UNIVERSITÀ DEGLI STUDI
DI MILANO – BICOCCA

Facoltà di Scienze Matematiche, Fisiche e Naturali
Dipartimento di Biotecnologie e Bioscienze

Dottorato di ricerca in Biotecnologie Industriali,
XXII ciclo

Study of the hepatitis C virus NS3 helicase domain for application in a chemiluminescent immunoassay

Dott. Gabriele Rebuzzini

Anno Accademico 2008/2009

Dottorato in Biotecnologie Industriali, XXII ciclo

Dott. Gabriele Rebuzzini
Matricola: 033809

Tutor: Prof. Danilo Porro
Dott. Andrea Dal Corso

Il lavoro presentato in questa tesi è stato realizzato presso i laboratori Nerviano Research Center di DiaSorin S.p.A., sotto la supervisione del Dott. Andrea Dal Corso.

Università degli Studi di Milano-Bicocca
Piazza dell'Ateneo Nuovo 1, 20126, Milano

Dipartimento di Biotecnologie e Bioscienze
P.za della Scienza 2, 20126, Milano

INDEX

1. Introduction

1.1. Discovery of Hepatitis C Virus

Hepatitis C virus (HCV) infects approximately 170 million people worldwide (WHO, 2003[113]). The first demonstration that most cases of transfusion-associated hepatitis were caused by neither hepatitis A virus HAV) nor hepatitis B virus (HBV), the only two known human hepatitis viruses at the time, came in 1975 [26]. This new form of disease was called non-A non-B hepatitis and the presumed etiologic agent was called non-A non-B hepatitis virus. However, it was only after many years of attempting to isolate the agent responsible for this so-called post-transfusion, non-A, non-B hepatitis that, in 1989, with the aid of modern techniques of molecular cloning and phage display, a new RNA virus, termed hepatitis C virus (HCV), was isolated [15]. HCV causes a persistent infection in the majority of infected people and can lead to severe manifestations as cirrhosis of the liver and hepatocellular carcinoma[95]. For this reason, and the high prevalence of infection worldwide, HCV is rightly classified as a major human pathogen.

1.2. Classification

HCV, a positive sense, single-stranded RNA virus, has been categorised as a member of the Hepacivirus genus within the Flaviviridae by genome analogy with other members of this family [16]. This family also includes the flaviviruses such as dengue virus and Japanese encephalitis virus, the pestiviruses such as bovine viral diarrhoea virus and classic swine fever virus, and the recently discovered GBV-A and B viruses and hepatitis G virus. The HCV genome encodes a polyprotein of approximately 3011 amino acids [16], which is comparable in size to other members of the Flaviviridae such as the flavivirus yellow fever virus (~3460 aa) and the pestivirus bovine viral diarrhoea virus (~3960 aa). The structural proteins of both the flavi- and pestiviruses are located at the N-termini of their polyproteins, beginning with a small, basic nucleocapsid protein [18]. The N-terminus of the HCV polyprotein is also highly basic. Furthermore, HCV, flaviviral and pestiviral polyproteins all share similar hydrophobic characteristics. Alignment of the HCV genome with other genomes of the Flaviviridae shows regions of sequence homology as well as comparable genomic organisation. As an example the region spanning amino acids 1230-1500 contains many residues identical to a putative NTP- binding helicase encoded by human flaviviruses, animal pestiviruses and plant potyviruses. Also, upstream from this lies a region sharing residues conserved among the putative trypsin-like serine proteases thought, by comparative sequence analysis with trypsin-like molecules, to be encoded by flaviviruses and

pestiviruses [34]. HCV and the pestiviruses have, within their 5'UTR, an internal ribosome entry site (IRES), which directs cap-independent translation of the open reading frame [78]. HCV has also been shown to share with flaviviruses and pestiviruses a large, conserved stem-loop structure within the 5'UTR [7]. However, despite similarities between HCV and both flaviviruses and pestiviruses, significant differences also exist, and these differences led to the proposal of a third Flaviviridae genus, the hepaciviruses.

Investigators in Japan were the first to suggest that HCV was a genetically diverse virus, with subsequent studies identifying at least 6 major genetic groups. To classify these groups, all presently known HCV isolates have been placed into one of 6 clades containing all 11 HCV subtypes (fig. 1.1).

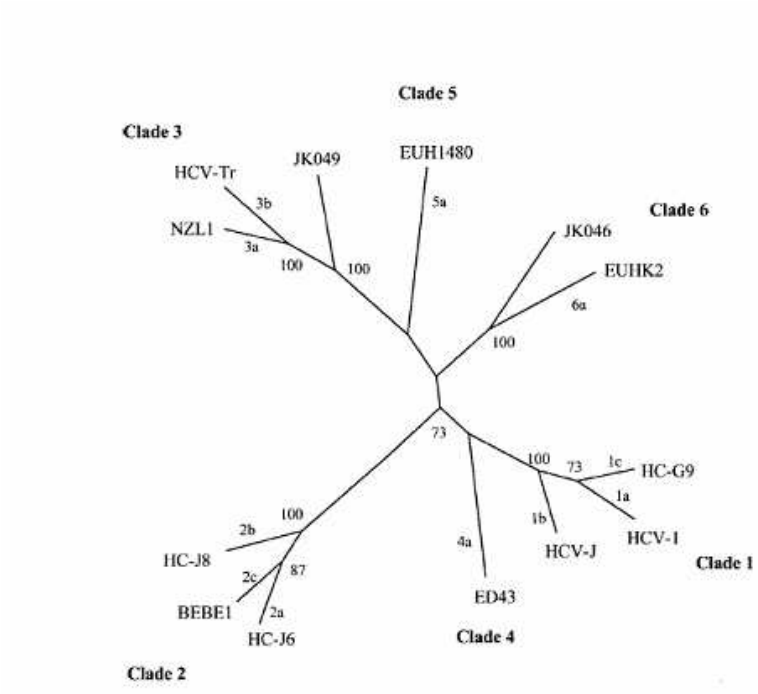


Figure 1.1: Cladogram of HCV genomes (taken from [85])

Furthermore, due to the high rate of mutations introduced during viral replication, a population of variant HCV species (termed quasispecies) can be isolated within a single patient [40]. The variability is distributed throughout the genome. However, the non-coding regions at either end of the genome (5'-UTR and 3'-UTR) are more conserved. The genes coding for the envelope E1 and E2 glycoproteins are the most variable. Aminoacid changes

may alter the antigenic properties of the proteins, thus allowing the virus to escape neutralizing antibodies [96].
Genotypes 1-3 have a worldwide distribution. Types 1a and 1b are the most common, accounting for about 60% of global infections. They predominate in Northern Europe and North America, and in Southern and Eastern Europe and Japan, respectively. Type 2 is less frequently represented than type 1. Type 3 is endemic in south-east Asia and is variably distributed in different countries. Genotype 4 is principally found in the Middle East, Egypt, and central Africa. Type 5 is almost exclusively found in South Africa, and genotypes 6-11 are distributed in Asia (WHO, 2003[113]).
Figure 1.2 reports a geographic distribution of the main six HCV genotypes (1-6) in 1999 [29].

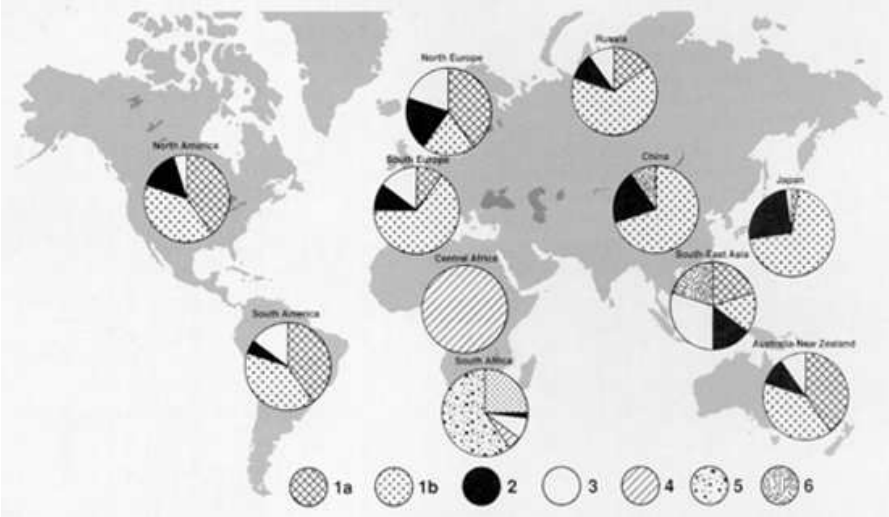


Figure 1.2: Geographic distribution of the main six HCV genotypes.

