were selected for further examination. Samples G, J, and M had proven to be AAV+ and the proteinase K digests had created copies of the samples that contained the antibodies but not the virus – VF-G, VF-J, and VF-M. The second pseudoparticle experiment used four concentrations of IgG to determine whether the concentration of IgG correlated to the neutralisation of the pseudoparticles.





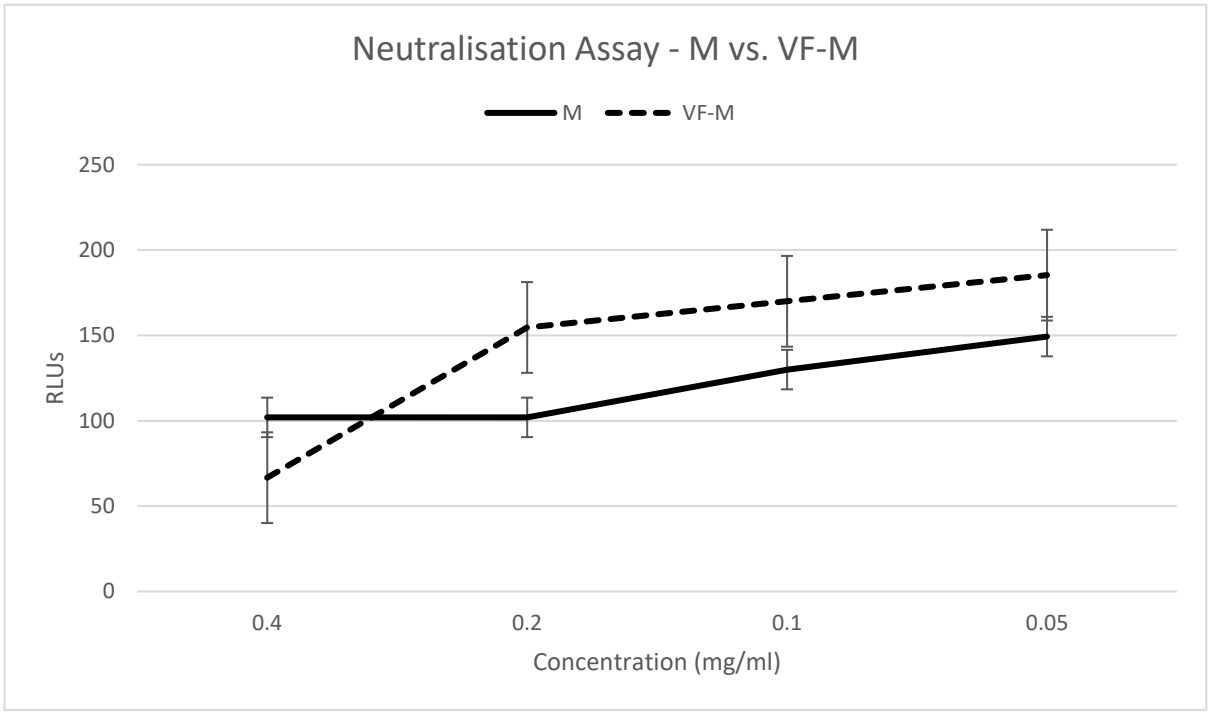


Figure 18 - Comparison of neutralisation between "whole" samples of antibodies and samples that have undergone proteinase K digestion (VF). On average, the "whole" samples demonstrated superior neutralisation – but the difference was not significant when compared to the control H77 sample. At the concentrations tested, the concentration of IgG did not appear to have any relationship to the antibodies' neutralisation. ***The control H77 sample was measured at 417 RLUs*** *– this was excluded from the graphs to make the figures more readable.*

The concentration of IgG did not appear to have any significant impact on

neutralisation. The “whole” samples appeared to have improved neutralisation compared to their VF counterparts, but this improvement was not significant when compared to the control H77 sample.

A third experiment was performed to compare the neutralisation of combined antibodies to single antibodies. The hypothesis was that increased neutralisation from combined samples would imply that the individual samples (G, J, or M) targeted separate epitopes on the H77 E1E2. If neutralisation was not increased, it would suggest that the samples share the same epitope on the H77 E1E2, and that combining the samples only caused the competition for a single epitope. The “whole” samples G, J, and M were used for the third experiment.

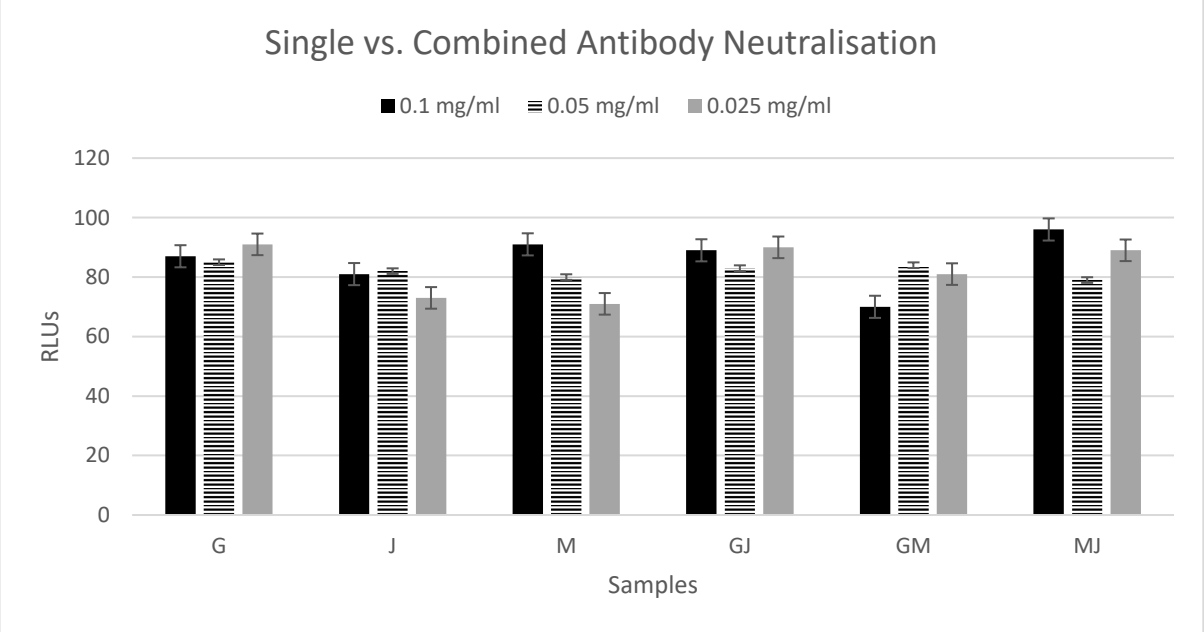


Figure 19 - Comparison between single and combined antibody samples. The combined samples did not demonstrate increased neutralisation when compared to the single samples. This implies that the antibodies isolated from the G, J, and M samples share the same epitope.

It was decided that samples G, J, and M would be sent to Pepscan Presto B. V. for conformational epitope mapping. AP33 was included to act as a standard antibody.

3.5 EPITOPE MAPPING

3.5.1 Hypothesis

Epitope mapping refers to the process of identifying the epitopes of antibodies to the corresponding antigens. For this project, “reverse” epitope mapping was used, meaning that antibodies that were known to bind to the HCV E2 glycoprotein were used to examine E2 epitopes.

3.5.2 Disclaimer

The discussion for Section 3.5 references research performed by Naik *et al*. and published as “*Reverse epitope mapping of the E2 glycoprotein in antibody associated hepatitis C virus*.” (PLoS One. 2017 May 30;12(5):e0175349. doi: 10.1371/journal.pone.0175349). As stated during Section 3.1, some site-directed mutagenesis was performed by Nicole Walsh for the purpose of this paper. However, the epitope mapping was performed completely by Naik *et al*. and Nicole Walsh does not claim to have performed any of the epitope mapping related to the paper.

3.5.3 Methodology

The following methodologies were used during this experiment. Please refer to the appropriate sections for further information on the protocols used.

- Epitope Mapping (Section 2.19)

3.5.4 Results

Several peptides were synthesised for the epitope mapping procedure. The order of the peptides on the cards was randomised. The epitope mapping report can be viewed in Appendix IV. The successful sets can be seen below (Table 17).

SET	Mimic Type	Label	Description	Sequences
SET 1	Linear	LIN	Linear peptides (length = 15aa) derived from the target sequence, with an offset of one residue. Cysteine residues were protected by acetamidomethyl (denoted “2”).	2GPVY2FTPSPVVVG TALN2NDSLNTGFLA NRTALN2NDSLNTGF SWGENETDVLLLNNT DTLT2PTD2FRKHPE STGFTKT2GGPP2NI GVPTYSWGENETDVL KHPEATYTR2GSGPW KFAQGWGPITHTEPP WFG2TWMNSTGFTKT
SET 2	Loop – mP2 CLIPS	LOOP	Constrained peptides (length = 17aa). On positions 2 – 16 are 15-mer sequences derived from the target sequence. To introduce structural constraints, cysteine residues were inserted on positions 1 and 17, and were then constrained by mP2 CLIPS. Native cysteine residues were protected by acetamidomethyl (denoted “2”).	CETHTVGGSASRAAHRC CTHTVGGSASRAAHRVC CHTVGGSASRAAHRVTC CTVGGSASRAAHRVTTC CVGGSASRAAHRVTTFC CGGSASRAAHRVTTFIC CGSASRAAHRVTTFITC CSASRAAHRVTTFITRC CASRAAHRVTTFITRGC CSRAAHRVTTFITRGPC

SET	Mimic Type	Label	Description	Sequences
SET 3	Helical – mP2 CLIPS	HEL	Structured peptides (length = 23aa) derived from the target sequence, with an offset of one residue. On positions 1 and 5 there are cysteine residues which are joined by mP2 CLIPS. Native cysteine residues were protected by acetamidomethyl (denoted “2”).	CETHCVGGSASRAAHRVTTFIT CTHTCGGSASRAAHRVTTFITR CHTVCGSASRAAHRVTTFITRG CTVGCSASRAAHRVTTFITRGP CVGGCASRAAHRVTTFITRGPS CGGSCSRAAHRVTTFITRGPSQ CGSACRAAHRVTTFITRGPSQN CSASCAAHRVTTFITRGPSQNI CASRCAHRVTTFITRGPSQNIQ CSRACHRVTTFITRGPSQNIQL
SET 4	β-turn –mP2 CLIPS	BET	Structured peptides (length = 22aa). On positions 2 – 21 there are 20-mer sequences derived from the target sequence, with an offset of one residue. “PG” residues supplant the residues present on positions 10 and 11 to induce the β-turn formation. On positions 1 and 22 there are cysteine residues which are joined by mP2 CLIPS. Native cysteine residues were protected by acetamidomethyl (denoted “2”).	CHRVTTFITRPGSQNIQLINTC CTYTR2GSGPPGTPR2MVHYPC CQNIQLINTNPGWHINRTALNC CNDTLT2PTDPGRKHPEATYTC CNETDVLLLNPGRPPRGNWFGC CETDVLLLNNPGPPRGNWFG2C CP2GIVPAAQPGGPVY2FTPSC C2RPIDKFAQPGGPITHTEPPC CDSLNTGFLAPGFYTHRFNASC CHYAPRP2GIPGAAQV2GPVYC

Table 17 - The successful peptides synthesised for epitope mapping.

The antibodies were tested against the sequence:

#H77_E2

THVTGGSAGRTTAGLVGLLTPGAKQNIQLINTNGSWHINSTALNCNESLNT
GWLAGLFYQHKFNSSGCPERLASCRRLTDFAQGWGPISYANGSGLDERP
YCWHYPPRPCGIVPAKSVCGPVYCFTPSPVVVGTTDRSGAPTYSWGAND
TDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGGVGNNTLLCPT
DCFRKHPEATYSRCGSGPWITPRCMVDYPYRLWHYPCTINYTIFKVRMYV
GGVEHRLEAACNWTRGERCDLEDRDRSELSPLLLSTTQWQVLPCSFTTLP
ALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEYVVLLFLLLADARVCSCL
WMMLLISQAEA

This amino acid sequence is based on the E2 glycoprotein of the H77 HCV genotype 1a isolate. The above sequence covers the hypervariable region of the E2 glycoprotein.

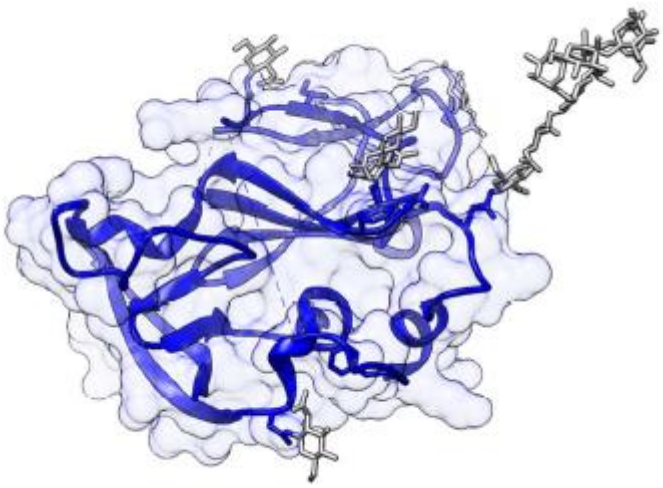


Figure 20 - A rendering of the HCV E2 glycoprotein structure, produced by Pepscan for the epitope mapping report. The blue ribbon represents the protein structure, and the grey structures are sugar moieties. The surface representation is transparent.

Sample AP33

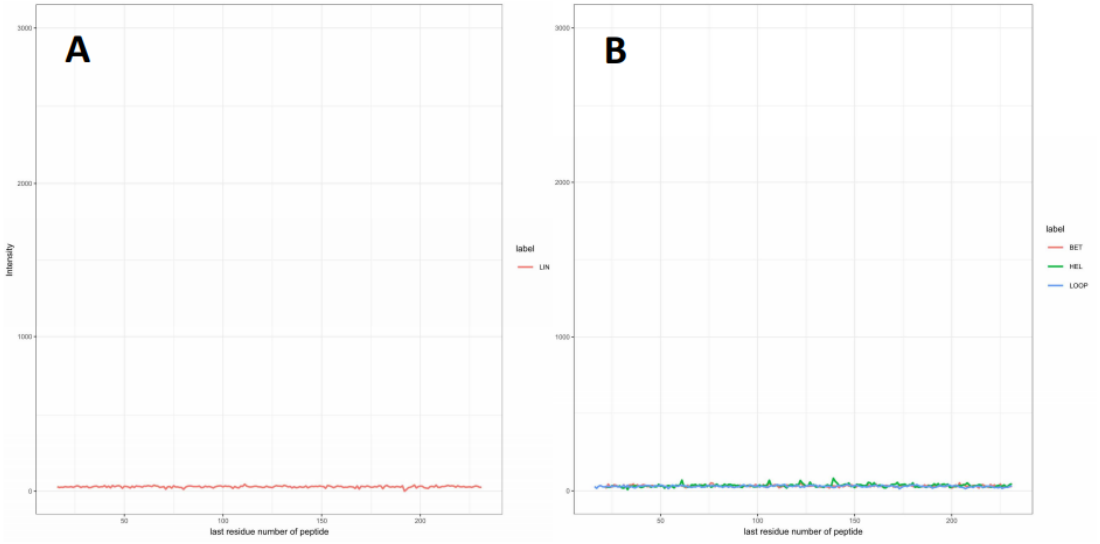


Figure 21 - The intensity profiles of the linear and conformational epitope mimics for Sample AP33. (A) represents the linear peptide mimics and (B) represents the conformational peptide mimics. The signal intensity is on the Y axis and positions of the last residues of a peptide with respect to the target sequence is on the X axis. LIN = linear; BET = beta-turn; HEL = helical; LOOP = looped mimic.

No binding peaks were observed for Sample AP33.

Sample J

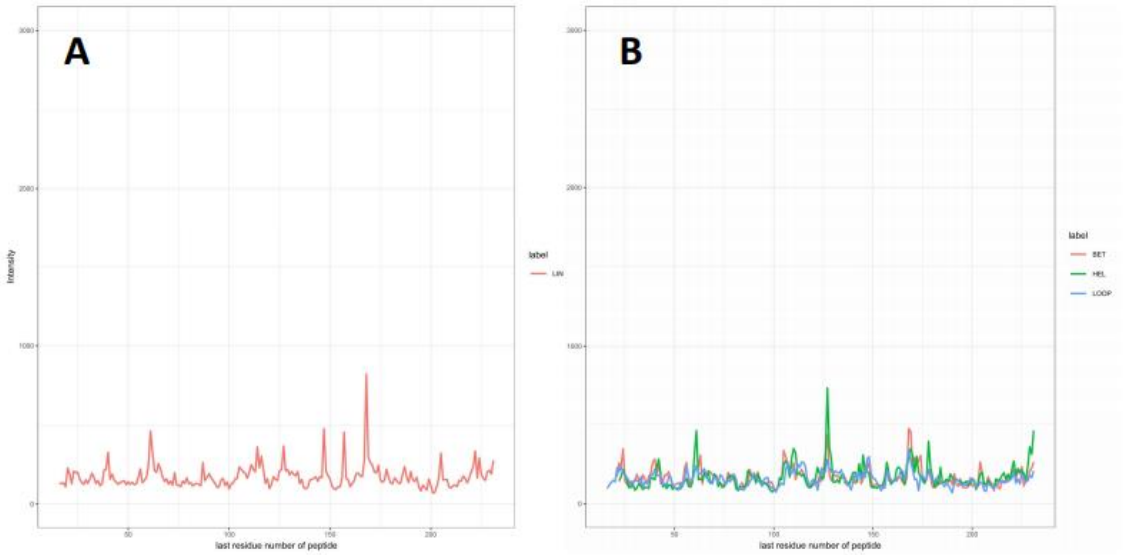


Figure 22 - The intensity profiles of the linear and conformational epitope mimics for Sample J. (A) represents the linear peptide mimics and (B) represents the conformational peptide mimics. The signal intensity is on the Y axis and positions of the last residues of a peptide with respect to the target sequence is on the X axis. LIN = linear; BET = beta-turn; HEL = helical; LOOP = looped mimic.

For Sample J, a number of binding peaks for linear and conformational mapping were observed, suggesting the presence of putative epitopes. Some binding peaks were observed during the linear mapping (Figure 22A), but these peaks were not consistent between repeated screenings. Several peaks were observed in the conformational mapping (Figure 21B); regions revolving around sequences $_{107}$YAPRPCGIVPAAQVCGPVYCF$_{127}$, $_{211}$TYTRCGSGPWLTPRCMVHYPY$_{231}$, and $_{149}$ENETDVLLLPGTRPPRGNWF$_{168}$ demonstrated binding to the helical mimics and/or beta-turn mimics. During the linear peptide screening, minor peaks were also observed at position $_{149}$ENETDVLLLPGTRPPRGNWF$_{168}$. Additionally, the conformational screening flagged sequence $_{5}$TVGGSASRAPGRVTTFITRG$_{24}$ as perhaps being part of a putative epitope.

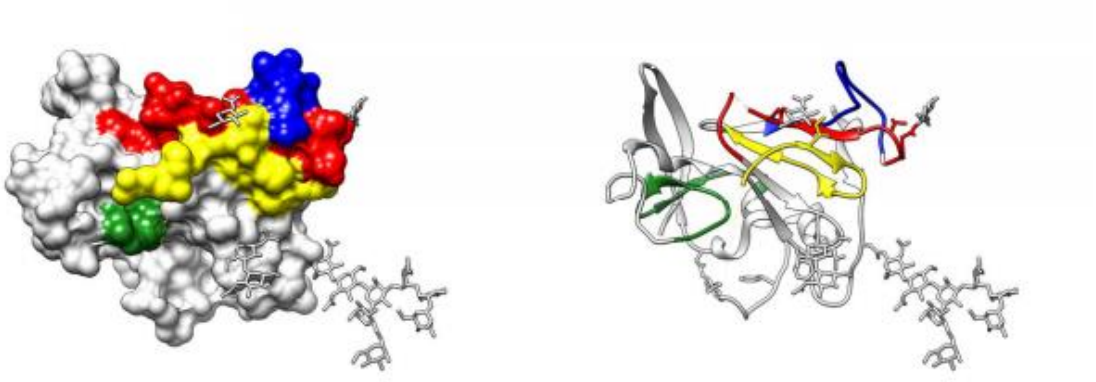


Figure 23 - Description of putative epitopes for Sample J on surface, and a ribbon representation.