1.3. Epidemiology of HCV

Despite being endemic worldwide, there is high geographic variability in the distribution of HCV (figure 1.3). Africa and Asia have the highest reported prevalence rates while the lowest are found in industrialised countries such as North America, Australia and those in Northern and Western Europe (WHO, 2003[113]). The highest reported seroprevalence rate is in Egypt where approximately 22% of the population are HCV seropositive, thought to be due to contaminated glass syringes used in nationwide schistosomiasis treatment campaigns from 1960 to 1987 [30].

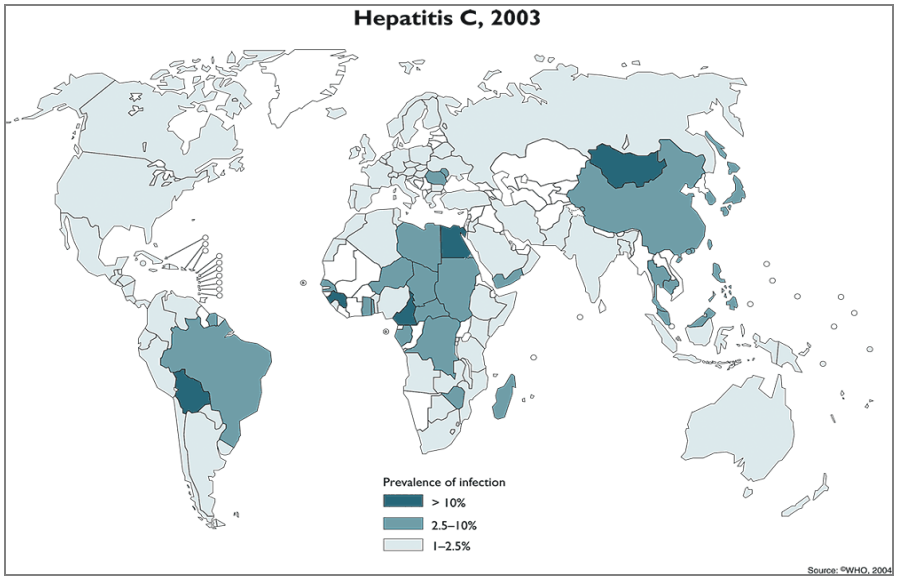


Figure 1.3: Prevalence rates of HCV infection worldwide

The most common factors responsible for HCV transmission worldwide are blood (transfusion from unscreened donors), intravenous drug abuse, unsafe therapeutic injections and other healthcare related procedures [98]. Within developed countries, the introduction of blood screening tests for HCV has effectively eradicated transmission by blood transfusion. Instead, injection drug use has been the predominant mode of transmission in recent times [63]. In the developing world however, unsafe therapeutic injections and blood transfusions are still major modes of transmission. HCV transmission via occupational, perinatal or sexual exposure is much less common. Perinatal transmission is estimated to occur in 2.7–8.4% of infants born to HCV infected mothers, with higher rates in those born to HIV/HCV co-infected mothers [117], though high rates of spontaneous clearance of neonatal infection during childhood has been observed [116]. Sexual mode of transmission has been reported, however it is far less efficient than for other sexually transmitted viruses [65].

1.4. Clinical manifestations of HCV infection

HCV infection is very often clinically silent with most acute infections being symptom-free and only a small number showing signs of jaundice [1]. Rapid, fulminant hepatitis associated with acute HCV infection has been reported in Japan, however this is not common elsewhere. This may reflect differences in genotype distribution as genotypes 2 and 3 are most prevalent in Japan while genotype 1 is prevalent in North America [118].

Infection becomes chronic in approximately 75% of patients, as confirmed by persistence of HCV RNA in serum[95]. Chronically infected patients may exhibit vague symptoms such as fatigue and joint aching, however it is more common for patients to be unaware until complications of chronic liver disease occur, often decades following infection. Almost all chronically infected patients develop histological features of chronic hepatitis such as portal inflammation, interface hepatitis and lobular injury[23]. Up to 20% of patients develop cirrhosis within the first two decades of HCV infection[115]. A wide range of factors can influence the development of cirrhosis, but it seems that being male, aged over 50 at time of infection and high alcohol intake increase susceptibility to cirrhosis[79]. Complications of chronic liver disease include liver failure and hepatocellular carcinoma (HCC). In patients with established cirrhosis, HCC may develop in up to 4% per year and up to 4% of patients infected with HCV may go on to develop HCC during their life[22] (fig. 1.4).

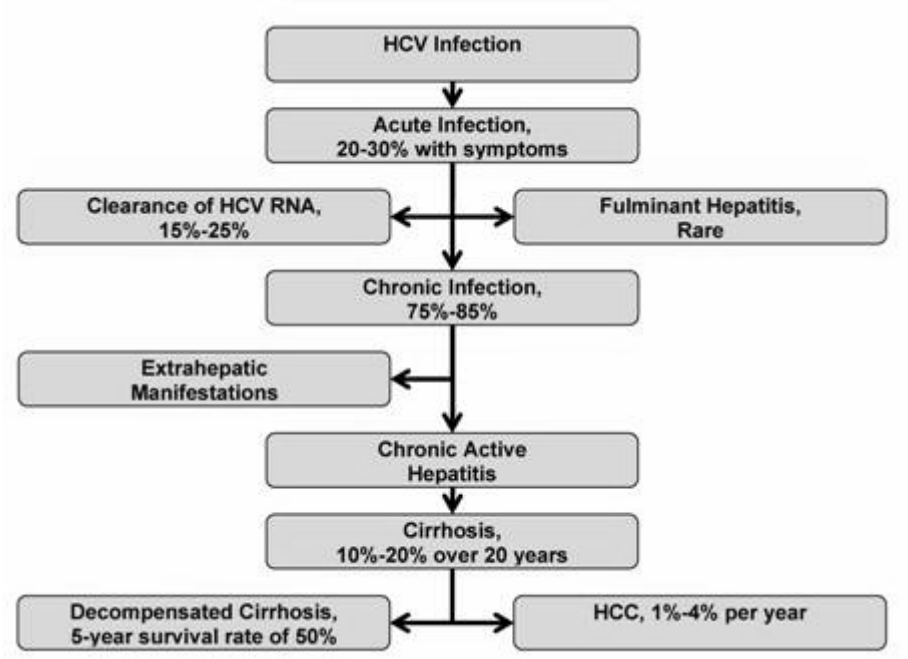


Figure 1.4: Various outcomes of HCV infection

Furthermore, there are many important extrahepatic manifestations of HCV infection. Most of them are associated with autoimmune or lymphoproliferative states and may be related to the possibility that HCV is able to replicate in lymphoid cells. Cryoglobulins can be found in 50% or more of patients with hepatitis C. Vasculitis, arthralgia and purpura are included as symptoms. Presently, it is well known that HCV is the chief cause of essential mixed cryoglobulinemia. Other diseases, including lichen

planus, Sjogren's syndrome, and porphyria cutanea tarda, have been linked to HCV infection [41]. However, a definite pathophysiological role of HCV in these diseases has been difficult to establish.

1.5. Immune response

Innate and acquired immune responses both play a role in HCV infection. The critical period, in terms of determining the outcome of infection appears to be the acute phase[36]. Natural killer cells may play a vital role as chimpanzees with asymptomatic HCV infection can eliminate the virus without any detectable HCV-specific T-cell response[105].

Within a few weeks of infection, the majority of patients show seroconversion, indicated by the presence of antibodies against both structural and non-structural proteins. The serology of HCV infection does not follow the classical pattern of IgM response observed in other viral infections because it may be absent, late or persistent after HCV infection and does not correlate with the histologic activity: positive anti-HCV IgM levels are found in 50-93% of patients with acute hepatitis C and in 50-70% of patients with chronic hepatitis C. Therefore, anti-HCV IgM cannot be used as a reliable marker of acute HCV infection[75]. IgG levels are used instead for monitoring of chronicized infections.