Sample M

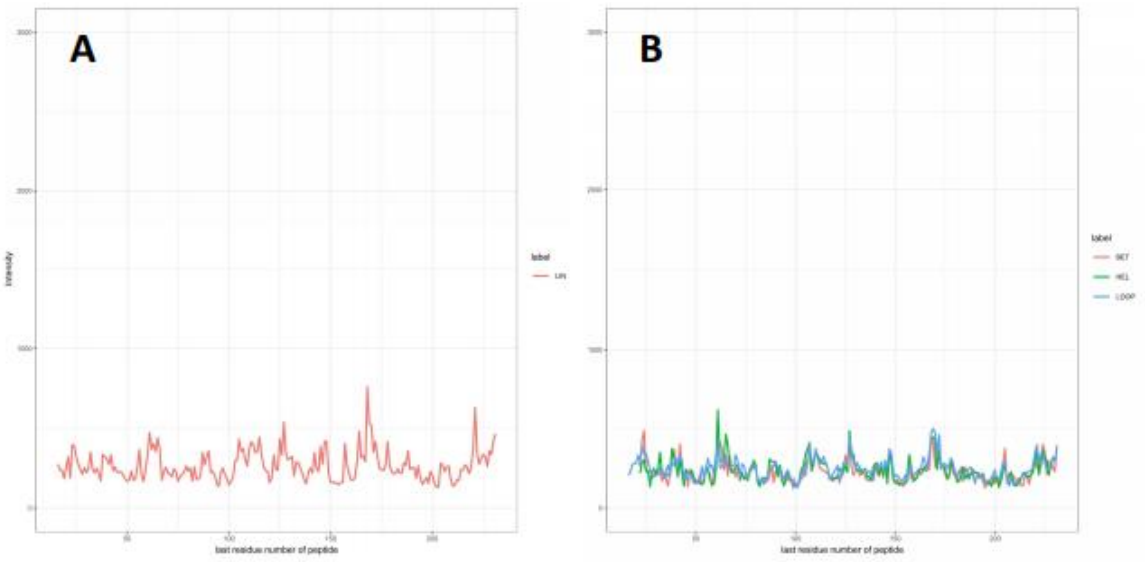


Figure 24 - The intensity profiles of the linear and conformational epitope mimics for Sample M. (A) represents the linear peptide mimics and (B) represents the conformational peptide mimics. The signal intensity is on the Y axis and positions of the last residues of a peptide with respect to the target sequence is on the X axis. LIN = linear; BET = beta-turn; HEL = helical; LOOP = looped mimic.

For Sample M, several binding peaks were observed during the linear and conformational mapping, suggesting the presence of putative epitopes. It should be noted that residues outside of the major peaks could contain residues involved with a putative epitope. Figure 24A shows the binding events observed during the linear mapping, but because of the low signal-to-noise ratio these peaks were not considered as putative epitopes. The best binding was observed on $_{154}$VLLLNNTRPPRGNWFGC$_{170}$. This region was also present during the conformational array (Figure 24B). The main peak for the conformational arrays was centred on the helical mimic at the sequence $_{40}$NRTALNCNDSLNTGFLAALFY$_{60}$. The β-turn mimic $_{5}$TVGGSASRPGHRVTTFITRG$_{24}$ was a consistent presence during multiple screens but did not meet the criteria for an epitope – however, it could form part an epitope.

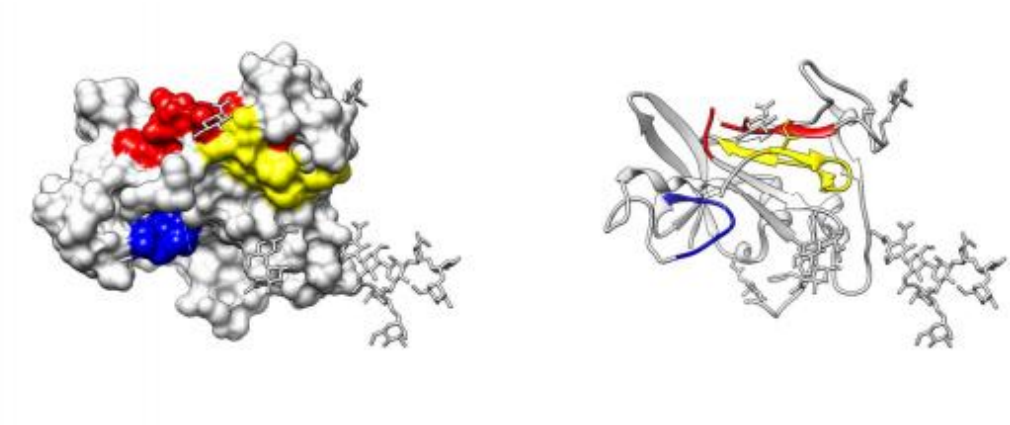


Figure 25 - Description of putative epitopes for Sample M on surface, and a ribbon representation.

Sample G

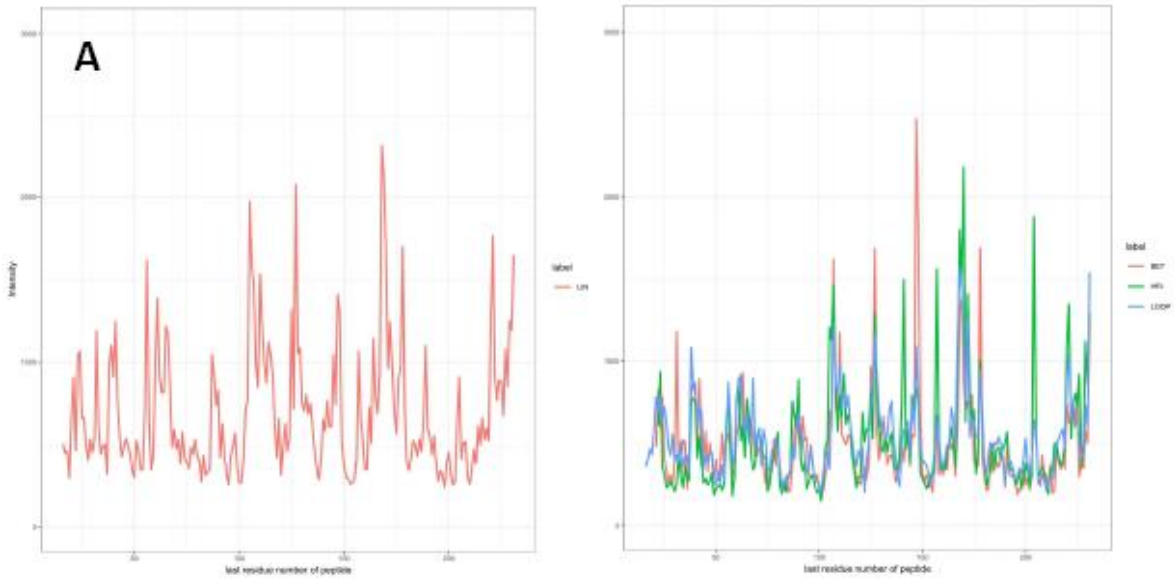


Figure 26 - The intensity profiles of the linear and conformational epitope mimics for Sample G. (A) represents the linear peptide mimics and (B) represents the conformational peptide mimics. The signal intensity is on the Y axis and positions of the last residues of a peptide with respect to the target sequence is on the X axis. LIN = linear; BET = beta-turn; HEL = helical; LOOP = looped mimic.

For Sample G, multiple peaks were observed for both the linear and conformational mapping, suggesting the presence of several putative epitopes. However, because of the presence of numerous peaks, the proper identification of epitopes was complicated. Pepscan identified the main peaks based on the intensity of the signal. (Table 17).

Sample	Area	Main Epitope Candidates

AP33	-	-
J	LIN	VLLLNNTRPPRGNWF VVGTTDRFGVPTYSW CNDSLNTGFLAALFY PTYSWGENETDVLLL
	BET	CENETDVLLLPGTRPPRGNWFC CNETDVLLLNPGRPPRGNWFGC
	HEL	CYAPCPCGIVPAAQVCGPVYCF CTYTCCGSGPWLTPRCMVHYPY CNRTCLNCNDSLNTGFLAALFY
M	LIN	VLLLNNTRPPRGNWF LLLNNTRPPRGNWFG LLNNTRPPRGNWFGC HPEATYTRCGSGPWL GIVPAAQVCGPVYCF
G	LIN	VLLLNNTRPPRGNWF LLLNNTRPPRGNWFG LLNNTRPPRGNWFGC RGNWFGCTWMNSTGF GIVPAAQVCGPVYCF PCGIVPAAQVCGPVY THTEPPSSDQKPYCW HTEPPSSDQKPYCWH TEPPSSDQKPYCWHY PSSDQKPYCWHYAPR HPEATYTRCGSGPWL SGPWLTPRCMVHYPY CGSGPWLTPRCMVHY RTALNCNDSLNTGFL CNDSLNTGFLAALFY VVGTTDRFGVPTYSW VGTTDRFGVPTYSWG QNIQLINTNGSWHIN
	BET	CTPSPVVVGTPGRFGVPTYSWC CPSPVVVGTTPGFGVPTYSWGC CNTRPPRGNWPGCTWMNSTGFC

CAPRPCGIVPPGQVCGPVYCFC

CGPITHTEPPPGDQKPYCWHYC

CNETDVLLLNPGRPPRGNWFGC
CENETDVLLLPGTRPPRGNWFC

CYTRCGSGPWPGPRCMVHYPYC

HEL CNETCVLLLNNTRPPRGNWFGC
CGENCTDVLLLNNTRPPRGNWF
CENECDVLLLNNTRPPRGNWFG
CTDVCLLNNTRPPRGNWFGCTW
CWGECETDVLLLNNTRPPRGNW

CGPPCNIGGVGNDTLTCPTDCF

CTDRCGVPTYSWGENETDVLLL

CCGPCYCFTPSPVVVGTTDRFG

CWGPCTHTEPPSSDQKPYCWHY
CQGWCPITHTEPPSSDQKPYCW
CGWGCITHTEPPSSDQKPYCWH
CHTECPSSDQKPYCWHYAPRPC

CTDCCRKHPEATYTRCGSGPWL
CEATCTRCGSGPWLTPRCMVHY
CPTDCFRKHPEATYTRCGSGPW

CYAPCPCGIVPAAQVCGPVYCF

CTYTCCGSGPWLTPRCMVHYPY

LOOP CVLLLNNTRPPRGNWFC
CLLLNNTRPPRGNWFGC
CNNTRPPRGNWFGCTWC
CLLNNTRPPRGNWFGCC
CDVLLLNNTRPPRGNWC

CSGPWLTPRCMVHYPYC

CTEPPSSDQKPYCWHYC
CPSSDQKPYCWHYAPRC
CEPPSSDQKPYCWHYAC

CGIVPAAQVCGPVYCFC
CIVPAAQVCGPVYCFTC
CCGIVPAAQVCGPVYCC

CVVGTTDRFGVPTYSWC

CGPSQNIQLINTNGSWC
CSQNIQLINTNGSWHIC
CHPEATYTRCGSGPWLC
CKHPEATYTRCGSGPWC

Table 18 - The epitope candidates identified by the screening process. The peptide groups are ordered by the intensity of the signal, and only binding events of >3 times the median value were included.

The four categories of putative epitopes (LIN, BET, HEL, and LOOP) were investigated further. The putative epitopes were examined, and "consensus" epitopes made by paring the putative epitopes down to their shared residues. These consensus epitopes were examined against the H77 sequence to see where on the E2 glycoprotein they targeted, and whether they overlapped against known antigenic sites.

Area	Positions
HVR1	384 - 410
AS412	412 - 423
AS434	434 - 446
HVR2	460 - 482
E2 β-Sandwich	492 - 566
IgVR	572 - 588

Table 19 – A guide to key regions of the E2 glycoprotein.

Name	Sequence	Notes
LIN_1	$_{536}$VLLLNNTRPPRGNWFGC$_{552}$	
LIN_2	$_{495}$GIVPAAQVCGPVY$_{507}$	E2 β-Sandwich
LIN_3	$_{481}$DQKPYCWHY$_{489}$	
LIN_4	$_{589}$HPEATYTRCGSGPWL$_{608}$	
LIN_5	$_{429}$CNDSLNTGFL$_{438}$	AS434
LIN_6	$_{516}$VGTTDRFGVPTYSW$_{529}$	

Table 20 – Linear epitopes identified by reverse epitope mapping.

Name	Sequence	Notes
AN1	$_{393}$SRAAHRVTTFITR$_{405}$	Inside HVR1
AN2	$_{433}$LNTGFLAALFYTH$_{445}$	AS434
AN3	$_{428}$NCNDSLNTGFLAALFYTHRF$_{447}$	AS434

AN4	539LLNNTRPPRGNWF551	
AN5	599SGPWLTPRCM608	Beside IgVR (572 – 588)

Table 21 – Epitopes identified by Naik et al.

When comparing the positions of the consensus linear epitopes and the epitopes identified by Naik *et al.*, several overlaps can be seen – AN2 and AN3 with LIN_5, AN4 with LIN_1, and AN5 with LIN_4. At the time the paper was released (May 2017), AN5 had not been described by any other research groups.

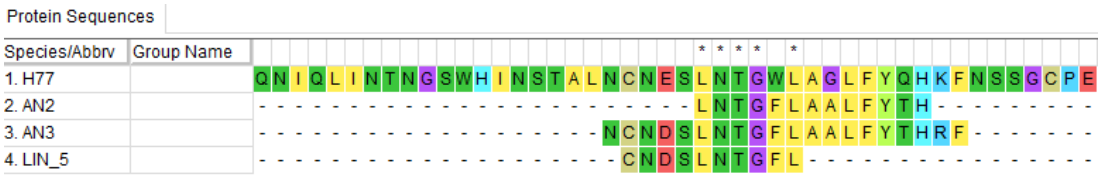


Figure 27 - Alignment of AN2, AN3, and LIN_5 against H77. The sequence "LNTG#L" is conserved across the sequences. The # symbol represents a residue that varies between the sequences. Note the conserved region of amino acids with hydrophobic side chains.

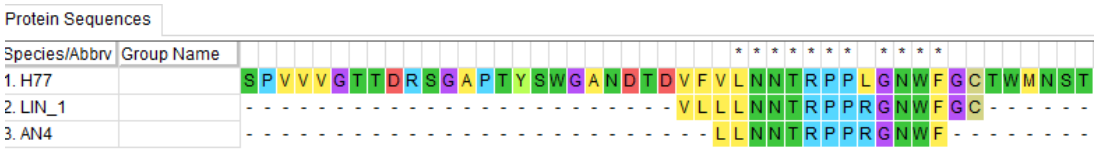


Figure 28 - Alignment of AN4 and LIN_1 against H77. The sequence "LNNTRPP#GNWF" is conserved across sequences. The # symbol represents a residue that varies between the sequences. Regarding the properties of the amino acid residues, note the pattern of hydrophobic -> polar uncharged -> charged side chains.

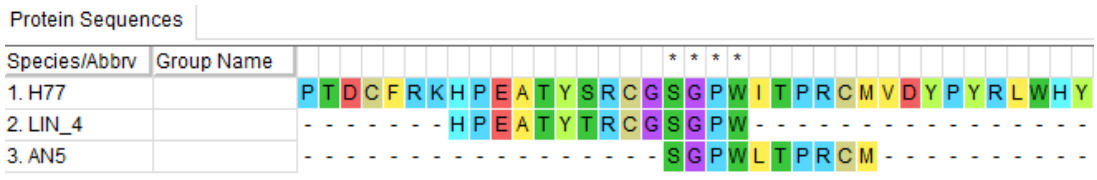


Figure 29 - Alignment of AN5 and LIN_4 against H77. The sequence "SGPW" is conserved across sequences.

Name	Sequence
C_LIN_1	LNTG_L
C_LIN_2	LNNTRPP#GNWF
C_LIN_3	SGPW

Table 22 - Final consensus binding sites. C_LIN meaning "consensus linear". The # symbol indicates an amino acid residue varies across the sequences.

β-Turn Epitopes

Refer to Table 19 for key areas of the E2 glycoprotein.

The E2 ectodomain is a stable protein with a Tm of 85°C. The E2 glycoprotein is modified post-translation by eleven N-glycans and nine conserved disulphide bonds. Four structural faces have been defined on the E2 glycoprotein – the glycan face, the occluded face, the non-neutralising face, and the neutralising face. The majority of the neutralising face is hydrophobic. It has been suggested that non-neutralising mAbs that target HVR1 could shield the neutralising face and protect the HVR1 from the binding of nAbs [181].

For the putative epitopes described by this study, both samples J and G displayed β-Turn epitopes. The most common conformation of AS412 is the β-hairpin conformation. A backbone of internal hydrogen bonds stabilises the conformation. The hydrophobic face of the β-hairpin conformation forms a "binding pocket" containing antibody heavy chains and light chains. Despite AS412 being conserved across many mutations of HCV, HCV evasion of nAbs that target AS412 has been reported for several mutations. One mutation is the N417S/T mutation, which changes glycosylation from N417 to N415. N415 is located inside the "binding pocket" of AS412, and a mutation of N415 causes steric clashes that prevent nAb binding.

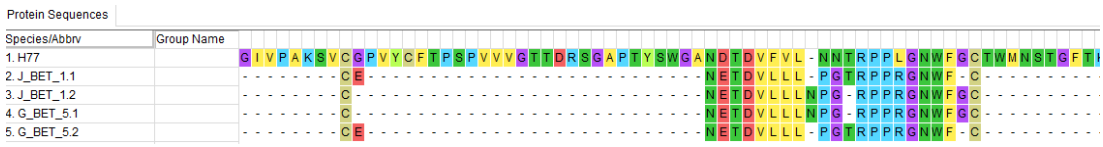


Figure 30 – Alignment of J and G putative β-Turn epitopes with H77.

The J β-Turn epitopes and one of the G epitopes shared a C residue at position 503. There was a "gap" of 29 residues and a conserved section of "NETDV". There was a conserved section of amino acids with hydrophobic side chains. The "RPP" section was conserved across the sequences, with a probable "GNWFGC" section at the end of the epitope. The putative epitope spans positions 503 – 553.

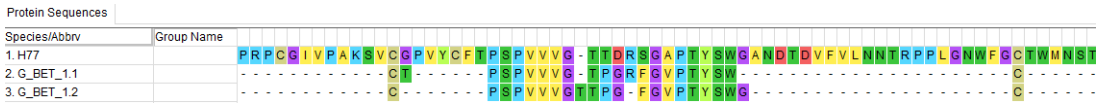


Figure 31 – Alignment of G_1 putative β-Turn epitopes with H77.

G_BET_1.1 and G_BET_1.2 both showed the same C residue at the start and end of the putative epitope as Figure 31. The “PSPVVVG” region and “PTYSW” was also conserved across the sequences. The putative epitope spans positions 503 – 553, the same as above. Combining the two sets of putative epitopes, we can create a consensus epitope. Some of the residues are not completely conserved, the chemical properties of the side chains remained the same.

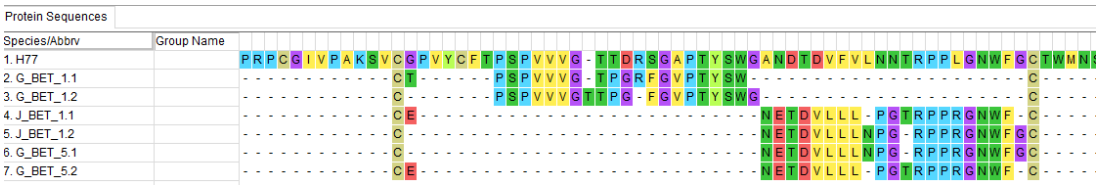


Figure 32 – Alignment of both sets of putative β-Turn epitopes with H77.

The remaining G β-Turn epitopes were not a good match to the H77 sequence, but did contain some of the residues between the 503 – 553 area, with some of them containing the conserved “RPP” region. The β-Turn epitopes are on the E2 β-Sandwich, showing some overlap with AN4, LIN_1, LIN_2, and LIN_6.

Helical Epitopes

Refer to Table 18 for key areas of the E2 glycoprotein.