Clearance of infection is associated with a strong CD4+ and CD8+ T-cell response[103]. However, the cellular immune response may occur at the expense of a long-lasting inflammatory reaction resulting in liver cirrhosis and HCC.

1.6. Therapy

The therapy for chronic hepatitis C has evolved steadily since interferon-α was first approved for use in this disease in the 90's.

Interferon-α has potent antiviral activity, not by acting directly on the virus or its replication cycle, but instead by inducing interferon-stimulated genes which in turn promote a non-virus-specific antiviral state within the infected cell [91]. Recombinant forms of interferon-α have been produced, and several formulations (α-2a, α-2b, consensus interferon) are available as therapy for hepatitis C. However, interferon-α therapy alone has limited success, with only 16-20% of patients producing sustained responses after 12 months of treatment[24]. The addition of the antiviral agent Ribavirin more than doubled the response rate seen with interferon-α alone[68]. Today, pegylated interferon, in which polyethylene glycol (PEG) is covalently attached to recombinant interferon-α, in combination with Ribavirin, is the treatment of choice. Pegylation changes the uptake, distribution, and excretion of interferon, prolonging its half-life. PEG-interferon-α can be given once weekly and provides a constant level of interferon-α in the blood, whereas standard interferon-α must be given several times weekly and provides intermittent and fluctuating levels. In addition, PEG-interferon-α is more

active than standard interferon-α in inhibiting HCV and yields higher sustained response rates with similar side effects. Because of its ease of administration, better pharmacokinetics and higher efficacy, PEG-interferon-α has replaced standard interferon-α both as monotherapy and as combination therapy for hepatitis C[119].

The outcome of treatment can be grouped into one of three categories. Virological Response is defined as the loss of detectable HCV RNA during (Early Virological Response, EVR), and continuing for 6 months after the end of therapy (Sustained Virological Response, SVR). End-of-treatment Response and Relapse is defined as a transient response followed by relapse, while Non-Response, which occurs in approximately one-third of chronically infected patients, defines those who never become HCV RNA negative. When PEG-interferon-α plus ribavirin is administered for the standard duration (48 weeks), a sustained virological response is achieved in around 50% of patients infected with HCV genotype 1 and around 80% of patients infected with HCV genotype 2 or 3. Data now suggest that treatment duration can be shortened or lengthened depending on baseline viral load and/or early on-treatment viral kinetics, offering the prospect of individualizing therapy further to improve response or to prevent treatment from being unnecessarily extended (figure 1.5)[119].

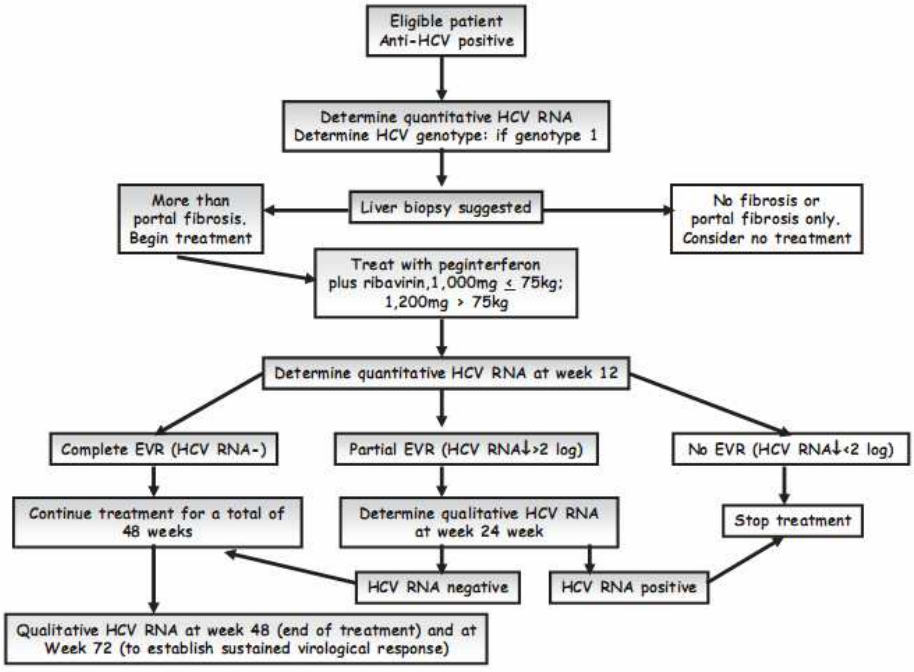


Figure 1.5: Algorithm for treatment of genotype 1 HCV infection (taken from [33])

Recently, research has been focused on new targets for HCV treatment. Agents have been discovered which inhibit specific processes in the virus life cycle including inhibitors of HCV enzymes as well as nucleic acid based molecules that interfere with the viral RNA. Small molecule inhibitors have been identified that block essential viral enzymes such as the NS3/4A protease and the NS5B polymerase[21][35]. Nucleic acid based antiviral agents have also been discovered recently including antisense oligonucleotides and small interfering RNAs (siRNAs)[83]. Potential has also been shown by novel immunomodulatory agents as candidates for treatment of HCV; a synthetic agonists of Toll-like receptor 7 has shown anti-HCV activity with very limited side effects[42].

1.7. Virus morphology

Electron microscopy studies on HCV positive plasma samples using monoclonal and polyclonal antibodies allowed visualisation of spherical particles of diameter 55-65 nm, with a low buoyant sucrose density between 1.08 and 1.16 g/ml [45]. HCV is believed to possess an envelope (derived from host cell membranes) as evidenced by its sensitivity to chloroform [27]. Non-enveloped HCV nucleocapsids, with higher buoyant density (~1.3 g/ml), have similar spherical morphology and can be found in significant quantities in serum[67]. The nucleocapsid is composed by the core protein, while E1 and E2 glycoproteins are integrated in the envelope and exposed on the outer side of the particle.

Upon synthesis, E1 and E2 proteins are targeted to the endoplasmic reticulum, where they fold and heterodimerize. In addition, they are retained in this compartment, where the virus particle is thought to acquire its envelope [73]. However, because HCV does not replicate efficiently in cell culture, it is currently impossible to study later events in the maturation of HCV envelope glycoproteins that might occur during budding and maturation of the particle. Non-structural proteins (NS) provide various functions necessary for viral entry, replication and processing (fig.1.6).

1.8. Genome structure

HCV has a single stranded, positive sense genome of approximately 9.6kb. The genome contains one long open reading frame (ORF) encoding a polyprotein of approximately 3011 amino acids, flanked by 5' and 3' non-coding regions[16]. The highly conserved 5'UTR is 341 nucleotides in length and contains extensive secondary structure. Similarities in structure between the HCV 5'UTR and that of picornaviruses led to the conclusion that HCV had an internal ribosome entry site (IRES), required for translation initiation [109]. The ORF encodes a single polyprotein which is processed co- and post-translationally by host and viral proteases to produce at least 10 viral proteins [11]. The structural proteins (core, E1, E2 and p7) are located within the amino-terminal third of the polyprotein, while the non-structural proteins

(NS2, NS3, NS4A, NS4B, NS5A and NS5B) are found within the carboxy-terminal two thirds.
The 3' UTR is a tripartite structure consisting of a conventional 3' end, a poly(U) tract and a 3' X-tail. This X-tail is highly conserved and forms an elaborate stem-loop structure, suggesting a possible role in replication[31].