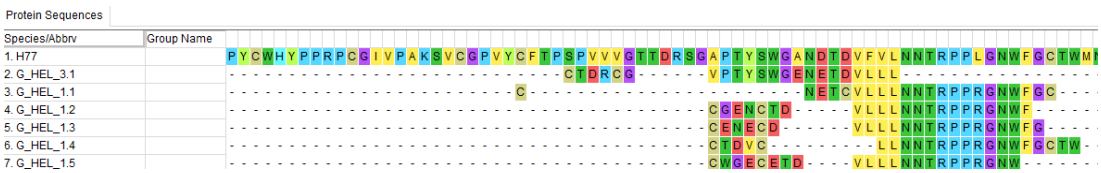


Figure 33 – Alignment of G_1 and G_3 putative helical epitopes with H77.

A number of the G_HEL sequences showed the conserved “LLNNTRPPRGNWF” sequence. The same sequence can be seen when examining the β-Turn epitopes.

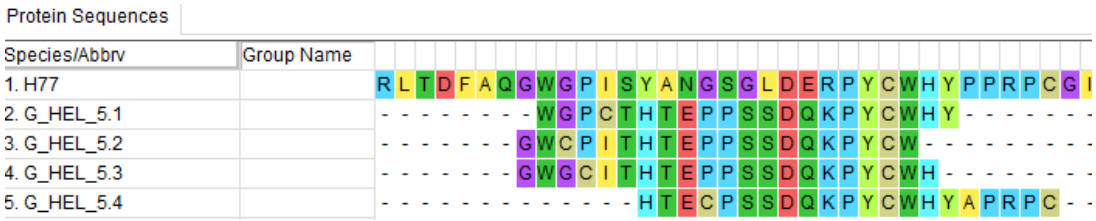


Figure 34 – Alignment of G_5 putative helical epitopes with H77.

The G_HEL_5 putative epitope was not a complete match to the H77 sequence, but did have matching sequences – “PYCWHY” and G_HEL_5.4 having the “PRP” motif. The G_HEL_5 putative epitopes spanned from position 468 – 494 – placing G_HEL_5 inside HVR2.

Loop Epitopes

Refer to Table 18 for key areas of the E2 glycoprotein.

Sample G was the only sample that displayed loop epitopes.

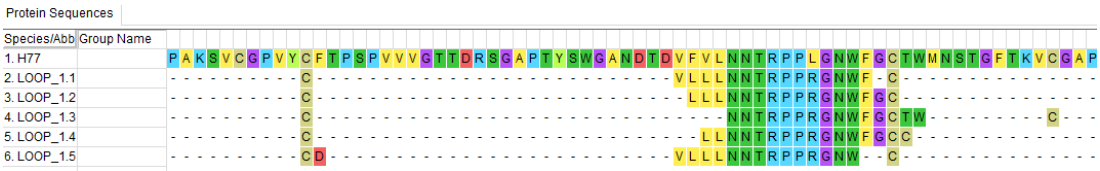


Figure 35 – Alignment of putative loop epitopes with H77.

The LOOP_1 samples showed a conservation of the hydrophobic -> polar uncharged -> charged pattern of side chain properties – similar to the LIN_1, β-Turn, and helical epitopes.

3.5.5 Discussion

There are several ways epitope mapping can be performed, with x-ray crystallography and multi-dimensional nuclear magnetic response (NMR) being the most popular methods. One the major advantages of x-ray crystallography is that it provides the atomic resolution of residues involved with the interaction of the epitope and the complementarity determining regions (CDRs) of the antibody paratope (antigen binding site). There are three CDRs (CDR1 – 3) present on the amino acid sequence of the variable domain of an antigen receptor, and these CDRs determine the specific antibody binding of that particular immunoglobulin [182]. Epitope mapping using the x-ray crystallography technique requires that both the antigen and the antibody are purified before crystallisation of the complex. The expense and the high level of expertise needed to perform x-ray crystallography are two of

the major disadvantages of this technique [183]. NMR can be used to generate precise epitope mapping information but requires that the previous structure of chosen antigen to already be determined. A rough epitope mapping can be performed using the hydrogen-deuterium exchange methodology - this technique can be used on impure antibodies and antigens but has the disadvantage of not giving much information about the sites examined. For instance, the hydrogen-deuterium exchange method might only be able to determine that a sequence of 10 – 20 residues does or does not contain an epitope.

Using a peptide-based approach to epitope mapping can avoid many of the obstacles associated with NMR and x-ray crystallography. As discussed during the methodology, peptides are synthesised to cover the whole sequence of the antigen and immobilised on a surface. This method is ideal for linear peptide sequences, but disulphide bonds can be used to mimic discontinuous and conformational epitopes. This method does not require the same expertise and expense as the NMR and x-ray crystallography methods and does not require purified antigens or antibodies. However, the peptide method might not work for complex conformational epitopes [184]. Another cheap and accessible method of determining epitopes is a mutagenesis approach. Epitopes are determined by examining the binding of the antibody to mutated forms of the antigen. This method can be used to screen thousands of proteins using “shotgun mutagenesis” or alanine scanning and can be combined with x-ray crystallography to determine the contribution of individual amino acids to an antibody-antigen interaction. A disadvantage of the mutagenesis method is that it can be hard to distinguish whether an amino acid change on one of the antigen mutants has changed the tertiary structure (blocking the interaction) or whether it’s a genuine contributory residue [184].

For the purposes of this study, linear epitope mapping was chosen. The epitope mapping was carried out by Pepscan Presto B. V., and four antibodies were used during the procedure – VF-G, VF-J, VF-M, and AP33. AP33 defines an epitope on the E2 glycoprotein ($_{412}$QLINTNGSWHIN$_{423}$ – AKA AS412) and carries a potential *N*-linked glycosylation site.[185]

The majority of bnAbs isolated from infected patients and tested against the E2 glycoprotein target three adjacent sites and block the E2 glycoprotein interacting with the host CD81 receptor. These epitopes are antigenic site 412 - 423 (AS412, AKA Antigenic Domain E or Epitope I), antigenic site 434 – 446 (AS434, AKA Antigenic Domain D or Epitope II), and antigenic region 3 (AR3). As mentioned above, the AS412 includes the epitope that was bound by AP33. The AR3 is a cluster of discontinuous epitopes formed by the E2 front layer (FL) and the CD81 binding loop (CD81bl). These sites have been found to cluster on an exposed surface of the E2 core domain that is devoid of glycans – this area has become known as the E2 neutralising face, and comprises one of the four structural surface regions of the E2 core domain, the other three surface regions being the glycan face, the occluded face, and the non-neutralising face [186] [187].

One region appeared to be conserved across the four categories of putative epitopes – $_{532}$NDTDVFVLNNTRPPLGNWFGC$_{552}$ located on the E2 β-Sandwich of the H77 E2 glycoprotein. Sample G, Sample J, and AN4 demonstrated amino acid variations, most showing a LL substitution for the VF residues and an additional P residue before the GNW section. L, V, and F are amino acids with hydrophobic side chains, meaning that the chemical properties of the section VFVL *could* have been preserved during the change to VLLL. On the other hand, Naik *et al.* found that a single amino acid change in the HVR1 of a HCV genotype 4a isolate was enough to significantly improve the ability of the isolate to evade and resist humoral immune responses. It was found that an F residue at position 399 (with the positions referencing GenBank NC_004102) improved antibody evasion when compared to an isolate with L at the same position. Despite F and L both having hydrophobic side chains, F is aromatic (containing a benzene ring) and L is aliphatic (containing a branched side chain), suggesting that the structure of individual amino acids is enough to impact the epitope, even when certain properties (such as hydrophobicity) are conserved [125].

Name	Origin	Sequence
HE1	H77	$_{532}$NDTDVF**VL**NNT**RPP**LGNWFGC$_{552}$

NW1	Experimental	$_{532}$NDTDVF**LL**NNT**RPPP**LGNWFGC$_{553}$

For ease of discussion, the sequences have been named HE1 and NW1. The changes between HE1 and NW1 are bolded.

Using the BLASTp software, searching HE1 indicated other matches to HCV genotype 1a sequences. However, searching the NW1 sequence using the BLASTp database brought up matches to genotype 1b sequences despite the fact the G and J were isolated from genotype 3a infections. In clinical settings, HCV infections are genotyped by examining a mixture of the core, 5'-UTR, and NS5B regions, but the high genetic variability of HCV can pose problems when distinguishing between subtypes – such as distinguishing between genotype 1a and 1b infections. The accurate genotyping and subtyping of HCV infections is critical for developing robust DAA treatment plans in the age of RAVs. Saludes *et al*. reported that when using a commercial assay that examined the 5'-UTR and NS5B regions, about 5.4% of genotype 1 infections were not assigned a subtype. A second commercial assay (GT *Plus*, Abbot Molecular Inc.) was designed to resolve ambiguous HCV genotype 1 samples and targets the core region – when applied to a selection of ambiguous HCV genotype 1 samples, it was able to assign a subtype to 88.8% of samples (142/160) [188]. Therefore, it is not completely surprising that the BLASTp database matched NW1 to genotype 1b sequences, considering that the sequence covers one of the least conserved regions of the entire HCV genome, and considering that genotyping and subtyping HCV samples can present a challenge even when modern techniques are used.

Kong *et al.* mentioned during their 2016 paper that when using E2 constructs during the course of their research that two sections of HE1 and NW1 are *N*-linked glycosylated[189]. As previously stated, *N*-linked glycosylation is the attachment of an oligosaccharide (a carbohydrate consisting of a number of sugar molecules and known as a glycan) to a nitrogen atom. There are three major types of N-glycans. The first type is the high mannose glycan, which consists of two *N*-acetylglucosamine molecules joined to several mannose residues. The second type is the complex oligosaccharide, which contains two *N*-acetylglucosamine molecules, mannose, and galactose. They can also

contain sialic acid. The third type of glycans are the hybrid glycans, which contain *N*-acetylglucosamine, galactose, mannose, and sialic acid. Hybrid and complex glycans are produced during the movement of the protein through the Golgi apparatus.[79] As shown from

Figure 36, the glycosylation sites present on the portion of the E2 glycoprotein that underwent epitope mapping have asparagine residues (asparagine contains an amide nitrogen).

Figure 36 - Figure from Kong et al. showing the N-linked glycosylation sites of Epitope1 and Epitope2.[189]

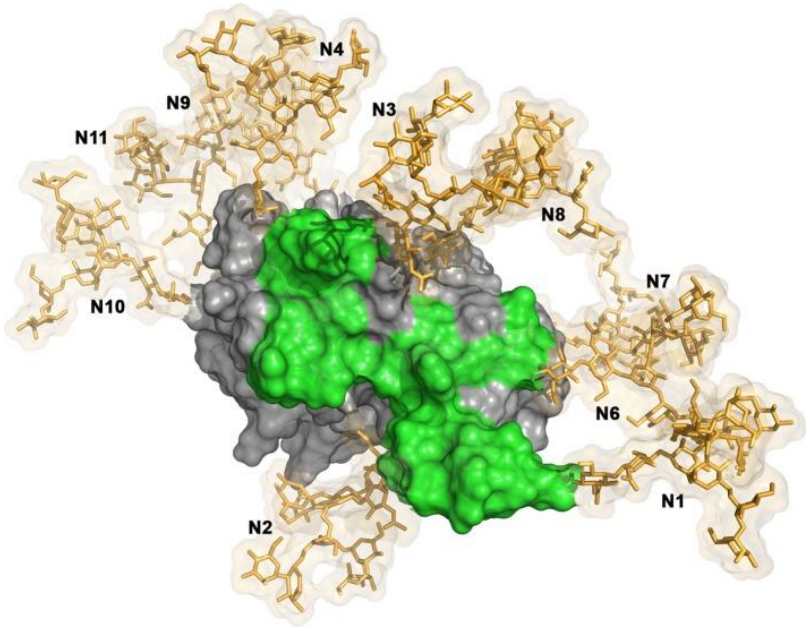


Figure 37 - The glycan shield masking the E2 neutralising epitopes (green area). The structure shows the E2 core and the N-terminal antigenic region (AS412). The HVR1 is not shown. The molecular surface is represented by the grey area, and the glycans are numbered N1 – N11 (gold sticks). Adapted from Lavie et al. [79]

The ectodomains of E1 and E2 are highly glycosylated – see section 1.2.4 for further information. Lavie *et al*. characterised the glycans of the E2 glycoprotein (see Figure 17). E2N6 corresponds to position 532 – the start of HE1 and NW1. E2N8 corresponds to position 556, the glycosylation site just four residues away from HE1 and three residues away from NW1. Both

glycosylation sites are conserved across genotypes. Both Lavie *et al.* and Helle *et al.* state that E2N7 (position 540) is absent in genotypes 3 and 6 - however position 540 corresponds to the NNT sequence displayed by HE1 (genotype 1a) and NW1 (antibodies derived from genotype 3a infections). It would be worth examining whether the same antibodies used for the reverse epitope mapping would be able to detect a NNT sequence on genotype 3a E2 glycoprotein card. Helle *et al.* have also observed the importance of glycosylated asparagine residues in E2 for masking epitopes recognised by bnAbs. For example, the glycosylated asparagine residues at N417 and N423 (termed E2N1 and E2N2) have been associated with decreased sensitivity to bnAbs in HCVcc and HCVpp and increased binding with soluble forms of CD81, suggesting the two glycans might mask the CD81-binding epitope on E2 [81]. In addition to glycan shielding, Pantua *et al*. observed glycan *shifting* where a N417S HCVcc variant was shown to be resistant to two common HCV bnAbs – AP33 and HCV1. Analysis of glycosylation states showed that glycosylation had transferred from N417 to the adjacent N415, enabling the variant to escape the bnAbs. The HCVcc used during this study were challenged with AP33 and underwent UDPS, showing that the immune pressure exerted by AP33 had caused a cumulative 82-fold increase in the number of N417 mutations as soon as five days post treatment [190].

HIV is one of the foremost examples of a virus that utilises glycan shielding. Similarly to HCV, the HIV glycan shield has also been shown to evolve under immune pressure. Key epitopes have been discovered on the HIV Env protein, but the glycan shield surrounding the Env protein prevents nAbs accessing the vulnerable epitopes. A 2017 study aiming to characterise HIV Env glycan shield reported the presence of glycan holes. The omission of specific Env glycans led to the production of glycan holes, exposing epitopes to nAbs. It was noted that the presence of a glycan hole impacted the conformational properties of surrounding glycans and the underlying protein.[191] The 241 glycan site is conserved across 97% of HIV-1 isolates, and the site was recognised by three separate nAb lineages. However, McCoy *et al.* found that even though many HIV-1 isolates lacked more than one conserved glycan site, nearby glycan sites were able to compensate for

missing glycans and withstand nAbs. McCoy *et al.* referenced Wei *et al.*, who found that viruses from one donor lacked a conserved glycan at position 234, but at day 212 the virus was able to escape from the autologous plasma, implying the existence of an evolving glycan shield [192] [193].

Alongside an evolving and adaptable glycan shield, the physical amount of glycan structures present on a virus appears to correlate to the amount of protection the structures grant. A 2020 study investigated the relationship between glycan and oligomannose abundance and the degree of immune evasion granted by glycosylation. The study first examined the glycosylation of the severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS) coronaviruses (CoVs). CoVs are enveloped RNA viruses that cause respiratory diseases in mammals and birds. Vaccine development against SARS-CoV-2 has largely focused on the primary target of the host humoral response; the spike (S) glycoprotein. The trimeric S glycoproteins protrude from the viral envelope and are extensively glycosylated, with 88 – 87 N-linked glycosylation sites per trimeric spike. The CoV S glycoproteins are among the largest Class I viral fusion proteins known, and are used to mediate viral entry and membrane fusion. Watanabe *et al.* utilised a combination of cryo-electron microscopy (cryo-EM) and computational modelling to characterise the glycan shields of the SARS and MERS viruses and then compared them to the glycan shields of “evasion strong” viruses, including HIV-1 and SIV. Glycan shield densities were calculated using analyses of glycoproteins and abundance of oligomannose structures on glycosylated models of the viruses, and by examining the number of aa residues interacting with N-linked glycans to determine the “spread” of the glycan shield over the viral particle. There appeared to be correlation between the density of the glycan shield and the protection that the glycan shield conferred against the humoral immune response. The SARS and MERS viruses were found to have sparse glycan shields and are known to elicit strong nAb responses in humans, whereas HIV-1 was found to have an extremely dense glycan shield and does not induce an effective nAb response in humans [194].

Despite the high degree of evasion conferred by glycosylation, a number of treatments strategies have been proposed for glycosylated viruses. Additionally, improved analytical techniques have helped develop the concept of a “glycome” – the aggregate of glycoproteins within an organism. In 2020, Jia *et al.* used a combination of mass spectrometry, Western blots, and staining techniques to characterise the N- and O-glycome from two human lung samples. Host receptor recognition by the influenza A virus (IAV) has been studied extensively as an example of pathogen-host interaction, and the glycans expressed on the surface of the host’s respiratory tissues are essential for guiding entry and infection by IAV. It was widely believed that hemagglutinin (the IAV envelope protein) recognised glycans with terminal sialic acids (a family of acidic sugars with a backbone of nine carbons) with α2,3- or α2,6-linkages, and that this recognition was required for proper IAV infection. Neuraminidase (present on the IAV surface) later cleaves sialic acid to help release new viral particles replication is completed. However, previous shotgun glycomics work by same research group found that IAVs can also recognise phosphorylated glycans, suggesting that IAVs can recognise and bind to a range of host receptors [195]. The characterisation of the human lung glycome by Jia *et al*. showed that the human lung expresses N-glycans with diverse structures and lengths (sialylated and fucosylated, and a minor population of phosphorylated glycans), and that the O-glycome was dominated by two sialylated structures. These insights could lead to improvements to IAV treatment and vaccine development, and could generate more accurate predictions of how IAV hemagglutinin and neuraminidase variants might bind to the human respiratory system – an approach that could be applied to HCV if the hepatic glycome was explored and characterised [196].

Glycosylation structures tend to be conserved among mammals, leading to the exploration of polysaccharides from non-mammalian sources as a potential therapeutic agents for infections caused by glycosylated viruses. Li *et al.* found that soluble HCV E2 produced from a *Drosophila* S2 culture had a novel glycosylation pattern and had increased immunogenic properties than soluble HCV E2 produced from mammalian cultures [197]. Ahmadi *et al*.

explored the use of algae polysaccharides as therapeutic agents, stating that a number of algae polysaccharides showed promising activity against enveloped viruses [198]. Of particular interest is griffithsin derived from red algae (GRFT), which has demonstrated strong anti-viral activity in vitro and in vivo against a number of enveloped viruses, with minimal toxicity to the host. Initially GRFT was isolated from an aqueous extract of the red algae *Griffithsia* sp. collected from waters off the shores of New Zealand. Researchers at the U.S. National Cancer Institute first reported the potent activity of GRFT against HIV-1 in 2005, noting that GRFT bound to glycoproteins gp120, gp41, and gp160 in a glycosylation-dependent manner – sites known to uses glycan shielding [199]. GRFT is a lectin of 121 amino acids, with its antiviral activity stemming from a unique structural feature that forms a homodimeric complex with three carbohydrate-binding domains on each monomer. These three carbohydrate-binding domains target high-mannose arrays present on many pathogenic enveloped viruses including HIV, SARS-CoV, MERS-CoV, and HCV. Wild-type GRFT has a non-standard amino acid at position 31, but it is replaced by alanine in recombinant GRFT. A number of mechanisms of action of GRFT against HIV-1 have been proposed, including exposure of the CD4 binding site of gp120 through modification of the glycan at position 386 and the inhibition of mannose-binding to gp120 [200]. GRFT has also shown activity against HCV, demonstrating prevention of HCV infection *in vitro* and mitigation of HCV infection for murine models [201,202].

α-glucosidases are ER enzymes necessary for the processing and production of glycoproteins. Qu *et al.* found that downregulating the expression of α-glucosidases in Huh7.5s inhibited the production of HCV and reduced HCV release. Qu *et al*. also explored the use of iminosugars (analogues of sugars where a nitrogen atom has replaced the oxygen atom), which are known to be α-glucosidase inhibitors. The iminosugars proved to inhibit the production of HCV, but the clinical development of iminosugars as drugs has been hampered by the fact that iminosugars struggle to reach therapeutic concentrations *in vivo* – an iminiosugar approved for the treatment of Gaucher's disease was terminated for this reason. Celgosivir

was an α-glucosidase inhibitor created for the treatment of HCV, but was terminated during phase II development because of Celogosivir's plasma drug concentration and poor reduction of viremia (5% of patients experienced a >1-log_{10} reduction). The downregulation of α-glucosidases and/or the use of α-glucosidases inhibitors would avoid the development of drug resistance, because they target host activities, but further research is needed to create viable treatments using this strategy [203] [204].

Another strategy for overcoming the barrier of glycosylation could be the modification or removal of glycosylation sites. Lavie *et al*. found that removal of N-glycans made HCVcc and HCVpp constructs more vulnerable to nAbs. Similarly, a 2019 study by Urbanowicz *et al.* explored the relationship between glycosylation and the antigenicity and immunogenicity of the E2 glycoprotein region. Urbanowicz *et al.* generated *in vitro* variations of 11 E2 *N*-glycosylation sites using mammalian and insect model systems, and but found that modifications to the glycosylation sites did not make them more vulnerable to nAbs. It was only the complete deletion of glycosylation sites that induced increased sensitivity to nAbs – such as the removal of the N1 and N6 glycans (see Figure 37). Urbanowicz *et al.* concluded that future vaccine development should perhaps focus on the complete deletion of specific *N*-glycans, as opposed to just modification [205].

In terms of prophylactic treatments for HCV and other glycosylated viruses, conjugate vaccines (glycans conjugated to an immunogenic protein) have proven to induce protection against encapsulated bacteria. For example, the development of the *Haemophilus influenzae* type B (Hib) conjugate vaccine heralded the almost complete eradication of diseases caused by the pathogen. It was discovered that polyribosyl ribitol phosphate (PRP) was a component of the Hib capsule, but that the administration of anti-PRP antibodies could not induce sustained protection. This was because the PRP antigen was only producing a partial response, activating B lymphocytes but not T lymphocytes – known as a "T independent" (TI) response. The TI response also prevent the formation of memory B lymphocytes, preventing lasting immunity. PRP was conjugated to a immunogenic protein carrier, which was able to induce a complete response and lead to the development

of a working vaccine [206]. This strategy could provide a framework for the development of a conjugate HCV vaccine.

4 DISCUSSION AND CONCLUSION

4.1.1 Discussion

Despite significant drug advancements HCV continues to cause an estimated 399,000 deaths per annum. The WHO estimates that “in the absence of additional efforts, 19 million hepatitis-related deaths are anticipated from 2015 to 2030 [20]. Modern DAA treatment regimens can cure >95% of those diagnosed with hepatitis C infections, but resistance to DAAs have already been observed. The emerging resistance to DAAs, coupled with the problems associated with proper access to diagnostic and treatment programmes for vulnerable populations, only further emphasises the importance of developing improved treatments and – ultimately – a HCV vaccine.

HCV is an important model virus. In a practical sense, the scarcity of suitable animal model systems drove the development of novel *in vitro* systems for HCV, expanding the range of systems that could be used to support other viruses. For instance, the pseudoparticle system has been used to safely study the Lassa virus (an arenavirus that causes Lassa haemorrhagic fever) under Biosafety Level 2 conditions, instead of the Biosafety Level 4 conditions the wildtype Lassa virus requires[207]. From an epidemiological perspective, HCV has served as an important lesson in disease prevention and tracing among vulnerable populations (IDUs, for example).

Understanding the humoral immune response to HCV has been key for shaping our approach to the disease, and epitope mapping has proven to be an important tool for studying the antibody response to HCV infection. Keck *et al.* used epitope mapping to identify epitope determinants for E2 antibody targeting, an important step for the development of a B lymphocyte HCV vaccine [208]. Beyond HCV, epitope mapping has also been used examine the human norovirus for epitopes that could serve as therapeutic targets for vaccines and immunotherapy [209]. During the course of this project, reverse epitope mapping was used to identify a conserved epitope on the HCV E2

glycoprotein. Further examination of the epitope structure and location lead to the conclusion that despite the epitope being conserved across a number of HCV samples, the epitope was protected from nAbs by glycan structures. Glycan shielding has been observed in many enveloped viruses, including HIV-1, and is correlated with high degrees of immune evasion.

To summarise, HCV remains an important virus from both a research perspective and a public health perspective, and the continued research into the humoral response to HCV is crucial for the future management and eradication of the disease.

4.1.2 Future Work

Regarding the continuation of the work done during this research project, the next stage would be isolating HCV E2 from the genotype 3a samples received from the NVRL and using those E2 glycoproteins to produce HCV pseudoparticles. E2N7, which has been stated to be absent from genotype 3 isolates, was detected during this project using antibodies from genotype 3a patients. Using epitope mapping, a larger sample of genotype 3a antibodies could be tested against various HCV genotypes to further examine the presence/absence of E2N7. Glycosylation has been found to impede access to vulnerable epitopes by nAbs, and subsequent work could incorporate some of the methods mentioned in Section 3.5.5 to make the conserved epitopes more accessible to nAbs. SDM could be used to evaluate the importance of various residues from the identified epitopes and the impact they have on resistance to antibodies. The basis of the project (using patient antibodies to examine epitopes through reverse epitope mapping) could also be applied to other viruses known to form AAV complexes – such as Zika virus and Dengue virus [210] [211].

4.1.3 Conclusion

Reverse epitope mapping using antibodies derived from patients with chronic HCV genotype 3a infections identified the conserved epitope $_{532}$NDTDVFLLNNTRPPPLGNWFGC$_{553}$. However, this sequence contains two glycosylation sites (E2N6 and E2N7), with a third glycosylation site (E2N8) present at position 556. The presence of these glycosylation sites may prevent nAbs from accessing this conserved epitope. In conclusion, the

research succeeded in identifying conserved epitopes from patient samples and contributed to the body of knowledge about HCV and the humoral immune response to chronic viral infections.

5 REFERENCES

1. Simmonds, P. *et al.* ICTV Virus Taxonomy Profile: Flaviviridae. doi:10.1099/jgv.0.000672

2. Bukh, J. The history of hepatitis C virus (HCV): Basic research reveals unique features in phylogeny, evolution and the viral life cycle with new perspectives for epidemic control. *J. Hepatol.* **65**, S2S21 (2016).

3. Alazard-Dany, N., Denolly, S., Boson, B. & Cosset, F. L. Overview of hcv life cycle with a special focus on current and possible future antiviral targets. *Viruses* **11**, (2019).

4. Messina, J. P. *et al.* Global distribution and prevalence of hepatitis C virus genotypes. *Hepatology* **61**, 77–87 (2015).

5. Murphy, D. G. *et al.* Hepatitis C virus genotype 7, a new genotype originating from central Africa. *J. Clin. Microbiol.* **53**, 967–72 (2015).

6. Kenny-Walsh, E. Clinical Outcomes after Hepatitis C Infection from Contaminated Anti-D Immune Globulin. *N. Engl. J. Med.* **340**, 1228–1233 (1999).

7. Hepatitis C - Annual Epidemiological Report for 2017. Available at: https://ecdc.europa.eu/en/publications-data/hepatitis-c-annual-epidemiological-report-2017. (Accessed: 28th June 2019)

8. Ocana, S., Casas, M. L., Buhigas, I. & Lledo, J. L. Diagnostic strategy for occult hepatitis B virus infection. *World J. Gastroenterol.* **17**, 1553 (2011).

9. Austria, A. & Wu, G. Y. Occult Hepatitis C Virus Infection: A Review. *J. Clin. Transl. Hepatol.* **6**, 155–160 (2018).

10. Maccallum, F. O. & Bauer, D. J. HOMOLOGOUS SERUM JAUNDICE TRANSMISSION EXPERIMENTS WITH HUMAN VOLUNTEERS. *Lancet* **243**, 622–627 (1944).

11. HOOFNAGLE, J. H. *et al.* Transmission of Non-A, Non-B Hepatitis. *Ann. Intern. Med.* **87**, 14 (1977).

12. Purcell, R. H., Alter, H. J. & Dienstag, J. L. Non-A, non-B hepatitis. *Yale J. Biol. Med.* **49**, 243–50 (1976).

13. Feinman, C. V, Berris, B., Sinclair, J. C. & Wrobel, D. Hepatitis non-A, non-B. *Can. Med. Assoc. J.* **123**, 181–4 (1980).

14. Feinstone, S. M., Kapikian, A. Z., Purcell, R. H., Alter, H. J. & Holland, P. V. Transfusion-Associated Hepatitis Not Due to Viral Hepatitis Type A or B. *N. Engl. J. Med.* **292**, 767–770 (1975).

15. Alter, H. *et al.* CLINICAL AND SEROLOGICAL ANALYSIS OF TRANSFUSION-ASSOCIATED HEPATITIS. *Lancet* **306**, 838–841 (1975).

16. Holland, P. V & Alter, H. J. Non-A, non-B viral hepatitis. *Hum. Pathol.* **12**, 1114–22 (1981).

17. Blum, H. E. & Vyas, G. N. Non-A, non-B hepatitis: a contemporary assessment. *Haematologia (Budap).* **15**, 153–73 (1982).

18. Choo, Q. L. *et al.* Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **244**, 359–62 (1989).

19. Lozano, R. *et al.* Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet (London, England)* **380**, 2095–128 (2012).

20. WHO | Combating hepatitis B and C to reach elimination by 2030. *WHO* (2016).

21. World Health Organization. *Global Hepatitis Report 2017*.

22. Petruzziello, A., Marigliano, S., Loquercio, G. & Cacciapuoti, C. Hepatitis C virus (HCV) genotypes distribution: an epidemiological up-date in Europe. *Infect. Agent. Cancer* **11**, 53 (2016).

23. CDC. HCV FAQs for Health Professionals _ Division of Viral Hepatitis _ CDC. (2015).

24. Schreiber, G. B., Busch, M. P., Kleinman, S. H. & Korelitz, J. J. The

Risk of Transfusion-Transmitted Viral Infections. *N. Engl. J. Med.* **334**, 1685–1690 (1996).

25. Elgharably, A. *et al.* Hepatitis C in Egypt - past, present, and future. *Int. J. Gen. Med.* **10**, 1–6 (2017).

26. Kenny-Walsh, E. Clinical Outcomes after Hepatitis C Infection from Contaminated Anti-D Immune Globulin. *N. Engl. J. Med.* **340**, 1228–1233 (1999).

27. Palmer, B. a *et al.* Analysis of the evolution and structure of a complex intrahost viral population in chronic hepatitis C virus mapped by ultradeep pyrosequencing. *J. Virol.* **88**, 13709–21 (2014).

28. *Comptroller and Auditor General Report on the Accounts of the Public Services 2017*. (2018).

29. Hser, Y.-I., Liang, D., Lan, Y.-C., Vicknasingam, B. K. & Chakrabarti, A. Drug Abuse, HIV, and HCV in Asian Countries. *J. Neuroimmune Pharmacol.* **11**, 383–393 (2016).

30. Martin, N. K. *et al.* Prioritization of HCV treatment in the directacting antiviral era: An economic evaluation the STOPHCV Consortium. (2016). doi:10.1016/j.jhep.2016.02.007

31. Inglis, S. K. *et al.* Randomised controlled trial conducted in injecting equipment provision sites to compare the effectiveness of different hepatitis C treatment regimens in people who inject drugs: A Direct obserVed therApy versus fortNightly CollEction study for HCV treatment - ADVANCE HCV protocol study. *BMJ Open* **9**, (2019).

32. Hutchinson, S. J. *et al.* Expansion of HCV treatment access to people who have injected drugs through effective translation of research into public health policy: Scotland's experience. *International Journal of Drug Policy* **26**, 1041–1049 (2015).

33. Hickman, M. *et al.* Evaluating the population impact of hepatitis C direct acting antiviral treatment as prevention for people who inject drugs (EPIToPe) - A natural experiment (protocol). *BMJ Open* **9**,

(2019).

34. El-Shabrawi, M. H. Burden of pediatric hepatitis C. *World J. Gastroenterol.* **19**, 7880 (2013).

35. Khaderi, S., Shepherd, R., Goss, J. A. & Leung, D. H. Hepatitis C in the pediatric population: transmission, natural history, treatment and liver transplantation. *World J. Gastroenterol.* **20**, 11281–11286 (2014).

36. Garcia-Tejedor, A. *et al.* Risk factors for vertical transmission of hepatitis C virus: A single center experience with 710 HCV-infected mothers. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **194**, 173–177 (2015).

37. Muñoz-Gámez, J. A., Salmerón, J. & Ruiz-Extremera, Á. Hepatitis C durante la gestación, transmisión vertical y nuevas posibilidades de tratamiento. *Med. Clin. (Barc).* **147**, 499–505 (2016).

38. Crowley, D. *et al.* Barriers and facilitators to hepatitis C (HCV) screening and treatment - A description of prisoners' perspective. *Harm Reduct. J.* **15**, (2018).

39. *NACDA - Study on the prevalence of drug use, including intravenous drug use, and blood-borne viruses among the Irish prisoner population*.

40. Martin, N. K. *et al.* Is increased hepatitis C virus case-finding combined with current or 8-week to 12-week direct-acting antiviral therapy cost-effective in UK prisons? A prevention benefit analysis. *Hepatology* **63**, 1796–1808 (2016).

41. Centers for Disease Control and Prevention. Testing for HCV Infection: An Update of Guidance for Clinicians and Laboratories. *Morb. Mortal. Wkly. Rep.* **62**, 362–365 (2013).

42. Pearlman, B. L. & Traub, N. Sustained virologic response to antiviral therapy for chronic hepatitis C virus infection: A cure and so much more. *Clin. Infect. Dis.* **52**, 889–900 (2011).

43. Ohya, K. *et al.* Three children treated with direct-acting antivirals for chronic hepatitis c virus genotype 1b infection. *Intern. Med.* **59**, 941–944 (2020).

44. What happens during hepatitis C treatment - HSE.ie. Available at: https://www.hse.ie/eng/national-hepatitis-c-treatment-programme/hepatitis-c-treatment/what-hepatitis-c-treatment-is/. (Accessed: 14th May 2020)

45. Chevaliez, S. Strategies for the improvement of HCV testing and diagnosis. *Expert Review of Anti-Infective Therapy* **17**, 341–347 (2019).

46. Galli, C., Julicher, P. & Plebani, M. HCV core antigen comes of age: A new opportunity for the diagnosis of hepatitis C virus infection. *Clinical Chemistry and Laboratory Medicine* **56**, 880–888 (2018).

47. Liu, W.-L. *et al.* Ribavirin enhances the action of interferon-α against hepatitis C virus by promoting the p53 activity through the ERK1/2 pathway. *PLoS One* **7**, e43824 (2012).

48. Huang, M., Jiang, J.-D. & Peng, Z. Recent advances in the anti-HCV mechanisms of interferon. *Acta Pharm. Sin. B* **4**, 241–7 (2014).

49. Kohli, A., Shaffer, A., Sherman, A. & Kottilil, S. Treatment of hepatitis C a systematic review. *JAMA - Journal of the American Medical Association* **312**, 631–640 (2014).

50. Baker, D. E. Pegylated interferon plus ribavirin for the treatment of chronic hepatitis C. *Rev. Gastroenterol. Disord.* **3**, 93–109 (2003).

51. Richard Thompson, J. Emerging therapeutic options for the management of hepatitis C infection. *World J. Gastroenterol.* **20**, 7079–7088 (2014).

52. Bertino, G. *et al.* Chronic hepatitis C: This and the new era of treatment. *World J. Hepatol.* **8**, 92–106 (2016).

53. Hinrichsen, H. *et al.* Short-term antiviral efficacy of BILN 2061, a hepatitis C virus serine protease inhibitor, in hepatitis C genotype 1 patients. *Gastroenterology* **127**, 1347–1355 (2004).

54. Kiser, J. J. & Flexner, C. Direct-acting antiviral agents for hepatitis C virus infection. *Annu. Rev. Pharmacol. Toxicol.* **53**, 427–49 (2013).

55. Solbach, P. & Wedemeyer, H. The New Era of Interferon-Free Treatment of Chronic Hepatitis C. *Viszeralmedizin* **31**, 290–296 (2015).

56. Zhang, X. Direct anti-HCV agents. *Acta Pharmaceutica Sinica B* **6**, 26–31 (2016).

57. Ghany, M. G. & Morgan, T. R. Hepatitis C Guidance 2019 Update: American Association for the Study of Liver Diseases–Infectious Diseases Society of America Recommendations for Testing, Managing, and Treating Hepatitis C Virus Infection. *Hepatology* **71**, 686–721 (2020).

58. EASL Recommendations on Treatment of Hepatitis C 2016. *J. Hepatol.* **66**, 153–194 (2017).

59. Chen, Z. W., Li, H., Ren, H. & Hu, P. Global prevalence of pre-existing HCV variants resistant to direct-acting antiviral agents (DAAs): Mining the GenBank HCV genome data. *Sci. Rep.* **6**, (2016).

60. Ramirez, S., Mikkelsen, L. S., Gottwein, J. M. & Bukh, J. Robust HCV Genotype 3a Infectious Cell Culture System Permits Identification of Escape Variants With Resistance to Sofosbuvir. *Gastroenterology* **151**, 973-985.e2 (2016).

61. Lawitz, E. *et al.* Sofosbuvir for previously untreated chronic hepatitis C infection. *N. Engl. J. Med.* **368**, 1878–87 (2013).

62. Goossens, N. & Negro, F. Is genotype 3 of the hepatitis C virus the new villain? *Hepatology* **59**, 2403–2412 (2014).

63. Chevaliez, S. & Pawlotsky, J.-M. *HCV Genome and Life Cycle*. *Hepatitis C Viruses: Genomes and Molecular Biology* (2006). doi:PMID: 21250393

64. Bukh, J., Purcell, R. H. & Miller, R. H. Sequence analysis of the 5' noncoding region of hepatitis C virus. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4942–6 (1992).

65. Shi, S. T. & Lai, M. M. C. HCV 5 ' and 3 ' UTR : When Translation Meets Replication. *Hepat. C Viruses Genomes Mol. Biol.* 49–87

(2006).

66. Honda, M., Brown, E. A. & Lemon, S. M. Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation on hepatitis C virus RNA. *RNA* **2**, 955–68 (1996).

67. Song, □ Yutong *et al.* The Hepatitis C Virus RNA 3J-Untranslated Region Strongly Enhances Translation Directed by the Internal Ribosome Entry Site. *J. Virol.* **80**, 11579–11588 (2006).

68. Yanagi, M., St Claire, M., Emerson, S. U., Purcell, R. H. & Bukh, J. 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**, 2291–5 (1999).

69. Gawlik, K. & Gallay, P. A. HCV core protein and virus assembly: what we know without structures. *Immunologic Research* **60**, 1–10 (2014).

70. Polyak, S. J., Klein, K. C., Shoji, I., Miyamura, T. & Lingappa, J. R. *Assemble and Interact: Pleiotropic Functions of the HCV Core Protein*. *Hepatitis C Viruses: Genomes and Molecular Biology* (Horizon Bioscience, 2006).

71. Li, H.-C., Ma, H.-C., Yang, C.-H. & Lo, S.-Y. Production and pathogenicity of hepatitis C virus core gene products. *World J. Gastroenterol.* **20**, 7104–22 (2014).

72. Irshad, M. & Dhar, I. Hepatitis C Virus Core Protein: An Update on Its Molecular Biology, Cellular Functions and Clinical Implications. *Med Princ Pr.* **15**, 405–416 (2006).

73. Voisset, C. & Dubuisson, J. Functional hepatitis C virus envelope glycoproteins. *Biol. Cell* **96**, 413–413 (2004).

74. Lavie, M., Goffard, A. & Dubuisson, J. *HCV Glycoproteins: Assembly of a Functional E1–E2 Heterodimer*. *Hepatitis C Viruses: Genomes and Molecular Biology* (Horizon Bioscience, 2006).

75. Zhang, M. *et al.* Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV

envelopes and influenza hemagglutinin. *Glycobiology* **14**, 1229–1246 (2004).

76. Castelli, M. *et al.* A Biologically-validated HCV E1E2 Heterodimer Structural Model. *Sci. Rep.* **7**, 214 (2017).

77. Freedman, H. *et al.* Computational Prediction of the Heterodimeric and Higher-Order Structure of gpE1/gpE2 Envelope Glycoproteins Encoded by Hepatitis C Virus. *J. Virol.* **91**, e02309-16 (2017).