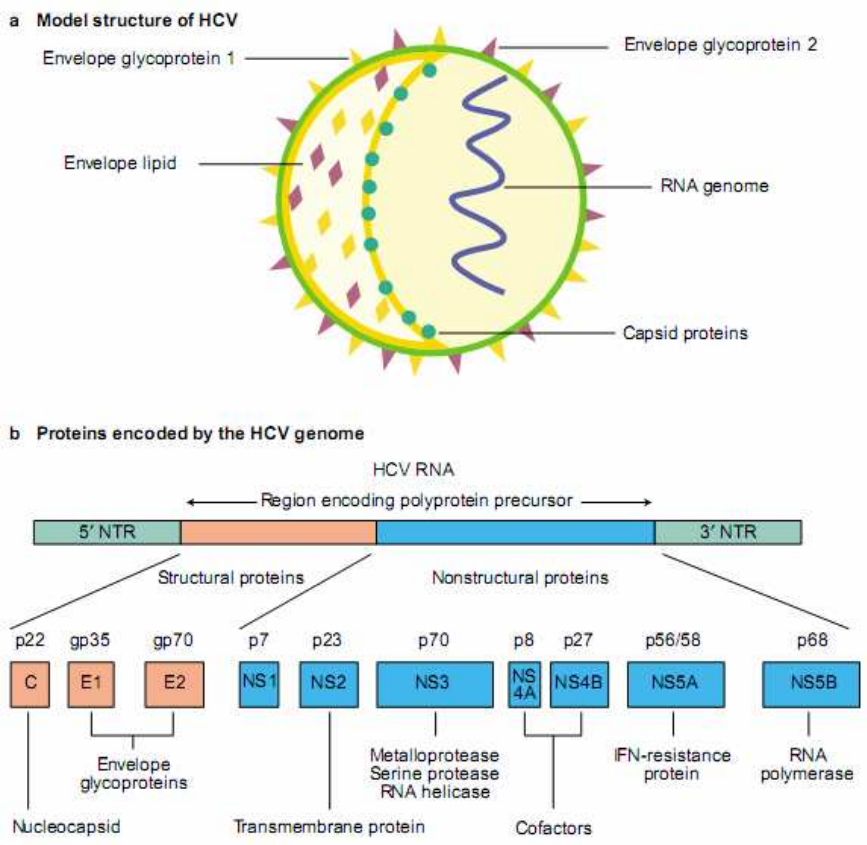


Figure 1.6: Model structure of HCV and genome organization (taken from [4])

1.9. NS3 protein

NS3 is a multifunctional protein of 68 kDa, containing a serine protease domain within its N-terminal third and an NTPase/helicase domain within its C-terminal two-thirds [6][50][59][60] (figure 1.7). The two functional domains are not cleaved from each other yet act independently of each other. The serine protease domain is required for 4 cleavage events, acting in cis to release itself from the HCV polyprotein and in trans to produce the N-termini of NS4B, NS5A and NS5B[108]. In addition to the N-terminal third of NS3, a C-terminal domain of NS4A has been described as an NS3 co-factor,

required for efficient cleavage of the downstream polyprotein, especially at the NS4B/5A cleavage site[25], resembling that seen in flaviviruses and pestiviruses. The NS3/4A protease has also been shown to control host cell antiviral defences by disrupting pathways that lead to activation of interferon regulatory factor 3 (IRF3) and subsequent induction of type I interferon. Identification of an interaction between NS3 and NS2 suggests NS2 may function as an anchor to retain NS3 at the ER membrane until cleavage of NS4A[49].

The C-terminal two-thirds of NS3 possess NTPase and helicase activity. Polynucleotide-stimulated NTPase activity, capable of hydrolysing all NTPs and dNTPs, has been shown, while RNA helicase activity, requiring ATP and a divalent ion, has also been identified: the NS3 C-terminal domain is capable of unwinding RNA-RNA, RNA-DNA and DNA-DNA substrates in a 3'-5' direction [101].

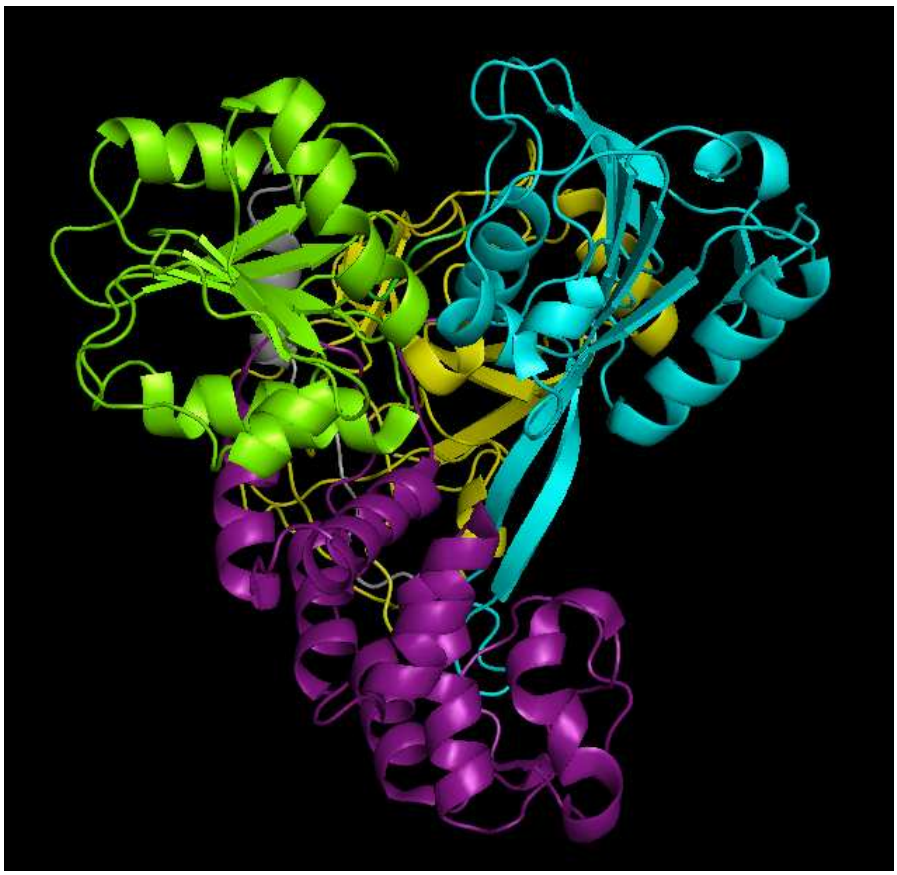


Figure 1.7: Structure of NS3 protein (Protein Data Bank code 1CU1). Protease domain is depicted in yellow, helicase subdomains 1,2 and 3 are in green, cyan and violet respectively

Sequence-based classification has placed the HCV NS3 C-terminal domain into helicase superfamily II (SF-II), along with nearly 100 proteins encoded by prokaryotic, eukaryotic and viral genes. The SF-II helicases play critical roles in transcription, RNA processing, translation, and RNA replication. Prototypical members of this family include translation factor eIF-4A, DNA repair enzymes UvrB, Rad-3 and Ercc-3, the human RNA processing enzyme helicase A, and the vaccinia virus RNA helicase NPH-II.

Crystal structures of the HCV NS3 helicase domain by itself[47], in complex with single-stranded DNA[53], and in the bifunctional protease-helicase NS3 protein complexed with NS4A co-factor peptide[52] showed that it is composed of three nearly equal-sized subdomains. Subdomain 1 (residues 181-326 of NS3) and subdomain 2 (residues 327-481 of NS3) have little sequence identity, but share the same structure composed of a large central β-sheet flanked by α-helices, and are homologous in structure to the central region of the RecA protein. Subdomain 3 (residues 482-631 of NS3) is mostly α-helical and contains part of the single-stranded nucleic acid binding site. Subdomains 1 and 3 share a more extensive interface than either share with subdomain 2. Therefore, subdomains 1 and 3 form a rigid unit, whereas subdomain 2 is connected to subdomains 1 and 3 by solvent-exposed polypeptide segments capable of supporting large scale, relative rotations of subdomain 2. In particular, an unusual molecular feature is a long antiparallel β-loop that extends from the central β-sheet of subdomain 2 to subdomain 3 where the end of the loop becomes an integral part of the subdomain 3 structure. Thus, similar to other helicases, subdomain motions are characteristic for the activity of the HCV helicase.

All helicases crystallized to date contain domains that resemble domains 1 and 2, but none share a domain that resembles domain 3. In some helicases, such as PcrA[97], two domains replace domain 3, one which extends from domain 1 (called domain 1B) and one that extends from domain 2 (called domain 2B). In several helicase structures that share domains similar to domains 1 and 2 of HCV helicase, such as the RecQ protein, DnaG, and eukaryotic translation initiation factor 4A[12], domain 3 is missing entirely, suggesting that domain 3 might not be required for HCV helicase movements. This is not the case, however, and although its role in unwinding is only beginning to be understood, domain 3 is clearly essential. Deletion of 97 amino acids from the C-terminus of NS3 results in an inactive helicase[51].