78. Sautto, G., Tarr, A. W., Mancini, N. & Clementi, M. Structural and antigenic definition of hepatitis C virus E2 glycoprotein epitopes targeted by monoclonal antibodies. *Clin. Dev. Immunol.* **2013**, 450963 (2013).

79. Lavie, M., Hanoulle, X. & Dubuisson, J. Glycan Shielding and Modulation of Hepatitis C Virus Neutralizing Antibodies. *Front. Immunol.* **9**, 910 (2018).

80. Sandrin, V. *et al.* Assembly of functional hepatitis C virus glycoproteins on infectious pseudoparticles occurs intracellularly and requires concomitant incorporation of E1 and E2 glycoproteins. *J. Gen. Virol.* **86**, 3189–3199 (2005).

81. Helle, F. *et al.* Role of N-Linked Glycans in the Functions of Hepatitis C Virus Envelope Proteins Incorporated into Infectious Virions. *J. Virol.* **84**, 11905–11915 (2010).

82. Madan, V. & Bartenschlager, R. Structural and Functional Properties of the Hepatitis C Virus p7 Viroporin. *Viruses* **7**, 4461–81 (2015).

83. Khaliq, S., Jahan, S. & Hassan, S. Hepatitis C virus p7: molecular function and importance in hepatitis C virus life cycle and potential antiviral target. *Liver Int.* **31**, 606–617 (2011).

84. Sakai, A. *et al.* The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 11646–51 (2003).

85. Gentzsch, J. *et al.* Hepatitis C Virus p7 is Critical for Capsid Assembly

and Envelopment. *PLoS Pathog.* **9**, (2013).

86. Welbourn, S. & Pause, A. *HCV NS2/3 Protease*. *Hepatitis C Viruses: Genomes and Molecular Biology* (Horizon Bioscience, 2006).

87. Kolykhalov, A. A., Mihalik, K., Feinstone, S. M. & Rice, C. M. 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**, 2046–51 (2000).

88. Welbourn, S. & Pause, A. The hepatitis C virus NS2/3 protease. *Curr. Issues Mol. Biol.* **9**, 63–9 (2007).

89. Vega, S. *et al.* NS3 Protease from Hepatitis C Virus: Biophysical Studies on an Intrinsically Disordered Protein Domain. *Int. J. Mol. Sci.* **14**, 13282–13306 (2013).

90. Dimitrova, M., Imbert, I., Kieny, M. P. & Schuster, C. Protein-Protein Interactions between Hepatitis C Virus Nonstructural Proteins. *J. Virol.* **77**, 5401–5414 (2003).

91. Sarrazin, C. *et al.* Prevalence of the hepatitis C virus NS3 polymorphism Q80K in genotype 1 patients in the European region. *Antiviral Res.* **116**, 10–16 (2015).

92. Dietz, J. *et al.* Evolution and function of the HCV NS3 protease in patients with acute hepatitis C and HIV coinfection. *Virology* **485**, 213–222 (2015).

93. Gouttenoire, J., Penin, F. & Moradpour, D. Hepatitis C virus nonstructural protein 4B: a journey into unexplored territory. *Rev. Med. Virol.* **20**, 117–129 (2010).

94. Nakamoto, S., Kanda, T., Wu, S., Shirasawa, H. & Yokosuka, O. Hepatitis C virus NS5A inhibitors and drug resistance mutations. *World J. Gastroenterol.* **20**, 2902–12 (2014).

95. Janardhan, S. V & Reau, N. S. Should NS5A inhibitors serve as the scaffold for all-oral anti-HCV combination therapies? *Hepat. Med.* **7**, 11–20 (2015).

96. Huang, J.-T. *et al.* Hepatitis C virus replication is modulated by the interaction of nonstructural protein NS5B and fatty acid synthase. *J. Virol.* **87**, 4994–5004 (2013).

97. Ahmed, A. & Felmlee, D. J. Mechanisms of Hepatitis C Viral Resistance to Direct Acting Antivirals. *Viruses* **7**, 6716–29 (2015).

98. Petracca, R. *et al.* Structure-Function Analysis of Hepatitis C Virus Envelope-CD81 Binding. *J. Virol.* **74**, 4824–4830 (2000).

99. Kapadia, S. B., Barth, H., Baumert, T., McKeating, J. A. & Chisari, F. V. Initiation of Hepatitis C Virus Infection Is Dependent on Cholesterol and Cooperativity between CD81 and Scavenger Receptor B Type I. *J. Virol.* **81**, 374–383 (2007).

100. Miao, Z. *et al.* Regulated Entry of Hepatitis C Virus into Hepatocytes. *Viruses* **9**, 100 (2017).

101. Kim, C. W. & Chang, K.-M. Hepatitis C virus: virology and life cycle. *Clin. Mol. Hepatol.* **19**, 17–25 (2013).

102. Miyanari, Y. *et al.* The lipid droplet is an important organelle for hepatitis C virus production. *Nat. Cell Biol.* **9**, 1089–1097 (2007).

103. Suzuki, T. Assembly of hepatitis C virus particles. *Microbiol. Immunol.* **55**, 12–18 (2011).

104. Gao, B., Jeong, W.-I. & Tian, Z. Liver: An organ with predominant innate immunity. *Hepatology* **47**, 729–736 (2007).

105. Yi, Z., Chen, J., Kozlowski, M. & Yuan, Z. Innate detection of hepatitis B and C virus and viral inhibition of the response. *Cell. Microbiol.* **17**, 1295–1303 (2015).

106. Schneider, W. M., Chevillotte, M. D. & Rice, C. M. Interferon-stimulated genes: a complex web of host defenses. *Annu. Rev. Immunol.* **32**, 513–45 (2014).

107. Dustin, L. B., Cashman, S. B. & Laidlaw, S. M. Immune control and failure in HCV infection--tipping the balance. *J. Leukoc. Biol.* **96**, 535–

48 (2014).

108. Heim, M. H. & Thimme, R. Innate and adaptive immune responses in HCV infections. *J. Hepatol.* **61**, S14–S25 (2014).

109. Xu, Y. & Zhong, J. Innate immunity against hepatitis C virus. *Curr. Opin. Immunol.* **42**, 98–104 (2016).

110. Lechner, F. *et al.* Analysis of successful immune responses in persons infected with hepatitis C virus. *J. Exp. Med.* **191**, 1499–512 (2000).

111. Larrubia, J. R. *et al.* Adaptive immune response during hepatitis C virus infection. *World J. Gastroenterol.* **20**, 3418–30 (2014).

112. Larrubia, J. R. J. *et al.* Adaptive immune response during hepatitis C virus infection. *World J. Gastroenterol.* **20**, 3418–3430 (2014).

113. Shoukry, N. H. *et al.* Memory CD8 $^{+}$ T Cells Are Required for Protection from Persistent Hepatitis C Virus Infection. *J. Exp. Med.* **197**, 1645–1655 (2003).

114. Penna, A. *et al.* Dysfunction and functional restoration of HCV-specific CD8 responses in chronic hepatitis C virus infection. *Hepatology* **45**, 588–601 (2007).

115. Fallahi, P. *et al.* Cytokines and HCV-related disorders. *Clin. Dev. Immunol.* **2012**, 468107 (2012).

116. Chen, P. J. *et al.* Transient immunoglobulin M antibody response to hepatitis C virus capsid antigen in posttransfusion hepatitis C: putative serological marker for acute viral infection. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 5971–5 (1992).

117. Nikolaeva, L. I. *et al.* Virus-specific antibody titres in different phases of hepatitis C virus infection. *J. Viral Hepat.* **9**, 429–37 (2002).

118. Yoon, J. C., Shiina, M., Ahlenstiel, G. & Rehermann, B. Natural killer cell function is intact after direct exposure to infectious hepatitis C virions. *Hepatology* **49**, 12–21 (2009).

119. Long, L. *et al.* Non-neutralizing epitopes induce robust hepatitis C virus

(HCV)-specific antibody-dependent CD56+ natural killer cell responses in chronic HCV-infected patients. *Clin. Exp. Immunol.* **189**, 92–102 (2017).

120. Cashman, S. B., Marsden, B. D. & Dustin, L. B. The Humoral Immune Response to HCV: Understanding is Key to Vaccine Development. *Front. Immunol.* **5**, 550 (2014).

121. Osburn, W. O. *et al.* Clearance of hepatitis C infection is associated with the early appearance of broad neutralizing antibody responses. *Hepatology* **59**, 2140–2151 (2014).

122. Sacks-Davis, R. *et al.* High Rates of Hepatitis C Virus Reinfection and Spontaneous Clearance of Reinfection in People Who Inject Drugs: A Prospective Cohort Study. *PLoS One* **8**, e80216 (2013).

123. Liu, L. *et al.* Acceleration of Hepatitis C Virus Envelope Evolution in Humans Is Consistent with Progressive Humoral Immune Selection during the Transition from Acute to Chronic Infection. *J. Virol.* **84**, 5067–5077 (2010).

124. Cashman, S. B., Marsden, B. D. & Dustin, L. B. The Humoral Immune Response to HCV: Understanding is Key to Vaccine Development. *Front. Immunol.* **5**, 550 (2014).

125. Naik, A. S., Palmer, B. A., Crosbie, O., Kenny-Walsh, E. & Fanning, L. J. A single amino acid change in the hypervariable region 1 of hepatitis C virus genotype 4a aids humoral immune escape. *J. Gen. Virol.* **97**, 1345–1349 (2016).

126. Helle, F., Duverlie, G. & Dubuisson, J. The Hepatitis C Virus Glycan Shield and Evasion of the Humoral Immune Response. *Viruses* **3**, 1909–1932 (2011).

127. Bassendine, M. F., Sheridan, D. A., Bridge, S. H., Felmlee, D. J. & Neely, R. D. G. Lipids and HCV. *Semin. Immunopathol.* **35**, 87–100 (2013).

128. Sheridan, D. A. *et al.* Maximum levels of hepatitis C virus lipoviral

particles are associated with early and persistent infection. *Liver Int.* **36**, 1774–1782 (2016).

129. Robinson, M. W., Harmon, C. & O'Farrelly, C. Liver immunology and its role in inflammation and homeostasis. *Cellular and Molecular Immunology* **13**, 267–276 (2016).

130. Bilzer, M., Roggel, F. & Gerbes, A. L. Role of Kupffer cells in host defense and liver disease. *Liver Int.* **26**, 1175–1186 (2006).

131. Li, K. *et al.* Activation of chemokine and inflammatory cytokine response in hepatitis C virus-infected hepatocytes depends on toll-like receptor 3 sensing of hepatitis C virus double-stranded RNA intermediates. *Hepatology* **55**, 666–675 (2012).

132. Horner, S. M. & Gale, M. Regulation of hepatic innate immunity by hepatitis C virus. *Nature Medicine* **19**, 879–888 (2013).

133. Kelly, A. *et al.* CD141+ myeloid dendritic cells are enriched in healthy human liver. *J. Hepatol.* **60**, 135–142 (2014).

134. Stone, A. E. L. *et al.* Hepatitis C Virus Pathogen Associated Molecular Pattern (PAMP) Triggers Production of Lambda-Interferons by Human Plasmacytoid Dendritic Cells. *PLoS Pathog.* **9**, (2013).

135. Takahashi, K. *et al.* Plasmacytoid dendritic cells sense hepatitis C virus-infected cells, produce interferon, and inhibit infection. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 7431–7436 (2010).

136. Baiocchini, A. *et al.* Liver sinusoidal endothelial cells (LSECs) modifications in patients with chronic hepatitis C. *Sci. Rep.* **9**, (2019).

137. Greuter, T. & Shah, V. H. Hepatic sinusoids in liver injury, inflammation, and fibrosis: new pathophysiological insights. *Journal of Gastroenterology* **51**, 511–519 (2016).

138. García-Mediavilla, M. V. *et al.* Liver X receptor α-mediated regulation of lipogenesis by core and NS5A proteins contributes to HCV-induced liver steatosis and HCV replication. *Lab. Investig.* **92**, 1191–1202 (2012).

139. Xing, Y., Zhao, T., Gao, X. & Wu, Y. Liver X receptor α is essential for the capillarization of liver sinusoidal endothelial cells in liver injury. *Sci. Rep.* **6**, (2016).

140. Drane, D. *et al.* Priming of CD4+ and CD8+ T cell responses using a HCV core ISCOMATRIX vaccine: a phase I study in healthy volunteers. *Hum. Vaccin.* **5**, 151–7 (2009).

141. Yokokawa, H. *et al.* Induction of humoural and cellular immunity by immunisation with HCV particle vaccine in a non-human primate model. *Gut* **67**, 372–379 (2018).

142. Frey, S. E. *et al.* Safety and immunogenicity of HCV E1E2 vaccine adjuvanted with MF59 administered to healthy adults. *Vaccine* **28**, 6367–73 (2010).

143. Dahari, H., Feinstone, S. M. & Major, M. E. Meta-analysis of hepatitis C virus vaccine efficacy in chimpanzees indicates an importance for structural proteins. *Gastroenterology* **139**, 965–74 (2010).

144. Fuerst, T. R., Pierce, B. G., Keck, Z.-Y. & Foung, S. K. H. Designing a B Cell-Based Vaccine against a Highly Variable Hepatitis C Virus. *Front. Microbiol.* **8**, 2692 (2018).

145. Vietheer, P. T. *et al.* The core domain of hepatitis C virus glycoprotein E2 generates potent cross-neutralizing antibodies in guinea pigs. *Hepatology* **65**, 1117–1131 (2017).

146. Center, R. J. *et al.* Enhancing the antigenicity and immunogenicity of monomeric forms of hepatitis C virus E2 for use as a preventive vaccine. *J. Biol. Chem.* **295**, jbc.RA120.013015 (2020).

147. Jardine, J. *et al.* Rational HIV immunogen design to target specific germline B cell receptors. *Science* **340**, 711–6 (2013).

148. Swadling, L., Klenerman, P. & Barnes, E. Ever closer to a prophylactic vaccine for HCV. *Expert Opin. Biol. Ther.* **13**, 1109–24 (2013).

149. Rosen, H. R. “Hep C, where art thou”: What are the remaining (fundable) questions in hepatitis C virus research? *Hepatology* **65**,

341–349 (2017).

150. Hahn, J. A. *et al.* Potential impact of vaccination on the hepatitis C virus epidemic in injection drug users. *Epidemics* **1**, 47–57 (2009).

151. Scott, N. *et al.* The case for a universal hepatitis C vaccine to achieve hepatitis C elimination. *BMC Med.* **17**, 175 (2019).

152. Vercauteren, K., de Jong, Y. P. & Meuleman, P. Animal models for the study of HCV. *Curr. Opin. Virol.* **13**, 67–74 (2015).

153. Dorner, M. *et al.* Completion of the entire hepatitis C virus life cycle in genetically humanized mice. *Nature* **501**, 237–241 (2013).

154. Tariq, H. *et al.* An overview: In vitro models of HCV replication in different cell cultures. *Infect. Genet. Evol.* **12**, 13–20 (2012).

155. Blight, K. J., McKeating, J. A. & Rice, C. M. Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. *J. Virol.* **76**, 13001–14 (2002).

156. Lohmann, V. *et al.* Replication of Subgenomic Hepatitis C Virus RNAs in a Hepatoma Cell Line. *Science (80-.).* **285**, (1999).

157. Ashfaq, U. A., Khan, S. N., Nawaz, Z. & Riazuddin, S. In-vitro model systems to study Hepatitis C Virus. *Genet. Vaccines Ther.* **9**, 7 (2011).

158. Catanese, M. T. & Dorner, M. Advances in experimental systems to study hepatitis C virus in vitro and in vivo. *Virology* **479480**, 221233 (2015).

159. Tello, D. *et al.* Expression and structural properties of a chimeric protein based on the ectodomains of E1 and E2 hepatitis C virus envelope glycoproteins. *Protein Expr. Purif.* **71**, 123–131 (2010).

160. Whidby, J. *et al.* Blocking hepatitis C virus infection with recombinant form of envelope protein 2 ectodomain. *J. Virol.* **83**, 11078–89 (2009).

161. Bartosch, B. & Cosset, F.-L. Studying HCV Cell Entry with HCV Pseudoparticles (HCVpp). in 279–293 (Humana Press, 2009). doi:10.1007/978-1-59745-394-3_21

162. Bartosch, B., Dubuisson, J. & Cosset, F.-L. Infectious Hepatitis C Virus Pseudo-particles Containing Functional E1–E2 Envelope Protein Complexes. *J. Exp. Med. J. Exp. Med* **036331000**, 633–642 (2003).

163. Urbanowicz, R. A. *et al.* A Diverse Panel of Hepatitis C Virus Glycoproteins for Use in Vaccine Research Reveals Extremes of Monoclonal Antibody Neutralization Resistance. *J. Virol.* **90**, 3288–301 (2015).

164. Lavillette, D. *et al.* Characterization of host-range and cell entry properties of the major genotypes and subtypes of hepatitis C virus. *Hepatology* **41**, 265–274 (2005).

165. Urbanowicz, R. A. *et al.* Novel functional hepatitis C virus glycoprotein isolates identified using an optimized viral pseudotype entry assay. doi:10.1099/jgv.0.000537

166. Bailey, J. R., Urbanowicz, R. A., Ball, J. K., Law, M. & Foung, S. K. H. Standardized method for the study of antibody neutralization of HCV pseudoparticles (HCVpp). in *Methods in Molecular Biology* **1911**, 441–450 (Humana Press Inc., 2019).

167. Soares, H. R. *et al.* Enhancing Hepatitis C virus pseudoparticles infectivity through p7NS2 cellular expression. *J. Virol. Methods* **274**, 113714 (2019).

168. Kato, T. *et al.* Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon. *Gastroenterology* **125**, 1808–1817 (2003).

169. Simister, P. *et al.* Structural and functional analysis of hepatitis C virus strain JFH1 polymerase. *J. Virol.* **83**, 11926–39 (2009).

170. Gottwein, J. M. *et al.* Robust Hepatitis C Genotype 3a Cell Culture Releasing Adapted Intergenotypic 3a/2a (S52/JFH1) Viruses. *Gastroenterology* **133**, 1614–1626 (2007).

171. Mcclure, C. P. *et al.* Flexible and rapid construction of viral chimeras applied to hepatitis C virus. doi:10.1099/jgv.0.000530

172. Mathiesen, C. K. *et al.* Production and characterization of high-titer

serum-free cell culture grown hepatitis C virus particles of genotype 1-6. *Virology* **458**–**459**, 190–208 (2014).

173. Binder, M. *et al.* Identification of determinants involved in initiation of hepatitis C virus RNA synthesis by using intergenotypic replicase chimeras. *J. Virol.* **81**, 5270–83 (2007).

174. Petruzziello, A., Marigliano, S., Loquercio, G., Cozzolino, A. & Cacciapuoti, C. Global epidemiology of hepatitis C virus infection: An up-date of the distribution and circulation of hepatitis C virus genotypes. *World J. Gastroenterol.* **22**, 7824–40 (2016).

175. Kuiken, C., Yusim, K., Boykin, L. & Richardson, R. The Los Alamos hepatitis C sequence database. *Bioinformatics* **21**, 379–384 (2005).

176. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. (2015).

177. Naik, A. S. *et al.* Reverse epitope mapping of the E2 glycoprotein in antibody associated hepatitis C virus. *PLoS One* **12**, e0175349 (2017).

178. Felsenstein, J. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evolution (N. Y).* **39**, 783–791 (1985).

179. Parra, M., Laufer, N., Manrique, J. M., Jones, L. R. & Quarleri, J. Phylogenetic Diversity in Core Region of Hepatitis C Virus Genotype 1a as a Factor Associated with Fibrosis Severity in HIV-1-Coinfected Patients. *Biomed Res. Int.* **2017**, 1728456 (2017).

180. Kao, C. C., Yi, G. & Huang, H.-C. The core of hepatitis C virus pathogenesis. *Curr. Opin. Virol.* **17**, 66–73 (2016).

181. Keck, Z. *et al.* Antibody Response to Hypervariable Region 1 Interferes with Broadly Neutralizing Antibodies to Hepatitis C Virus. *J. Virol.* **90**, 3112–22 (2016).

182. Polonelli, L. *et al.* Antibody complementarity-determining regions (CDRs) can display differential antimicrobial, antiviral and antitumor activities. *PLoS One* **3**, e2371 (2008).