An active NS3 helicase domain has recently been reported to be required for replication of an HCV subgenomic replicon[61]. In spite of these evidences, since there is no DNA stage in the HCV lifecycle, the exact function of HCV helicase is still unclear. It could unwind duplex RNA that is formed when the single-stranded HCV genome is copied, or it might smooth RNA secondary structures, which impede the NS5B RNA-dependent RNA polymerase. Alternatively, the ability of HCV helicase to move like a motor

along RNA could be used for a process not linked to bona fide helicase activity (i.e. the disruption of base pairs). HCV helicase could strip RNA binding proteins from viral RNA, assist translation, or even help coordinate translation and polyprotein processing. Moreover, NS3 has been shown to bind many cellular factors including protein kinase A, protein kinase C, tumour suppressor p53 and histones H2B & H4 and has effects on various processes such as cell metabolism, differentiation and tumour promotion[102].
The multifunctionality of NS3, together with its interactions with numerous cellular factors, indicates that this protein plays a pivotal role in the life cycle of HCV, not just in terms of replication of the viral RNA but also by interacting with host proteins which may result in some of the pathogenic effects associated with HCV infection.
Whether NS3 exists *in vivo* as a monomer or as an oligomer is still debated.
Ever since the HCV helicase portion of NS3 was first purified, it was apparent that it behaved as a monomer and did not need to oligomerize to cleave ATP. Initial studies found HCV helicase to act as a monomer in solution based on gel filtration data [80]. The monomeric enzyme has all the residues necessary to catalyze ATP hydrolysis, and as a result, no decrease in turnover constant is observed when HCV helicase is diluted [64], in stark contrast with other helicases. There is also yet no direct structural evidence for dimerization. Although the protein exists as a dimer in the crystallographic asymmetric unit in PDB files 1HEI and 1CU1, it is a monomer in 1A1V and 8OHM. No evidence has been presented that interfaces seen in PDB files 1CU1 or 1HEI are biologically relevant. In addition, all HCV-encoded proteins are translated as a single polyprotein and therefore produced in equimolar amounts. It is therefore unlikely that NS3 exists in large excess relative to other viral proteins involved in RNA replication, e.g. the RNA polymerase NS5B and the NS5A protein.
While the evidence that HCV helicase acts as a monomer is convincing, there is also evidence that multiple subunits interact with each other to efficiently unwind RNA. Yeast two-hybrid assays provide the most persuasive evidence that NS3 interacts with itself [48]. In these experiments, three residues were identified that are critical for dimer formation, Thr266, Tyr267 and Met288. Mutations of these residues not only influence dimer formation that can be assayed using gel filtration, but also the ability of the protein to unwind DNA.
The latest evidence for oligomerization has emerged from measurements of rates of HCV helicase catalyzed DNA and RNA unwinding. Notably, unwinding rates are not linearly dependent on the amount of protein present in the reaction, but rather, accelerate greatly once a critical protein concentration is reached. Several kinetic models have been proposed that take into account and explain this cooperativity[64][92]. Ensemble and single molecule studies have established that NS3 is a processive helicase, capable of making multiple unwinding steps of well defined size without dissociating

from the substrate. This stepping behaviour involves alternating pauses and rapid unwinding events and is similar to the behaviour displayed by cytoskeletal motor proteins. The pauses are likely to be caused by the necessity to reset the conformation of the helicase-RNA complex after each unwinding step, consistent with an inch-worm model of unwinding. The kinetic models cited above also attempt to calculate the number of base pairs unwound in a single turnover event (called "step size"). Levin *et al*. have calculated a step size of 9 base pairs using DNA, and using a long RNA substrate, Serebrov and Pyle have determined that 18 base pairs are unwound by HCV helicase in a single step. The number of base pairs unwound by NS3 per step appears to be one of the largest reported for helicases to date, and it significantly exceeds the footprint of a monomeric NS3 bound to an RNA substrate (6-8 nt).
A model that represents a possible synthesis of the monomeric and oligomeric unwinding theories has been recently presented[93]. The experiments in this work revealed distinct active states of NS3, one of which is a nonprocessive but highly reactive monomer. This population of NS3 dominates the unwinding reaction under conditions in which the measured kinetic step size is lower, whereas more processive states of NS3 dominate under conditions in which the step size is higher, thereby demonstrating a link between oligomeric state and kinetic step size of the helicase. These experiments reconcile the apparent discrepancy in reported step sizes, and they establish a model for unwinding that is consistent with both ensemble and single molecule studies. These findings implicate an unwinding mechanism in which NS3 monomer is the minimal functional unit capable of RNA remodeling, and where increased processivity appears to be the only characteristic of NS3 unwinding that is strongly affected by oligomerization, as kinetic parameters for unwinding remain essentially the same.

1.10. Diagnosis of HCV infections

Historically, alanine aminotransferase (ALT) and aspartate aminotransferase (AST) levels were used as markers for suspect HCV infection, even though they are nothing more than surrogate markers for HCV, and are now considered obsolete by many[9].
The first proper marker of HCV infection is serum HCV RNA detectable by PCR as early as 1 week after infection and increasing up to 10^6-10^8 genomes/ml. Non-enveloped circulating nucleocapsid particles also rapidly appear and represent an early marker of acute HCV infection[67].
Different antibodies appear in serum at different intervals from the time of initial inoculation. Unlike the usual pattern of immunoglobulin response, presence of anti-HCV IgM is not a reliable marker of acute infection, as anti-HCV IgM can be detected in most patient with chronic HCV infection. Anti-core IgG directed to the nucleocapsid and NS3 proteins are generally

the first to appear and can be detected by the time ALT is peaking, within days to weeks after onset of clinical symptoms[9]. However, protective antibodies have not been identified yet.

Diagnostic tests for hepatitis C can be divided into the following two general categories: serological assays that detect antibodies to hepatitis C virus (anti-HCV) or directly HCV antigens; and molecular assays that detect, quantify, and/or characterize HCV RNA genomes within an infected patient. Serological assays have been subdivided into screening tests for anti-HCV, such as the enzyme immunoassay (EIA), and supplemental tests such as the recombinant immunoblot assay (RIBA). While screening assays are widely used for initial screening of large populations (e.g. blood donors), supplemental anti-HCV tests are designed as confirmation tests to resolve false-positive testing by EIA. Detection of HCV RNA in patient specimens provides evidence of active HCV infection and is potentially useful for confirming the diagnosis and monitoring the antiviral response to therapy. Optimal HCV PCR assays at present have a sensitivity of less than 100 copies of HCV RNA per milliliter of plasma or serum. Molecular tests have also been developed to classify HCV into distinct genotypes; the clinical importance of HCV genotype determination resides in the different response of the HCV genotypes to interferon-ribavirin treatment.