183. Saul, F. A. & Alzari, P. M. Crystallographic Studies of Antigen–Antibody Interactions. in *Epitope Mapping Protocols* 11–24 (Humana Press, 1996). doi:10.1385/0-89603-375-9:11

184. Abbott, W. M., Damschroder, M. M. & Lowe, D. C. Current approaches to fine mapping of antigen-antibody interactions. *Immunology* **142**, 526–35 (2014).

185. Owsianka, A. *et al.* Monoclonal antibody AP33 defines a broadly neutralizing epitope on the hepatitis C virus E2 envelope glycoprotein. *J. Virol.* **79**, 11095–104 (2005).

186. Tzarum, N., Wilson, I. A. & Law, M. The Neutralizing Face of Hepatitis C Virus E2 Envelope Glycoprotein. *Front. Immunol.* **9**, 1315 (2018).

187. Kong, L. *et al.* Hepatitis C virus e2 envelope glycoprotein core structure. *Sci. (New York, NY)* **342**, 1090–1094 (2013).

188. Saludes, V. *et al.* Reliable resolution of ambiguous hepatitis C virus genotype 1 results with the Abbott HCV Genotype Plus RUO assay. *Sci. Rep.* **9**, (2019).

189. Kong, L. *et al.* Structural flexibility at a major conserved antibody target on hepatitis C virus E2 antigen. *Proc. Natl. Acad. Sci.* **113**, 12768–12773 (2016).

190. Pantua, H. *et al.* Glycan shifting on hepatitis C virus (HCV) E2 glycoprotein is a mechanism for escape from broadly neutralizing antibodies. *J. Mol. Biol.* **425**, 1899–1914 (2013).

191. Yang, M. *et al.* Conformational Heterogeneity of the HIV Envelope Glycan Shield. *Sci. Rep.* **7**, 4435 (2017).

192. McCoy, L. E. *et al.* Holes in the Glycan Shield of the Native HIV Envelope Are a Target of Trimer-Elicited Neutralizing Antibodies. *Cell Rep.* **16**, 2327–38 (2016).

193. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).

194. Watanabe, Y. *et al.* Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat. Commun.* **11**, 1–10 (2020).

195. Byrd-Leotis, L. *et al.* Influenza binds phosphorylated glycans from human lung. *Sci. Adv.* **5**, (2019).

196. Jia, N. *et al.* The Human Lung Glycome Reveals Novel Glycan Ligands for Influenza A Virus. *Sci. Rep.* **10**, (2020).

197. Li, D. *et al.* Altered Glycosylation Patterns Increase Immunogenicity of a Subunit Hepatitis C Virus Vaccine, Inducing Neutralizing Antibodies Which Confer Protection in Mice. *J. Virol.* **90**, 10486–10498 (2016).

198. Ahmadi, A., Zorofchian Moghadamtousi, S., Abubakar, S. & Zandi, K. Antiviral Potential of Algae Polysaccharides Isolated from Marine Sources: A Review. *Biomed Res. Int.* **2015**, 825203 (2015).

199. Mori, T. *et al.* Isolation and characterization of Griffithsin, a novel HIV-inactivating protein, from the red alga Griffithsia sp. *J. Biol. Chem.* **280**, 9345–9353 (2005).

200. Lee, C. Griffithsin, a highly potent broad-spectrum antiviral lectin from red algae: From discovery to clinical application. *Marine Drugs* **17**, (2019).

201. Takebe, Y. *et al.* Antiviral Lectins from Red and Blue-Green Algae Show Potent In Vitro and In Vivo Activity against Hepatitis C Virus. *PLoS One* **8**, (2013).

202. Meuleman, P. *et al.* Griffithsin has antiviral activity against hepatitis C virus. *Antimicrob. Agents Chemother.* **55**, 5159–5167 (2011).

203. Qu, X. *et al.* Inhibitors of endoplasmic reticulum alpha-glucosidases potently suppress hepatitis C virus virion assembly and release. *Antimicrob. Agents Chemother.* **55**, 1036–44 (2011).

204. Durantel, D. Celgosivir, an alpha-glucosidase I inhibitor for the potential treatment of HCV infection. *Curr. Opin. Investig. Drugs* **10**, 860–70 (2009).

205. Urbanowicz, R. A. *et al.* Antigenicity and Immunogenicity of Differentially Glycosylated Hepatitis C Virus E2 Envelope Proteins Expressed in Mammalian and Insect Cells. *J. Virol.* **93**, 1403–1421 (2019).

206. Robbins, J. B., Schneerson, R., Anderson, P. & Smith, D. H. The 1996 Albert Lasker Medical Research Awards. Prevention of systemic infections, especially meningitis, caused by Haemophilus influenzae type b. Impact on public health and implications for other polysaccharide-based vaccines. *JAMA* **276**, 1181–5 (1996).

207. Cosset, F.-L. *et al.* Characterization of Lassa virus cell entry and neutralization with Lassa virus pseudoparticles. *J. Virol.* **83**, 3228–37 (2009).

208. Keck, Z.-Y. *et al.* Broadly neutralizing antibodies from an individual that naturally cleared multiple hepatitis C virus infections uncover molecular determinants for E2 targeting and vaccine design. *PLOS Pathog.* **15**, e1007772 (2019).

209. van Loben Sels, J. M. & Green, K. Y. The Antigenic Topology of Norovirus as Defined by B and T Cell Epitope Mapping: Implications for Universal Vaccines and Therapeutics. *Viruses* **11**, 432 (2019).

210. Dai, L., Wang, Q., Song, H. & Gao, G. F. Zika Virus Envelope Protein and Antibody Complexes. in *Sub-cellular biochemistry* **88**, 147–168 (2018).

211. Moi, M. L., Takasaki, T. & Kurane, I. Detection of Virus-Antibody Immune Complexes in Secondary Dengue Virus Infection. in *Methods in molecular biology (Clifton, N.J.)* **1604**, 331–337 (2018).

APPENDIX I – SEQUENCES USED

HCV Genotype 1a H77 Reference Sequence

GenBank: NC_004102.1 – Hepatitis C Virus Genotype 1 – Complete Genome

HCV Genotype 3a Reference Sequence

GenBank: AF046866.1 – Hepatitis C Virus Subtype 3a Strain CB Polyprotein Gene – Complete CDS

320bp Sequence Analysis – Accession Numbers of Sequences Used for Figure 9

n = 20

KU897989.1	KU897985.1	KU897981.1	KU897976.1
KU897988.1	KU897984.1	KU897979.1	KU897975.1
KU897987.1	KU897983.1	KU897978.1	KU897974.1
KU897986.1	KU897982.1	KU897977.1	KU897973.1
KU897972.1	KU897970.1	KU897969.1	KU897947.1

320bp Sequence Analysis – Accession Numbers of Sequences Used for Figure 10

n = 100

NC_009824.1	KY620418.1	DQ506019.1	KC967545.1
D14305.1	KY620398.1	DQ506014.1	GQ356216.1
KU180722.1	DQ506538.1	DQ506003.1	AM423059.1
HQ108106.1	DQ506530.1	DQ505999.1	DQ506536.1
KY620509.1	DQ506518.1	DQ505996.1	DQ506529.1
KJ437318.2	DQ506515.1	DQ505978.1	DQ506517.1
KY620513.1	DQ506506.1	DQ505977.1	DQ506514.1

EU399722.1	DQ506505.1	DQ505969.1	DQ506507.1
KY620492.1	DQ506503.1	DQ505958.1	DQ506481.1
DQ506468.1	DQ506480.1	DQ505923.1	DQ506464.1
KY704893.1	DQ506466.1	DQ505917.1	DQ506026.1
DQ506516.1	DQ506461.1	DQ505913.1	DQ506000.1
DQ506473.1	LC368352.1	DQ505881.1	DQ505993.1
DQ506462.1	KY620603.1	DQ505878.1	DQ505986.1
DQ506460.1	KY620474.1	DQ505876.1	DQ505965.1
DQ506459.1	KY620375.1	AF416523.1	DQ505959.1
DQ506458.1	DQ506533.1	KJ470613.1	DQ505841.1
DQ506030.1	DQ506528.1	DQ506498.1	DQ505818.1
D14311.1	DQ506526.1	LC368368.1	DQ437509.2
KJ470614.1	DQ506510.1	KY704903.1	KY620387.1
DQ506476.1	DQ506509.1	KY620462.1	MN628590.1
KY620601.1	DQ506504.1	KY620396.1	KY704902.1
KY620593.1	DQ506501.1	KY120332.1	KY704900.1
KY620514.1	DQ506491.1	KX621475.1	KY620760.1
KY620459.1	DQ506483.1	KX621470.1	KT735468.1

Core Protein Sequence Analysis – Accession Numbers of Sequences Used for Figure 11 and Figure 12

n = 100

NC_009824.1	MK612031.1	MN026653.1	MN026688.1
D14305.1	KY620776.1	MN026580.1	MN026652.1
KY620603.1	KY620601.1	MN026568.1	MN026624.1
KY704903.1	KY620490.1	KY620616.1	MN026571.1
KY704902.1	KX621518.1	KY620516.1	MN026734.1
KY620510.1	KF292092.1	LC368352.1	AM263092.1
JX826592.1	KC143935.1	MN026674.1	AM263080.1
MN026630.1	HQ912953.1	MN026597.1	AM263079.1
AY522117.1	KY620602.1	MN026560.1	MN026676.1
AY522116.1	KY620561.1	KY620652.1	MN026570.1

AY522110.1	KY620524.1	KY620585.1	KY704901.1
AY522106.1	KY620502.1	KY620501.1	KY620719.1
AY522102.1	KJ437307.2	KY620391.1	KY620634.1
U10191.1	AY522115.1	KJ437314.2	KY620625.1
D14311.1	AY522114.1	KF292108.1	KY620613.1
MN026612.1	AY522111.1	MT632179.1	KY620606.1
MN026573.1	AY522108.1	GU814263.1	KY620605.1
KY620560.1	AY522104.1	AY522118.1	KY620554.1
LC368368.1	AY522103.1	U10210.1	KY620539.1
MN026668.1	AY522101.1	U10208.1	KY620537.1
MN026561.1	U10197.1	L12355.1	KY620532.1
KY620525.1	KY620461.1	KF292056.1	JF343784.2
KY620508.1	KY620456.1	KF589884.1	GU814264.1
KY620500.1	KY620366.1	KF134008.1	EU204645.1
KY620494.1	KX621530.1	JN050971.1	MF039291.1

APPENDIX II – EPITOPE MAPPING REPORT RECEIVED FROM PEPSCAN PRESTO B. V.

FINAL REPORT FOR UNIVERSITY COLLEGE CORK

CONFORMATIONAL EPITOPE MAPPING

HCV-E1E2 GLYCOPROTEIN

STUDY REPORT

STUDY NUMBER EM20180905068

TABLE OF CONTENTS

TABLE OF FIGURES

EXECUTIVE SUMMARY

GOAL

To map the epitopes of antibodies that recognize HCV-E1E2 glycoprotein using linear and conformational epitope mapping.

RESULTS

Pepscan analysis established tentative epitopes for three out of four samples investigated in this study. Details of the epitope information is illustrated in Table 6.

REFERENCES

The study was conducted at Pepscan Presto BV, (Zuidersluisweg 2, 8243RC Lelystad, The Netherlands). All Pepscan data is stored in the software package Peplab™, a proprietary database application developed in-house and built on a PostgreSQL storage back-end. The database has daily off-site back-ups on a private secure server. The screening department has additional lab books and hard-copies of all results. These are stored in the Pepscan laboratory which is electronically secured.

STUDY DATES

Table 1. Study credentials

Order accepted	November 2018
Synthesis of the array	Feb.-March 2016
Screening of peptides	PEP264: pp. 94, 96, 98, 101, 106, 108, 110, 114, 122 PEP265: pp. 2, 5, 8, 11, 14

PERSONNEL & FACILITIES

Table 2. Personnel

Study supervisor	Dr. D. Fliermann
Study review	Dr. T. Satav
Peptide synthesis	J. Buijnink, C. Lasthuis, R. Teeuwen
Array screening	D. Parohi, E. Fijlstra

MATERIALS

General materials used: See Methods section.

PROVIDED DESCRIPTION OF ANTIBODIES

Samples provided by University College Cork that recognize HCV-E1E2 glycoprotein are listed below (Figure 3).

Table 3. Sample information

Name	Origin	Concentration	Location
G	Human	707 µg/ml	-20°C/118
J	Human	5590 µg/ml	-20°C/118
M	Human	3690 µg/ml	-20°C/118
AP33	Mouse	25000 µg/ml	-20°C/118

TARGET PROTEIN

PROVIDED PROTEIN SEQUENCE

>fragment of HCV-E1E2 glycoprotein (plasmid 5)

```
  1 ETHTVGGSAS RAAHRVTTFI TRGPSQNIQL INTNGSWHIN RTALNCNDSL 50
 51 NTGFLAALFY THRFNASGCP ERMASCRPID KFAQGWGPIT HTEPPSSDQK 100
101 PYCWHYAPRP CGIVPAAQVC GPVYCFTPSP VVVGTTDRFG VPTYSWGENE 150
151 TDVLLLNNTR PPRGNWFGCT WMNSTGFTKT CGGPPCNIGG VGNDTLTCPT 200
201 DCFRKHPEAT YTRCGSGPWL TPRCMVHYPY 230
```

DATA VISUALIZATION

Data will be visualized using intensity plots, and will be available as raw data. If deemed useful, putative epitopes will be depicted onto the structure of the HCV E2 protein from file 4mwf.pdb.

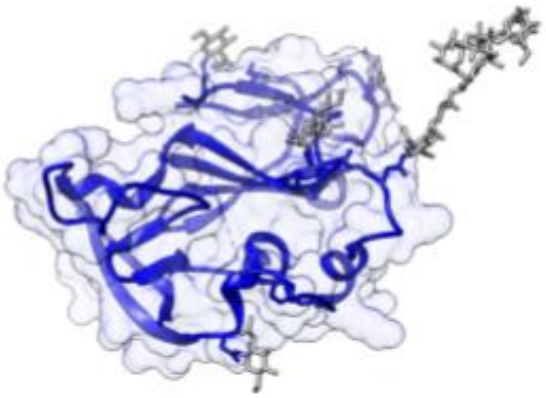


Figure 1. Rendering of 4mwf.pdb In blue, HCV E2 protein structure depicted as ribbon with transparent surface representation. Sugar moieties are depicted in gray.

CONCEPTS & PRINCIPLES

LINEAR EPITOPE MAPPING WITH ALL OVERLAPPING PEPTIDES

The concept of mapping linear epitopes using libraries of overlapping synthetic peptides was pioneered by Pepscan founders Geysen and Meloen (PNAS, 1984). As the inventor of the technology Pepscan has long standing expertise in addressing linear epitopes by directly synthesizing libraries of linear peptides on a solid support covered with a proprietary hydrogel formulation, which allows working with biomolecules and can be easily regenerated for profiling big sample sets. To generate a library of linear mimics, the correct amino acid sequence of the immunogen (or target protein) is split in overlapping fragments *in silico*, which are then synthesized on a solid support as shown in Figure 2.

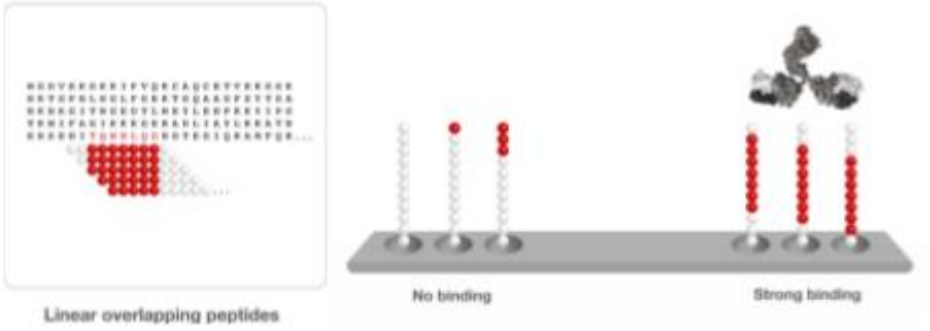


Figure 2. Constructing a linear peptide library. The target linear sequence is converted into a library of all overlapping linear peptides directly synthesized on a proprietary solid support called "mini-card". Binding of antibodies is quantified using an automated ELISA-type read-out. Constructs containing right amino acid sequence in the correct conformation best bind the antibody.

THE PRINCIPLES OF CLIPS TECHNOLOGY

The majority of biomolecules of therapeutic interest recognize conformational or discontinuous epitopes on their cognate target. To mimic this situation in peptides derived from the target, CLIPS technology structurally fixes peptides into defined 3D structures. The CLIPS reaction takes place between bromo groups of the CLIPS scaffold and thiol sidechains of cysteines introduced into peptide constructs. The reaction is ultra-fast, very specific and is undertaken under mild conditions. Using this elegant chemistry, native protein sequences are transformed into CLIPS constructs with a range of structures (Figure 3). CLIPS technology is now routinely used to shape peptide libraries into single, double or triple looped structures as well as sheet- and helix-like folds, which allows mimicking of conformational and discontinuous binding sites.

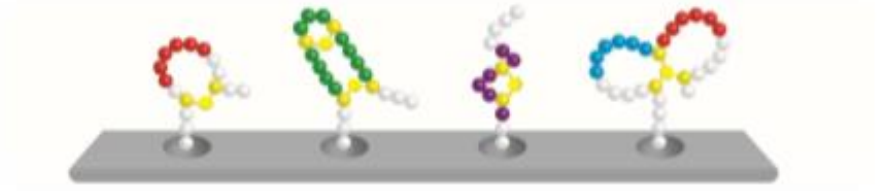


Figure 3. Examples of conformational peptide mimics. Using CLIPS technology, peptides derived from native proteins are transformed into CLIPS constructs with a range of structures. From left to right: single mP2 loop, stabilized beta sheet, alpha helix, T3 double loop.

CLIPS LIBRARY FOR PROFILING CONFORMATIONAL EPITOPES

Conformational epitopes can be mimicked using CLIPS chemistry. Simple secondary structure mimics can be stabilized by application of CLIPS scaffolds that allow selecting thermodynamically favored peptide conformations. CLIPS peptide libraries can mimic secondary structure elements, such as loops, α-helices, and β-strands. A schematic of this approach is drawn in Figure 4, where a single loop element is mimicked using mP2 CLIPS chemistry.

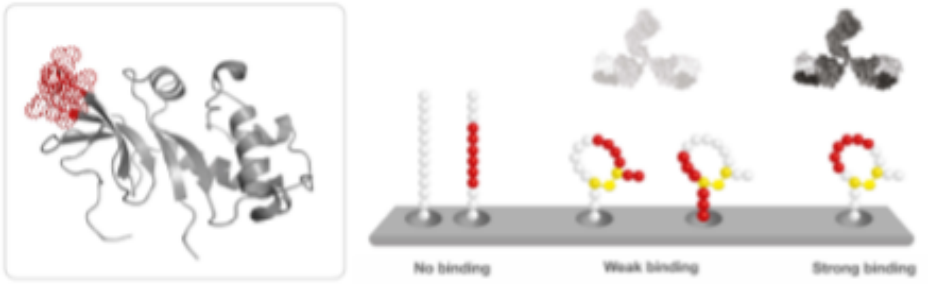


Figure 4. Constructing a conformational peptide library using CLIPS. The target protein contains α-helixes, ß-sheets separated by loops is converted into different conformational libraries using a CLIPS scaffold. In this example only single loop mimics are shown. Peptides are synthesized on a proprietary minicard and chemically converted into spatially defined CLIPS constructs (right). Binding of antibodies is quantified using an automated ELISA-type read-out. Constructs containing the right amino acid sequence in the correct conformation best bind the antibody.

COMBINATORIAL CLIPS LIBRARY FOR PROFILING DISCONTINUOUS EPITOPES

Combinatorial screening starts with the conversion of the target protein into a library of up to 10,000 overlapping peptide constructs representing both parts of the discontinuous epitope in the correct conformation bind the antibody with increased affinity. Constructs presenting the incomplete epitope may bind the antibody with lower affinity, whereas constructs not containing the epitope do not bind at all.