Since the discovery of hepatitis C virus (HCV) in 1989, there have been many advances in the area of diagnostic testing for hepatitis C. The problem of HCV transmission by blood products has spurred extensive research in the area of blood product testing for antibody to HCV (anti-HCV). In this setting, anti-HCV screening assays have been designed to optimize sensitivity of detecting infected carriers of HCV to prevent the transmission of infectious virus. Although first generation screening assays have been extremely effective in reducing the risk of posttransfusion hepatitis C, they have also created a situation in which false-positive testing has occurred in some individuals who have no other clinical or laboratory evidence of HCV infection. In a low-prevalence setting like the healthy blood donor population, the false-positive rate for a screening assay actually exceeded the true-positive rate, creating a need for supplemental or confirmatory tests. Because of this need, supplemental anti-HCV antibody tests, such as the recombinant immunoblot assay (RIBA), were developed to aid in the diagnostic evaluation of seemingly healthy individuals who test positive in the anti-HCV screening assay. Supplemental anti-HCV assays have been used widely in the diagnostic workup of patients with suspected hepatitis C, under the assumption that supplemental tests can be used to confirm the diagnosis. However, the unequivocal diagnosis of hepatitis C requires either clinicopathologic correlation, such as typical histologic changes on liver biopsy, and/or demonstration of active HCV replication in an individual with a positive antibody screen, such as demonstration of viremia by reliable molecular assays[5]. In many infections, isolation of the pathogen by tissue

culture is useful for establishing the diagnosis. However, tissue culture propagation of HCV has been extremely difficult, with only a few reports of low-efficiency propagation in the research setting. Therefore, molecular assays have been developed to detect, quantify, and/or characterize HCV RNA genomes within infected patients. Demonstration of HCV RNA in a patient's blood by molecular assays can be extremely useful in the diagnostic workup. However, in other cases, such as a patient with a known risk factor for HCV infection plus a history of chronic abnormalities in serum aminotransferase levels, molecular assays are probably not necessary to make the diagnosis of chronic hepatitis C in a seropositive patient. Patients who test positive for HCV RNA in the blood are known to be actively infected by HCV and are at increased risk for developing significant liver disease. Common algorithms for HCV investigation currently suggest use of both serological and molecular tests in different phases of the infection (figure 1.8)

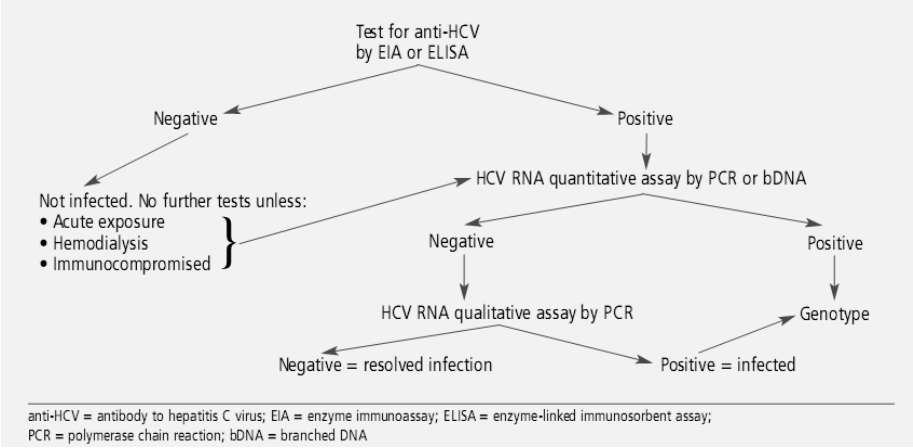


Figure 1.8: An algorithm for investigation of HCV infection (taken from [9])

1.11. Serological assays

1.11.1. EIA screening assays for anti-HCV

The main screening assay for detecting anti-HCV antibodies is the enzyme immunoassay (EIA), that has evolved into more sophisticated automated immunoassays (i.e. the CLIA format). Nevertheless, for the sake of simplicity, they can be treated as similar as the principle behind their mechanism is essentially the same.

The EIA has many advantages in the diagnostic setting, including ease of use, low variability, ease of automation, and relatively low expense. At least

four generations of anti-HCV tests have been developed, and each generation has resulted in an improvement in the sensitivity of detecting HCV infections.

The first generation anti-HCV test (EIA-1) contained a single HCV recombinant antigen derived from the nonstructural NS4 gene, designated c100-3. Although development of this test represented a dramatic breakthrough in terms of identifying patients with serologic evidence of HCV infection and reducing HCV transmission via blood transfusion, EIA-1 lacked optimal sensitivity and specificity. For example, in a high prevalence setting such as a tertiary care center, only approximately 80% of patients with clinical and molecular evidence of HCV infection tested positive for anti-HCV by the EIA-1 test[38].

On the other hand, in the low-prevalence blood donor setting, up to 70% of individuals who tested positive for anti-HCV by the EIA-1 test had false-positive test results, and were not infected with HCV. For this reason, supplemental anti-HCV tests (e.g. immunoblot assays) were developed to help reduce the number of individuals with false-positive anti-HCV screening tests.

The EIA-1 test was subsequently replaced in 1992 by the second-generation test (EIA-2). The EIA-2 test contains HCV antigens from the core and NS3 genes in addition to the NS4-derived antigen, and thus represents a multiantigen EIA. Introduction of the new antigens led to a substantial improvement in sensitivity and a slight increase in specificity relative to the EIA-1[54] (table 1.1)

TABLE 1. Sensitivity and Positive Predictive Value of EIA for Anti-HCV

		Positive Predictive Value† (%)	
Assay	Sensitivity* (%)	Low Prevalence	High Prevalence
EIA-1	70-80	30-50	70-85
EIA-2	92-95	50-61	88-95
EIA-3	97	25	Not Done

*Based on clinical findings and detection of HCV RNA by PCR.
†Compared with RIBA.

Table 1.1: Sensitivity and positive predictive value of anti-HCV EIA of generations 1,2 and 3 (taken from [37])

The use of core and NS3 antigens in the EIA-2 test shortened the average ‘‘window period’’ for HCV seroconversion, so that the mean time to seroconversion after blood transfusion was 10 weeks with the EIA-2 test versus 16 weeks with the EIA-1 test. In the high-prevalence setting, EIA-2 testing allows detection of approximately 95% of individuals with molecular evidence of HCV. Immunocompromised individuals, such as organ transplantation recipients or human immunodeficiency virus–infected

patients, may lack detectable antibodies by EIA-2 even in the presence of active viral infection. In the low-prevalence blood donor population, EIA-2 testing resulted in a further reduction in the incidence of posttransfusion hepatitis C[37]. A third-generation anti-HCV test (EIA-3) was launched in 1995. The EIA-3 test contains reconfigured core and NS3 antigens plus an additional HCV antigen (NS5) not present in the EIA-2 test. Several studies demonstrated an incremental improvement in sensitivity for detecting HCV infection in blood donors and liver clinic populations. The mean time to seroconversion in transfusion recipients was shortened by 2 to3 weeks, and the sensitivity for detecting anti-HCV in the high prevalence setting was improved to 97%[110]. Specificity of EIA-3 in the blood donor population appears to be slightly better than the EIA-2. Of interest, the slight improvement in EIA-3 sensitivity compared with the EIA-2 test has been attributed to reconfigured antigens already present in the EIA-2 test and not to the NS5 antigen[13]. A progressive improvement in sensitivity of detection of anti-HCV has been accomplished by the three generations of EIA screening assays (Table 1.1). However, testing in high-prevalence populations has indicated that not all patients with active HCV infection (eg, HCV RNA positive) are identified with the EIA screening tests.

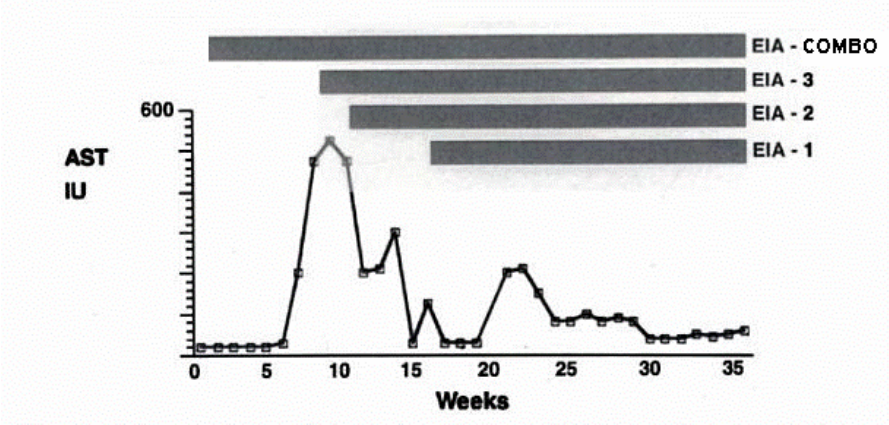


Figure 1.9: First detection of anti-HCV (and/or core antigen, for combo tests) after onset of acute HCV infection (adapted from[10])

Recently another generation of assays for testing of HCV infections has been established. These new generation is composed of the so-called “combo” immunoassays, that simultaneously detect anti-HCV antibodies and circulating HCV core antigen[3][94]. Inclusion of the core antigen detection system sensibly shortened seroconversion window by 6-7 weeks compared to 3rd generation tests, almost closing the gap between immunoassay and NAT technology (figure 1.9). Considering the lower sensitivity of HCV core antigen detection in comparison to NAT, the HCV combo assay is not

practical for the determination of the end of treatment response and sustained viral response, but could be useful for the determination of early viral response in the pegylated interferon-alpha and ribavirin treated patients infected with HCV. Furthermore, the HCV core antigen can be used as a marker of HCV replication in anti-HCV positive individuals in the areas of the world that cannot afford NAT and/or in the settings that are not equipped or competent to perform HCV RNA testing.

1.11.2. Supplemental tests for anti-HCV

Supplemental tests for anti-HCV were developed to help resolve false-positive EIA test results. The prototype supplemental test is the RIBA (Recombinant ImmunoBlot Assay), which contains the same HCV antigens as EIA in an immunoblot format.

Results are interpreted either as positive (two or more positive antigens), indeterminate (one positive antigen), or negative. Interpretation of HCV serology depends on the patient risk status. For example, in the low-prevalence blood bank setting, approximately 35% of specimens with positive EIA results used to be false positives (ie, RIBA negative) according to first and second generation tests[2]. In contrast, in a high-prevalence setting such as a university referral laboratory, approximately 97% of EIA–positive specimens were also positive by immunoblot, and less than 1% of EIA-positive specimens test negative by immunoblot [120]. Therefore, supplemental anti-HCV testing was not used for confirmation of infection in high-risk patients with a positive anti-HCV screen. However, improvement in false-positive detection by more advanced EIA tests is rapidly making supplemental tests unnecessary and obsolete.

1.11.3. Molecular assays

The presence of antibodies to HCV cannot distinguish between current and resolved infection. Moreover, only core antigen levels have any bearing on the likelihood of successful antiviral treatment. The advent of serum-based tests of viral presence represents a watershed in the evaluation and management of hepatitis C. However, several assays are in use for the detection of HCV RNA, and infected patients may be tested in different laboratories, each using a different test procedure. Until recently, even the units of expression lacked standardization. Many studies of HCV therapy expressed the amount of virus present (viral load) in copies per mL. Several studies selected 2 million copies per mL as the threshold separating “low” from “high” viral load[56]. However, because there was no comparability of quantified viral loads between assays, it was virtually impossible to interpret viral loads when different test systems were used. Assays now are standardized and expressed in IU per mL of serum[86]. A viral load greater than 800,000 IU/mL is currently considered high, regardless of the assay used[76]. Due to high sensitivity and virtual absence of a seroconversion

window, NAT test for the detection of HCV infections are becoming increasingly popular in western countries, sometimes substituting the immunoassay methods even for screening phase. Nevertheless, their elevated cost precludes their use in many developing countries [106].

1.11.3.1. Target vs signal amplification.

The NAT test most commonly used to determine the presence or absence of HCV is based on the polymerase chain reaction (PCR). This test detects the presence of minute quantities of HCV by first amplifying the quantity of HCV RNA in the sample, a technique referred to as target amplification. Transcription-mediated amplification (figure 1.10) is another target amplification test[28][88].

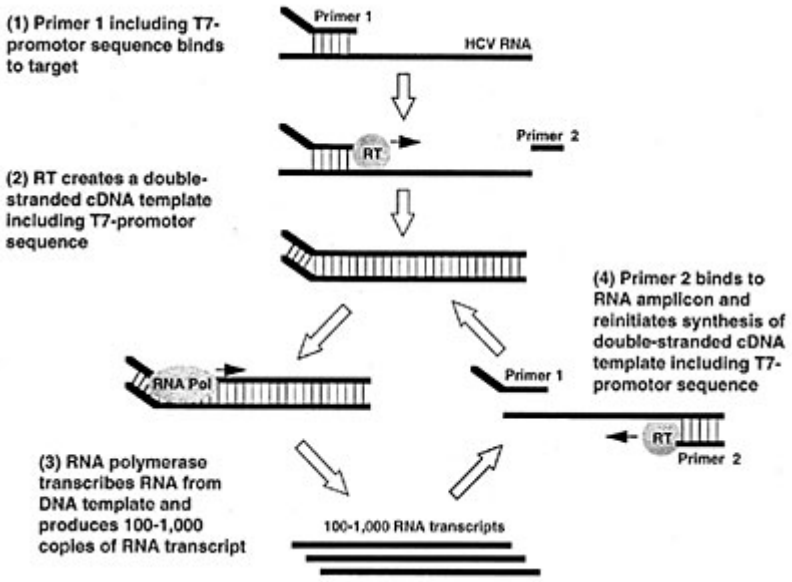


Figure 1.10: Mechanism of transcription-mediated amplification assay

Other tests, such as the branched DNA assay[71], operate by a different mechanism, referred to as signal amplification (figure 1.11). The principle of branched DNA testing is to bind branched probes to the target RNA, raising the number of signal molecules that can bind each target molecule. Most target amplification tests such as PCR are more sensitive than currently available signal amplification tests, and so yield fewer false-negative results. Target amplification tests are more complicated and more costly than signal amplification tests, and also take longer to perform. Signal amplification tests are technically simple, highly automated, rapid, and easily reproduced.

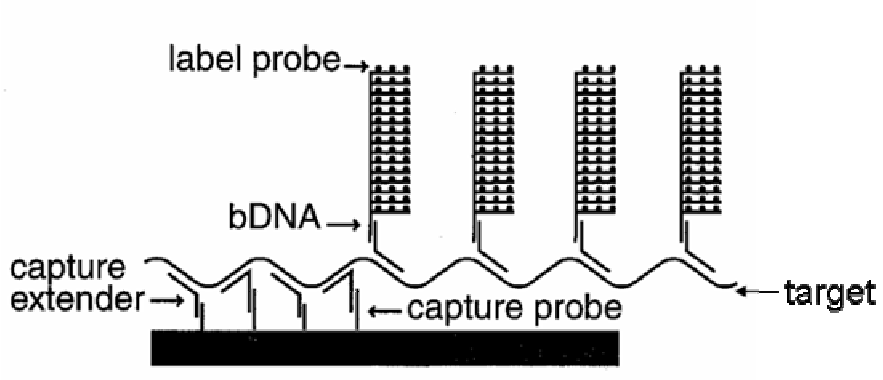


Figure 1.11: Mechanism of branched DNA assay

Their relative lack of sensitivity is their main drawback. Both types of tests are extremely specific. Apart from a contaminated system, false-positive results are rare. The current generation of PCR tests is quite sensitive, detecting HCV viral loads as low as 25 IU/mL. Levels of circulating HCV in individuals with untreated infection usually range from 50,000 to 5,000,000 IU/mL.

1.11.3.2. Qualitative vs quantitative assays.

The HCV RNA test kits designed to indicate viral load are not quite as sensitive as those that provide only a qualitative (present/absent) result. Because untreated individuals with HCV have viral levels so much higher than the threshold of detection, this small loss in sensitivity is not important, and quantitative HCV RNA testing should be ordered for these patients.
Results of qualitative PCR tests for HCV RNA are expressed as either positive or negative; viral load is not provided. Because of the slight loss in sensitivity with quantitative assays, a negative result on a quantitative PCR or branched DNA assay may be falsely negative and, in a person with suspected HCV infection, should be confirmed with a qualitative PCR test for HCV RNA. This is especially true when assessing treatment response. Patients with high viral load have a lesser likelihood of responding to available antiviral therapy[112]. In addition, viral load has implications for therapeutic "stopping rules." It is now clear that patients with HCV genotype 1 who do not achieve a 100-fold reduction in viral load after 12 weeks of antiviral therapy have less than a 5% chance of achieving such a response if therapy is continued for an entire year. As a result, antiviral therapy generally should be stopped after 12 weeks in such patients, since continuing treatment is usually not worth the associated cost and morbidity, given the low response rate. However, this criterion of a 100-fold reduction in viral load at 12 weeks does not apply to patients with HCV genotype 2 or 3, since such patients require only 6 months of antiviral therapy. [20]

1.12. The Liaison® system

The Liaison® system (Diasorin S.p.A., Saluggia, Italy) is an instrument designed to perform immunometric analyses of biological fluid samples (such as serum or plasma) in a completely automated way (figure 1.12). Up to 15 different tests can be performed at once on up to 144 samples in a sequential or random access mode. The output of the analysis is generated through the formation of an immune complex, followed by a chemiluminescent reaction that produces an emission of light.