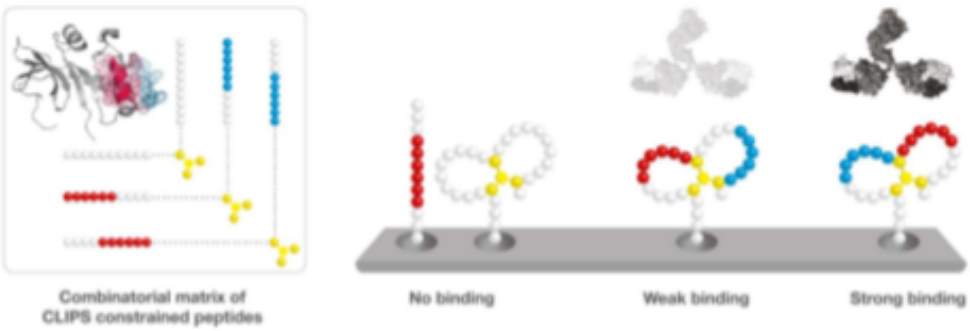


Figure 5. Constructing a discontinuous peptide library using CLIPS. The target protein containing a discontinuous epitope (left cartoon) is converted into a library of linear peptides as well as CLIPS constructs via a combinatorial matrix design. Constructs representing both parts of the discontinuous epitope in the correct orientation best bind the antibody.

EXPERIMENTAL METHODS

SYNTHESIS OF PEPTIDES

To reconstruct epitopes of the target molecule a library of peptide-based epitope mimics was synthesized using Fmoc-based solid-phase peptide synthesis. An amino functionalized polypropylene support was obtained by grafting with a proprietary hydrophilic polymer formulation, followed by reaction with t-butyloxycarbonyl-hexamethylenediamine (BocHMDA) using dicyclohexylcarbodiimide (DCC) with *N*-hydroxybenzotriazole (HOBt) and subsequent cleavage of the Boc-groups using trifluoroacetic acid (TFA). Standard Fmoc-peptide synthesis was used to synthesize peptides on the amino-functionalized solid support by custom modified JANUS liquid handling stations (Perkin Elmer).

Synthesis of structural mimics was done using Pepscan's proprietary Chemically Linked Peptides on Scaffolds (CLIPS) technology. CLIPS technology allows to structure peptides into single loops, double-loops, triple loops, sheet-like folds, helix-like folds and combinations thereof. CLIPS templates are coupled to cysteine residues. The side-chains of multiple cysteines in the peptides are coupled to one or two CLIPS templates. For example, a 0.5 mM solution of the P2 CLIPS (2,6-bis(bromomethyl)pyridine) is dissolved in ammonium bicarbonate (20 mM, pH 7.8)/acetonitrile (1:3(v/v)). This solution is added onto the peptide arrays. The CLIPS template will bind to side-chains of two cysteines as present in the solid-phase bound peptides of the peptide-arrays (455 wells plate with 3 µl wells). The peptide arrays are gently shaken in the solution for 30 to 60 minutes while completely covered in solution. Finally, the peptide arrays are washed extensively with excess of H_2O and sonicated in disrupt-buffer containing 1 % SDS/0.1 % 2,2'-(Ethylenedioxy)diethanethiol in PBS (pH 7.2) at 70°C for 30 minutes, followed by sonication in H_2O for another 45 minutes. The T3 CLIPS carrying peptides were made in a similar way but now with three cysteines.

ELISA SCREENING

The binding of antibody to each of the synthesized peptides was tested in a pepscan-based ELISA. The peptide arrays were incubated with primary antibody solution (overnight at 4°C). After washing, the peptide arrays were incubated with a 1/1000 dilution of an appropriate antibody peroxidase conjugate (SBA; Table 4) for one hour at 25°C. After washing, the peroxidase substrate 2,2'-azino-di-3-ethylbenzthiazoline sulfonate (ABTS) and 20 µl/ml of 3 percent H_2O_2 were added. After one hour, the color development was measured. The color development was quantified with a charge coupled device (CCD) - camera and an image processing system.

Table 4. Details of the antibodies

Name	Supplier	Cat. No
goat anti-human HRP conjugate	Southern Biotech	2010-05
rabbit anti-mouse IgG(H+L) HRP conjugate	Southern Biotech	6175-05

DATA PROCESSING

The values obtained from the CCD camera range from 0 to 3000 mAU, similar to a standard 96-well plate ELISA-reader. The results are quantified and stored into the Peplab database. Occasionally a well contains an air-bubble resulting in a false-positive value, the cards are manually inspected and any values caused by an air-bubble are scored as 0.

SYNTHESIS QUALITY CONTROL

To verify the quality of the synthesized peptides, a separate set of positive and negative control peptides was synthesized in parallel. These were screened with commercial antibodies 3C9 and 57.9 (*ref.* Posthumus *et al.* (1990) *J. Virol.* **64**:3304–3309)

LITERATURE REFERENCES

Timmerman *et al.* (2007). Functional reconstruction and synthetic mimicry of a conformational epitope using CLIPS™ technology. *J. Mol. Recognit.* **20**:283–299

Langedijk *et al.* (2011). Helical peptide arrays for lead identification and interaction site mapping, *Analytical Biochemistry* **417**: 149–155

DATA ANALYSIS AND INTERPRETATION

BOXPLOT ANALYSIS

The box-and-whisker plot is an exploratory tool allowing displaying batches of data (Tukey 1970, 1977). Box plots give a first impression on data distribution of a given data set. Each box plot displays data set extremes (max & min), upper and lower quartiles, median and outliers (Figure 6).

The main application of this boxplot is in finding the baseline levels and estimating signal over noise to allow data quality assessment.

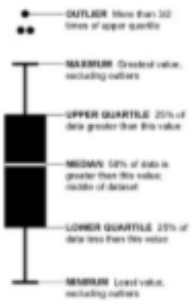


Figure 6. Box plot schematic. Source: https://flowingdata.com

LINEAR INTENSITY PROFILES

Throughout the report linear intensity profiles (e.g. top panel in Figure 7) depict a graphical representation of the intensity profile recorded for a given sample. Often overlays of intensity profiles recorded on different peptide sets are plotted in one graph to allow direct comparison between responses of a given antibody to epitope mimics of different types.

HEAT MAP ANALYSIS

A heat map is a graphical representation of data where the values taken by a variable in a two-dimensional map are represented as colors.

Heatmaps can be used to visualize data:

1). Representation or comparison of ELISA intensities recorded for sample(s) on set(s) of peptides;

2). Visualization of data recorded on a discontinuous peptide library for a given sample.

1. HEATMAP ANALYSIS OF SIMPLE INTENSITY PROFILES

The magnitude of a response is color-coded and the key is included in the plot. Figure 7 below illustrates a response from a polyclonal sample recorded on a library of overlapping peptides. The response is displayed in two complementary ways: as histogram (Figure 7, top panel) and a heatmap (Figure 7, bottom panel). While for a single sample the histogram is easily grasped, a comparison of responses from numerous samples is easier, when analyzed in a heatmap, where each sample is plotted as a row (or column) and each peptide as a column (or row).

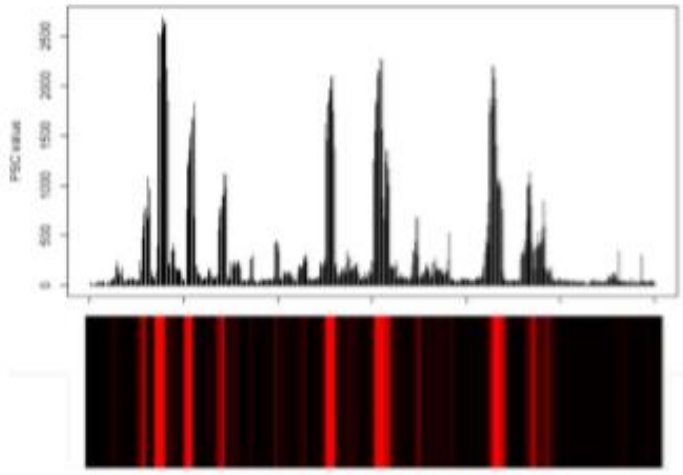


Figure 7. Example of intensity profile and corresponding heatmap representation of a response from a polyclonal serum screened on a library of overlapping linear peptides.

DESIGN OF PEPTIDES

Different sets of peptides were synthesized (see also Methods section) according to the following designs. Note that actual order of peptides on mini-cards in some was randomized. Note that not all experimental datasets yielded optimal results. These sets are not mentioned in detail in the results section, but are provided in the raw data file.

SET 1

Mimic Type	linear
Label	LIN
Description	Linear peptides of length 15 derived from the target sequence on page 8 with an offset of one residue. Cys are protected by acetamidomethyl (Acm; denoted "2").
Sequences (first 10)	2GPVY2FTPSPVVVG TALN2NDSLNTGFLA NRTALN2NDSLNTGF SWGENETDVLLLNNT DTLT2PTD2FRKHPE STGFTKT2GGPP2NI GVPTYSWGENETDVL KHPEATYTR2GSGPW KFAQGWGPITHTEPP WFG2TWWNSTGFTKT

SET 2

Mimic Type	Loop, mP2 CLIPS
Label	LOOP
Description	Constrained peptides of length 17. On positions 2 – 16 are 15-mer sequences derived from the target sequence of HCV-E1E2 glycoprotein. To introduce structural constraints Cys were inserted on positions 1 and 17, which then were constrained by mP2 CLIPS. Native Cys are protected by Acm (denoted "2").
Sequences (first 10)	CETHTVGGSASRAAHRC CTHTVGGSASRAAHRVC CHTVGGSASRAAHRVTC CTVGGSASRAAHRVTTC CVGGSASRAAHRVTTFC CGGSASRAAHRVTTFIC CGSASRAAHRVTTFITC CSASRAAHRVTTFITRC CASRAAHRVTTFITRGC CSRAAHRVTTFITRGPC

SET 3

Mimic Type	helical mimic, mP2 CLIPS
Label	HEL
Description	Structured peptides of length 23 derived from the target sequence on page 8 with an offset of one residue. On positions 1 and 5 are Cys residues which are joined by mP2 CLIPS. Native Cys are protected by Acm (denoted "2").
Sequences (first 10)	CETHCVGGSASRAAHRVTTFIT CTHTCGGSASRAAHRVTTFITR CHTVCGSASRAAHRVTTFITRG CTVGCSASRAAHRVTTFITRGP CVGGCASRAAHRVTTFITRGPS CGGSCSRAAHRVTTFITRGPSQ CGSACRAAHRVTTFITRGPSQN CSASCAAHRVTTFITRGPSQNI CASRCAHRVTTFITRGPSQNIQ CSRACHRVTTFITRGPSQNIQL

SET 4

Mimic Type	β-turn mimic, mP2 CLIPS
Label	BET
Description	Structured peptides of length 22. On positions 2- 21 are 20-mer sequences derived from the target sequence of HCV-E1E2 glycoprotein with an offset of one residue. "PG" residues supplant the residues present on positions 10 and 11 in order to induce β-turn formation. On positions 1 and 22 are Cys residues which are joined by mP2 CLIPS. Native Cys are protected by Acm (denoted "2").
Sequences (first 10)	CHRVTTFITRPGSQNIQLINTC CTYTR2GSGPPGTPR2MVHYPC CQNIQLINTNPGWHINRTALNC CNDTLT2PTDPGRKHPEATYTC CNETDVLLLNPGRPPRGNWFGC CETDVLLLNNPGPPRGNWFG2C CP2GIVPAAQPGGPVY2FTPSC C2RPIDKFAQPGGPITHTEPPC CDSLNTGFLAPGFYTHRFNASC CHYAPRP2GIPGAAQV2GPVYC

SCREENING DETAILS

Antibody binding depends on a combination of factors, including concentration of the antibody and the amounts and nature of competing proteins in the ELISA buffer. Also, the pre-coat conditions (the specific treatment of the peptide arrays prior to incubation with the experimental sample) affect binding. These details are summed up in Table 5. The Pepscan Buffer is (PBS-Tween-based) and Preconditioning (SQ), the numbers indicate the relative amount of competing protein (a combination of horse serum and ovalbumin).

Table 5. Screening conditions

Label	Dilution	Sample buffer	Pre-conditioning
G	5ug/ml	PBS+Tween	PBS+Tween
J	5ug/ml	PBS+Tween	PBS+Tween
M	5ug/ml	PBS+Tween	PBS+Tween
AP33	10ug/ml	PBS+Tween	PBS+Tween

SCREENING RESULTS

PRIMARY EXPERIMENTAL RESULTS AND SIGNAL TO NOISE RATIO DETERMINATION

The raw ELISA results of the screening are provided in the accompanying Excel data file. A graphical overview of the complete dataset is given in Figure 8. Here a box plot depicts each dataset and indicates the average ELISA signal, the distribution, and the outliers within each dataset. Depending on experiment conditions (amount of antibody, blocking strength, etc.) different distributions of ELISA data are obtained.

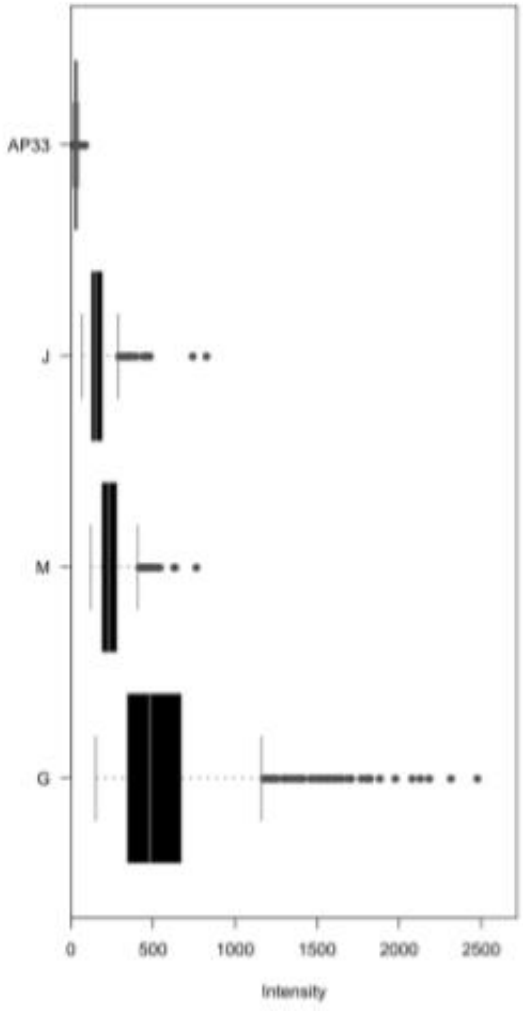


*Figure 8. Box plot graphs of raw data of antibody screening. The bottom and top of the boxes are the 25th and 75th percentile of the data. The band near the middle of the box is the 50th percentile (the median). The whiskers are at 1.5 the inter-quantile range, an indication of statistical outliers within the dataset (ref McGill et al. (1978) The American Statistician, **32**: 12-16).*

Results and Discussion

For each antibody sample the screenings were optimized for blocking conditions and sample concentration. For all samples tested, blocking prevented good signals, thus final screens were done in the absence of blocking reagent. For epitope validation, we look for consistency and redundancy between screens, and for a binding intensity higher than 3 times the median value, unless otherwise indicated.

SAMPLE AP33

For sample AP33 no clear peaks were observed with linear or conformational epitope mimics. Therefore, no epitope could be elucidated. For clarity, a graphical depiction of the results is shown in Figure 9.

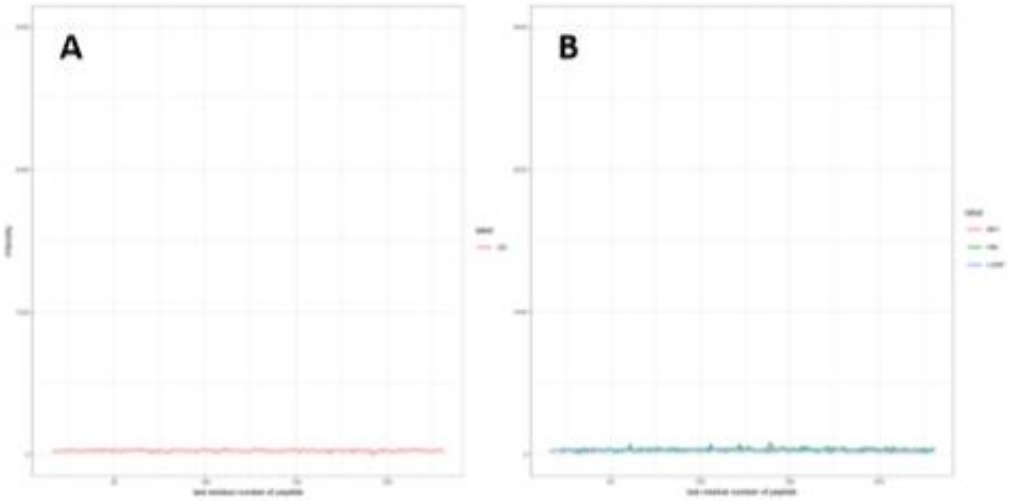


Figure 9. Overlay of intensity profiles of linear and conformational epitope mimics recorded for sample AP33. Binding intensities for linear peptide mimics (A) and conformational peptide mimics (B) are plotted per peptide. Intensities recorded with different sequence variants are color coded. Signal intensities are plotted on the y axis and positions of the last residues of a peptide with respect to the target sequence is on the x axis. LIN, linear; BET, beta-turn; HEL, helical; LOOP, looped mimic.

SAMPLE J

For sample J, multiple binding peaks were observed with linear and conformational mapping, suggesting the presence of multiple putative epitopes. The top binding peaks are displayed in Table 6. Residues lying just outside these peaks may still be the part of the epitope. While some binding peaks were observed in the linear peptide array (Fig. 10A), these peaks were not consistent between repeated screenings. Several peaks were observed in the conformational peptide array (Fig. 10B); regions revolving around sequences $_{107}$YAPRPCGIVPAAQVCGPVYCF$_{127}$ and $_{211}$TYTRCGSGPWLTPRCMVHYPY$_{231}$ and $_{149}$ENETDVLLLPGTRPPRGNWF$_{168}$ showed binding in the helical mimics and/or beta-turn mimics. In the linear peptide array, small peaks were observed at this position as well. In the conformational array, a peak at sequence $_{5}$TVGGSASRAPGRVTTFITRG$_{24}$ is observed and possibly part of a putative epitope. Putative epitopes are depicted in the structure in Figure 11.

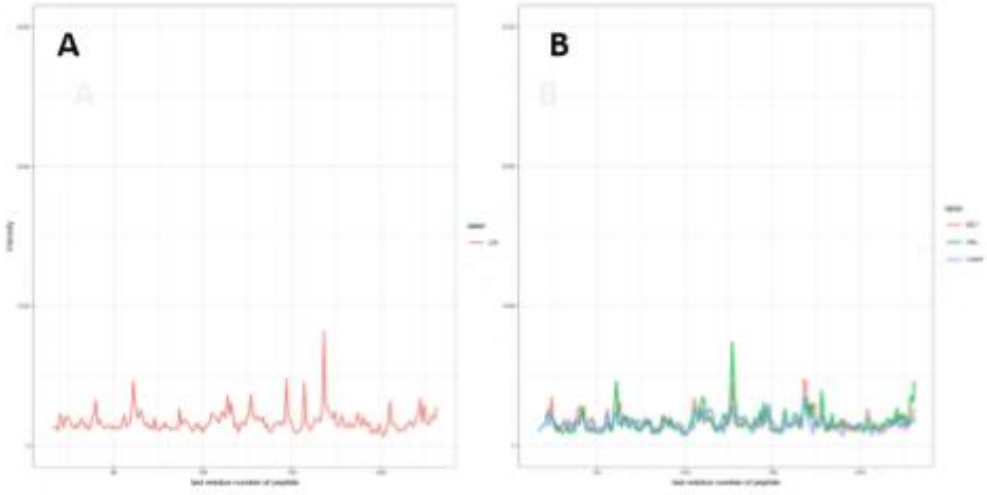


Figure 10. Overlay of intensity profiles of linear and conformational epitope mimics recorded for sample J. Binding intensities for linear peptide mimics (A) and conformational peptide mimics (B) are plotted per peptide. Intensities recorded with different sequence variants are color coded. Signal intensities are plotted on the y axis and positions of the last residues of a peptide with respect to the target sequence is on the x axis. LIN, linear; BET, beta-turn; HEL, helical; LOOP, looped mimic.

Figure 11. Depiction of putative epitopes for sample J on surface and ribbon representation of 4mwf.pdb.

SAMPLE M

For sample M, multiple binding peaks were observed with linear and conformational mapping, suggesting the presence of multiple putative epitopes. The top binding peaks are displayed in Table 6. Residues lying just outside this peak may still contain residues involved in the epitope. In the linear peptide array, we observed several binding events (Fig. 12A). However, due to a low signal-to-noise ratio they were not considered for putative epitope definition. Best binding was observed within $_{154}$VLLLNNTRPPRGNWFGC$_{170}$. This region is present in the conformational array as well (Fig. 12B). The main peak for the conformational arrays lies in the helical mimic and revolves around sequence $_{40}$NRTALNCNDSLNTGFLAALFY$_{60}$. Although not meeting all our criteria for epitope consideration, the beta-turn mimic $_{5}$TVGGSASRPGHRVTTFITRG$_{24}$ is consistently present between different screens, thus may contain residues part of an epitope. Putative epitopes are depicted in the structure in Figure 13.

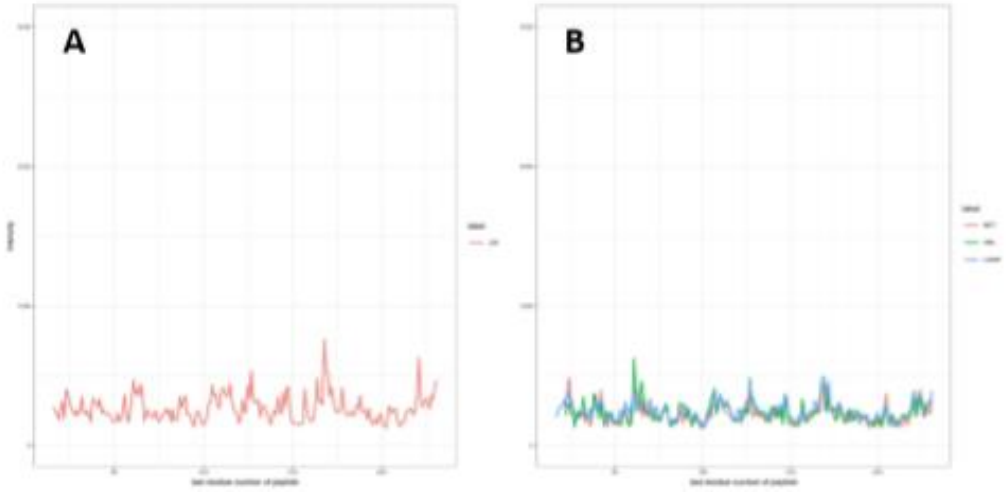


Figure 12. Overlay of intensity profiles of linear and conformational epitope mimics recorded for sample M. Binding intensities for linear peptide mimics (A) and conformational peptide mimics (B) are plotted per peptide. Intensities recorded with different sequence variants are color coded. Signal intensities are plotted on the y axis and positions of the last residues of a peptide with respect to the target sequence is on the x axis. LIN, linear; BET, beta-turn; HEL, helical; LOOP, looped mimic.

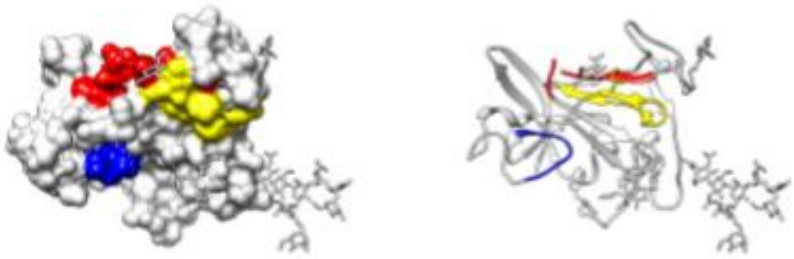


Figure 13. Depiction of putative epitopes for sample M on surface and ribbon representation of 4mwf.pdb.

SAMPLE G

For sample G, multiple binding peaks were observed with linear and conformational mapping, suggesting there are multiple putative epitopes present. A graphical depiction of the results is shown in Figure 14. Due to presence of multiple peaks throughout the arrays, identification of epitopes using our criteria based on the median intensity value was difficult. Therefore, we looked at peak definition to assign the main peaks. Main peaks are listed in table 6.

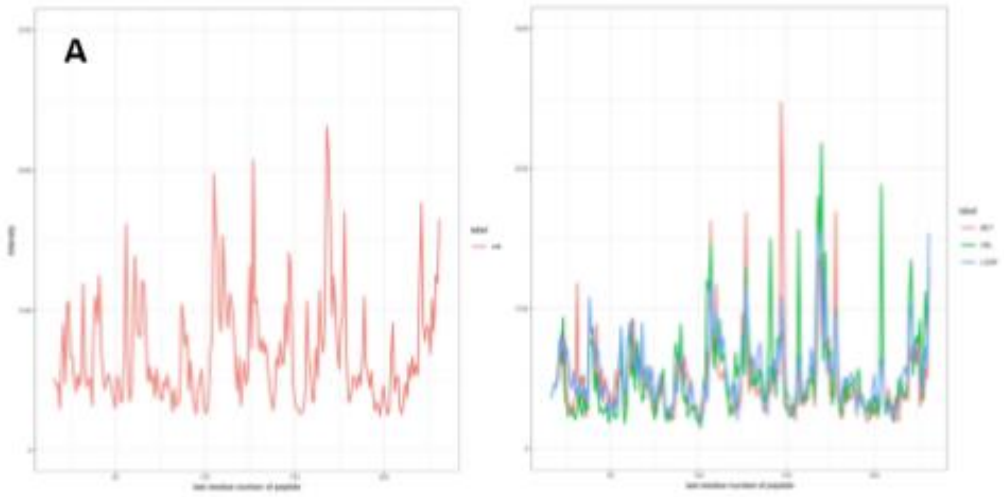


Figure 14. Overlay of intensity profiles of linear and conformational epitope mimics recorded for sample G. Binding intensities for linear peptide mimics (A) and conformational peptide mimics (B) are plotted per peptide. Intensities recorded with different sequence variants are color coded. Signal intensities are plotted on the y axis and positions of the last residues of a peptide with respect to the target sequence is on the x axis. LIN, linear; BET, beta-turn; HEL, helical; LOOP, looped mimic.

CONCLUSIONS

The screening of polyclonal antibodies was performed on linear and conformational peptide arrays. For three samples in this study, binding events were observed. Sample AP33 did not show sufficient binding to confidently identify a putative epitope. Clearly, sample G showed greatest promise as relatively good binding was observed compared to the other samples. For direct comparison of the binding events, we plotted all samples per peptide mimic in Figures 15-18. The main binding peaks are listed in table 6. Interestingly, some overlap in binding was observed between samples.

Table 6. List of epitope candidates found in this study. Binding events of at least 3 times the median value are listed. For sample G, the main binding peaks were listed. The peptide groups are ordered by intensity of the signal.

Sample		Main Epitope candidates
AP33		-
J	LIN	VLLLNNTRPPRGNWF VVGTTDRFGVPTYSW CNDSLNTGFLAALFY PTYSWGENETDVLLL
	BET	CENETDVLLLPGTRPPRGNWFC CNETDVLLLNPGRPPRGNWFGC
	HEL	CYAPCPCGIVPAAQVCGPVYCF CTYTCCGSGPWLTPRCMVHYPY CNRTCLNCNDSLNTGFLAALFY
M	LIN	VLLLNNTRPPRGNWF LLLNNTRPPRGNWFG LLNNTRPPRGNWFGC HPEATYTRCGSGPWL GIVPAAQVCGPVYCF
G	LIN	VLLLNNTRPPRGNWF LLLNNTRPPRGNWFG LLNNTRPPRGNWFGC RGNWFGCTWMNSTGF GIVPAAQVCGPVYCF PCGIVPAAQVCGPVY THTEPPSSDQKPYCW HTEPPSSDQKPYCWH TEPPSSDQKPYCWHY PSSDQKPYCWHYAPR HPEATYTRCGSGPWL SGPWLTPRCMVHYPY CGSGPWLTPRCMVHY RTALNCNDSLNTGFL CNDSLNTGFLAALFY VVGTTDRFGVPTYSW VGTTDRFGVPTYSWG

	QNIQLINTNGSWHIN
BET	CTPSPVVVGIPGRFGVPTYSWC
	CPSPVVVGTIPGFGVPTYSWGC
	CNTRPPRGNWPGCTWMNSTGFC
	CAPRPCGIVPPGQVCGPVYCFC
	CGPITHTEPPPGDQKPYCWHYC
	CNETDVLLLNPGRPPRGNWFGC
	CENETDVLLLPGTRPPRGNWFC
	CYTRCGSGPWPGPRCMVHYPYC
HEL	CNETCVLLLNNTRPPRGNWFGC
	CGENCTDVLLLNNTRPPRGNWF
	CENECDVLLLNNTRPPRGNWFG
	CTDVCLLNNTRPPRGNWFGCTW
	CWGECETDVLLLNNTRPPRGNW
	CGPPCNIGGVGNDTLTCPTDCF
	CTDRCGVPTYSWGENETDVLLL
	CCGPCYCFTPSPVVVGTTDRFG
	CWGPCTHTEPPSSDQKPYCWHY
	CQGWCPITHTEPPSSDQKPYCW
	CGWGCITHTEPPSSDQKPYCWH
	CHTECPSSDQKPYCWHYAPRPC
	CTDCCRKHPEATYTRCGSGPWL
	CEATCTRCGSGPWLTPRCMVHY
	CPTDCFRKHPEATYTRCGSGPW
	CYAPCPCGIVPAAQVCGPVYCF
	CTYTCCGSGPWLTPRCMVHYPY
LOOP	CVLLLNNTRPPRGNWFC
	CLLLNNTRPPRGNWFGC
	CNNTRPPRGNWFGCTWC
	CLLNNTRPPRGNWFGCC
	CDVLLLNNTRPPRGNWC
	CSGPWLTPRCMVHYPYC
	CTEPPSSDQKPYCWHYC
	CPSSDQKPYCWHYAPRC
	CEPPSSDQKPYCWHYAC
	CGIVPAAQVCGPVYCFC

CIVPAAQVCGPVYCFTC CCGIVPAAQVCGPVYCC CVVGTTDRFGVPTYSWC CGPSQNIQLINTNGSWC CSQNIQLINTNGSWHIC CHPEATYTRCGSGPWLC CKHPEATYTRCGSGPWC

Figure 15. Comparison of intensity profiles of linear epitope mimics for all samples. *Binding intensities for linear peptide mimics are plotted per peptide and plots for all samples were stacked for comparison between samples.*

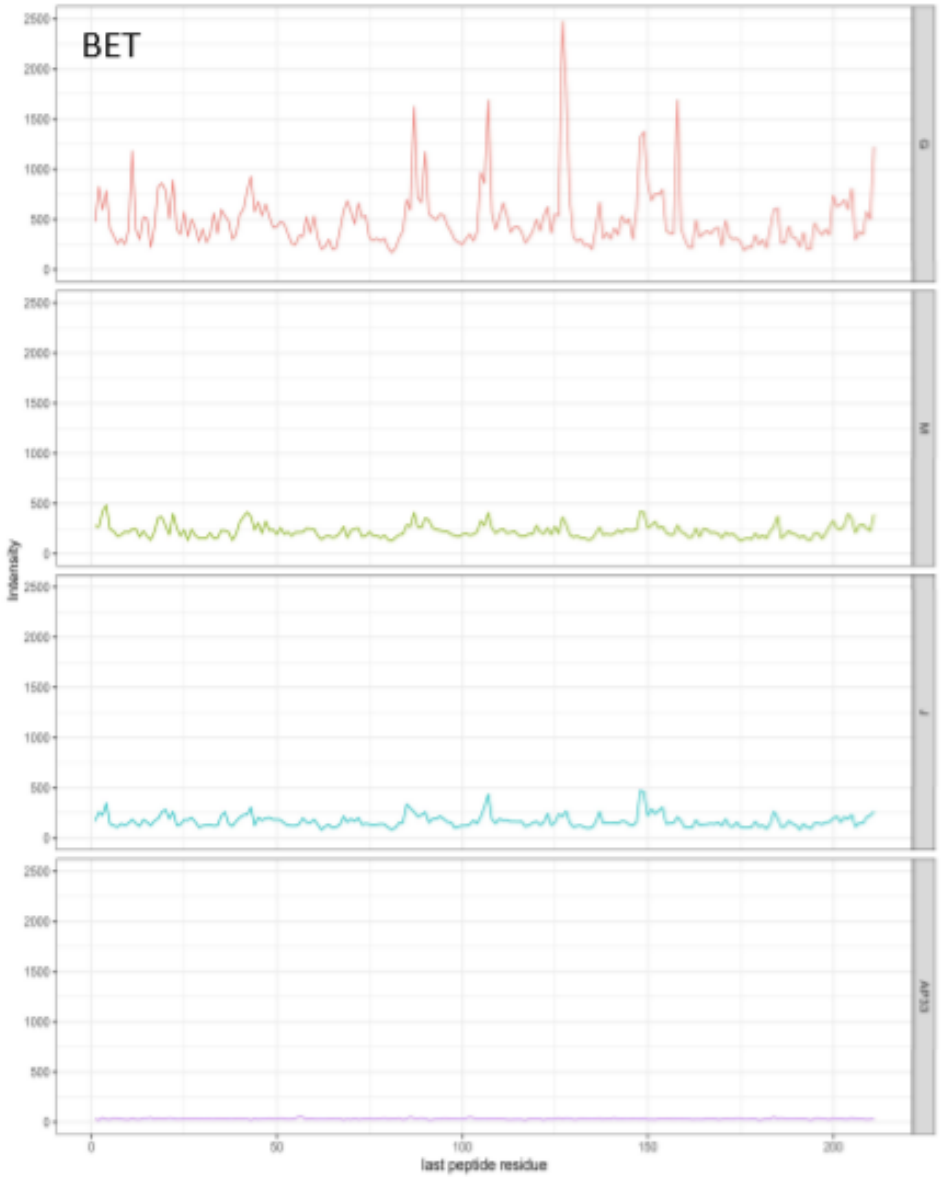


Figure 16. Comparison of intensity profiles of beta-turn mimics for all samples. *Binding intensities for beta-turn peptide mimics are plotted per peptide and plots for all samples were stacked for comparison between samples.*

Figure 17. Comparison of intensity profiles of helical mimics for all samples. *Binding intensities for helical peptide mimics are plotted per peptide and plots for all samples were stacked for comparison between samples.*

Figure 18. Comparison of intensity profiles of looped mimics for all samples. *Binding intensities for looped peptide mimics are plotted per peptide and plots for all samples were stacked for comparison between samples.*

APPENDIX III – ABBREVIATIONS USED

3D	Three Dimensional
aa	Amino Acid
AASLD	American Association for the Study of Liver Disease
Abs	Antibodies
AKA	Also Known As
bnAbs	Broadly-Neutralising Antibodies
CARD	Caspase Activation and Recruitment Domain
CARDIF	CARD Adaptor Inducing IFN-β
CD-81	Cluster of Differentiation - 1
cDNA	Complementary DNA
CDC	Centre for Disease Control (USA)
CTL	Cytotoxic T Lymphocyte
DAA	Direct-Acting Antiviral
DC	Dendritic Cells
DC-SIGN	Dendritic Cell SIGN
DNA	Deoxyribonucleic Acid
EASL	European Association for the Study of the Liver
ECDC	European Centre for Disease Control
EGF	Epithelial Growth Factor
ER	Endoplasmic Reticulum
FDA	Food and Drug Administration (USA)
GFP	Green Fluorescent Protein
GLE	Glecaprevir
GRFT	Griffithsin Derived from Red Algae
GT	Genotype
HAV	Hepatitis A Virus
HBV	Hepatitis B Virus
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C Virus
HCVAg	HCV Core Antigen
HDL	High Density Lipoprotein
HGV	Hepatitis G Virus
HIV	Human Immunodeficiency Virus

HSC	Hepatic Stellate Cells
HSE	Health Service Executive (Ireland)
HVR	Hypervariable Region
IAV	Influenza A Virus
IDU	Intravenous Drug User
IFN	Interferon
IRES	Internal Ribosome Entry Site
IRF	Interferon Regulatory Factor
JAK	Janus Kinase
KC	Kupffer Cells
LD	Lipid Droplet
LDL	Low Density Lipoprotein
LDL-R	Low Density Lipoprotein Receptor
LEL	Large Extracellular Loop
LSEC	Liver Endothelial Sinusoidal Cells
L-SIGN	Liver/Lymph Node SIGN
LVP	Lipoviral Particle
LXRα	Liver X Receptor Alpha
mAb	Monoclonal Antibody
MAVS	Mitochondrial Antiviral Signalling Protein
MDA5	Melanoma Differentiation Antigen – 5
mDC	Myeloid Dendritic Cell
MERS	Middle East Respiratory Syndrome
MHC	Major Histocompatibility Complex
MLV	Murine Leukaemia Virus
mRNA	Messenger RNA
miRNA	Micro RNA
MSM	Men-Who-Have-Sex-With-Men
nAb	Neutralising Antibody
NANBH	Non-A, Non-B Hepatitis
NIH	National Institutes of Health
NK	Natural Killer Cells
NK-ADCC	NK-Mediated Antibody-Dependent Cellular Cytotoxicity
NKT	Natural Killer T Cells
NLR	NOD-Like Receptor
NOD	Nucleotide-Binding Oligomerisation Domain

NPC	Non-Parenchymal Cells
NS	Non-Structural
NSP	Needle and Syringe (Exchange) Programme
OBV	Ombitasvir
ORF	Open Reading Frame
PAMP	Pathogen-Associated Molecular Patterns
PBMC	Peripheral Blood Mononuclear Cell
PCR	Polymerase Chain Reaction
pDC	Plasmacytoid Dendritic Cell
peg-IFN	Pegylated Interferon
PI	Protease Inhibitor
PIB	Pibrentasvir
POC	Point-of-Care
PTV	Paritaprevir
PRR	Pathogen Recognition Receptor
PWID	People Who Inject Drugs
QALY	Quality-Adjusted Life-Year
r	Ritonavir
RAV	Resistance-Associated Variant
RBV	Ribavirin
RdRp	RNA-Dependent RNA Polymerase
Rh	Rhesus
RIG-I	Retinoic Acid Inducible Gene – I
RLR	RIG-I-Like Receptor
RNA	Ribonucleic Acid
SARS	Severe Acute Respiratory Syndrome
SIGN	Specific Intercellular Adhesion Molecule-3-Grabbing Non-Integrin
SMV	Simeprevir
SOC	Standard-of-Care
SOCS	Suppressor of Cytokine Signalling
SOF	Sofosbuvir
SR-B1	Scavenger Receptor Class B – Member 1
ssRNA	Single-Stranded RNA
STAT	Signal Transducer and Activator of Transcription
STI	Sexually-Transmitted Infection
STING	Stimulator of Interferon Genes

SVR	Sustained Virologic Response
T_C	Cytotoxic T Cells
TC-PTP	T Cell Protein Tyrosine Phosphate
TLR	Toll-Like Receptor
TND	Target Not Detected
TNF	Tumour Necrosis Factor
TRAF	TNF Receptor Associated Factor
Tregs	Regulatory T Cells
TRIF	Toll-Interleukin-1 Receptor Domain-Containing Adaptor Inducing IFN
UDPS	Ultra-Deep Pyrosequencing
UTR	Untranslated Region
VLDL	Very Low Density Lipoprotein
vRNA	Viral RNA
WHO	World Health Organisation

APPENDIX IV – AMINO ACID CODE

Amino Acid	Three Letter Code	One Letter Code
Alanine	ala	A
Arginine	arg	R
Asparagine	asn	N
Aspartic Acid	asp	D
Asparagine/Aspartic Acid	asx	B
Cysteine	cys	C
Glutamic Acid	glu	E
Glutamine	gln	Q
Glutamine/Glutamic Acid	glx	Z
Glycine	gly	G
Histidine	his	H
Isoleucine	ile	I
Leucine	leu	L
Lysine	lys	K
Methionine	met	M
Phenylalanine	phe	F
Proline	pro	P
Serine	ser	S
Threonine	thr	T
Tryptophan	trp	W
Tyrosine	tyr	Y
Valine	val	Z

APPENDIX V – IUPAC NUCLEOTIDE BASE CODE

One Letter Code	Base
A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil
R	A or G
Y	C or T
S	G or C
W	A or T
K	G or T
M	A or C
B	C or G or T
D	A or G or T
H	A or C or T
V	A or C or G
N	Any Base
. or -	Gap