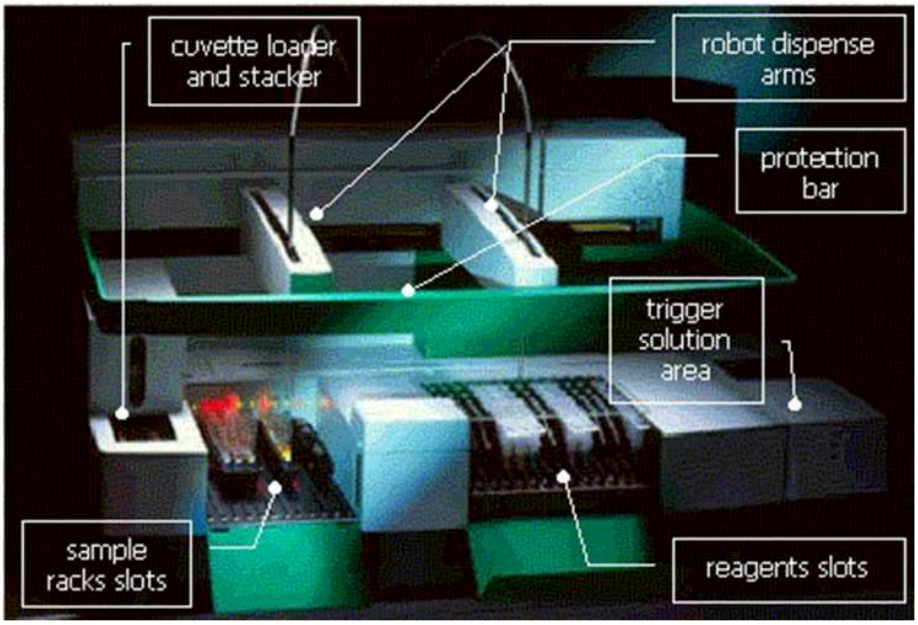


Figure 1.12: The Liaison® system

The instrument is composed of two modules, designed to be both allocated on a single workbench; the first module is a personal computer with touch screen, that hosts the user interface software and all the system data (assay protocols, reagent cartridges database, output of analyses, calibration history, network controls etc.). The second module is the actual analyzer, that performs the analysis from sample loading all the way to the final output for the user. Key components of the analyzer include:

- Cuvette loader and stacker: two conveyor belts allow continuous loading of the reaction modules, that are stored on a multilevel rack (7 levels).
- Sample rack slots: in the left-hand part of the instrument, a storage area can hold up to 12 sample racks, each carrying up to 12 samples. A barcode reader allows error-free catalogation of samples.

- Reagent slots: in the right-hand part of the instrument, another storage area can hold up to 15 different reagent cartridge simultaneously. This area is kept at a constant temperature of 15°C for optimal conservation of the reagents, while a stirring device keeps the microbeads always in homogenous suspension. Barcode reader for cartridge identification.
- Robot dispense arms: two robotic arms each carrying a dispensing needle. One arm is usually dedicated to dispensing samples, another to dispensing reagents. Each one has a separate washing well to clean the needle after each pipetting.
- Incubator: this area hosts the reaction cuvettes during incubation times, at a constant temperature of 37°C.
- Washing station: through the application of a magnetic field, this part allows retention of the paramagnetic microbeads and removal of the reaction liquids. Any number of washing steps with the desired washing buffer can be set.
- Read area: contains the injection devices of trigger reagents and the photomultiplier tube

The Liaison® system is based on two key features: the use of paramagnetic microbeads as the solid phase and the generation of signal by means of chemiluminescence.
The adoption of microbeads as the solid phase instead of the classic immunoassay supports, as the ELISA microwells gives a clear edge in terms of available reaction surface, which in turn increases the kinetic rates of the antigen-antibody complex formation.

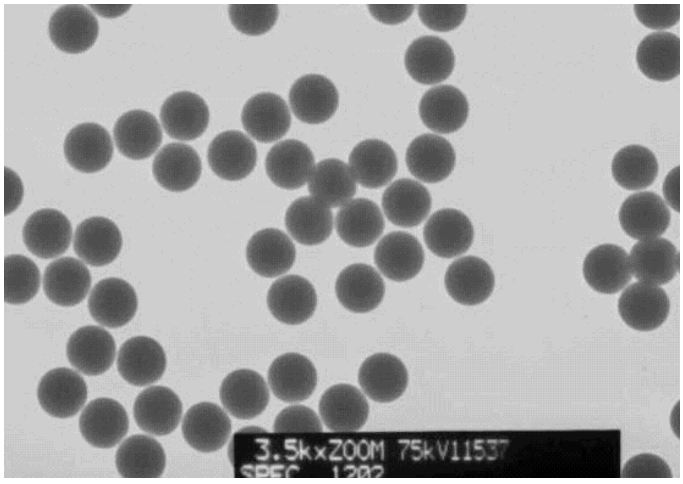


Figure 1.13: The paramagnetic microbeads used in the Liaison® system

Moreover, diffusion of both the analyte and the solid phase in the reaction volume is allowed, while in ELISA system only the analyte can diffuse, decreasing the possibility of the immune complex formation. The microbeads adopted in the Liaison® system (figure 1.13) are colloidal particles composed of a ferric oxide core covered by a polystyrene layer formed by spontaneous coalescence of polystyrene linear chains. This structure is in turn coated with another layer composed of polyurethane activated with tosyl- groups. The tosyl- group (4-toluenesulfonyl chloride) can undergo nucleophilic attack (figure 1.14), allowing the beads to covalently bind proteins through their available aminic groups (ε-amino groups of lysines, N-terminal end).

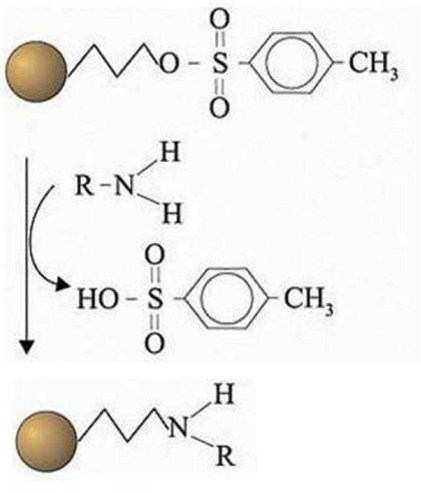


Figure 1.14: Chemistry of the covalent binding of amines to the tosyl-activated beads.

The paramagnetic properties of these microbeads allow easy manipulation through the application of a magnetic field. The particles respond to a magnet but are not magnetic themselves and retain no residual magnetism after removal of the magnet.
The tracer molecule is an antigen or antibody conjugated to a signal generating compound. Chemiluminescent tracers are formed by conjugating the antibody (or antigen) to a molecule that can generate a photon emission upon addition of certain reagents. The entity of this photon emission is measured with a luminometer, usually equipped with a photomultiplier tube.
The chemiluminescent molecule used in the Liaison® system is the luminol derivate ABEI (N-(4-Amino-Butyl)-N-Ethyl-Isoluminol), which is converted to its activated ester to allow conjugation with the antibody or antigen (figure 1.15).

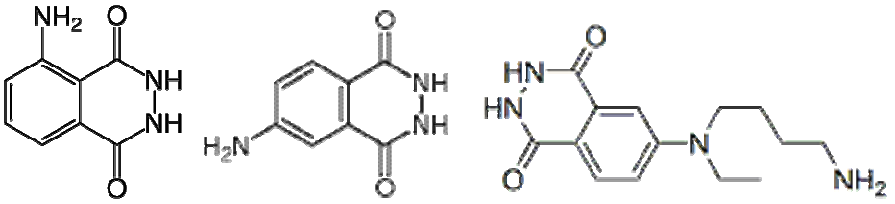


Figure 1.15: Luminol, isoluminol, and ABEI

In presence of H_20_2 and a microperoxidase (deuteroferriheme), ABEI achieves an excited state in consequence of a chemical reaction. This excited level decays to the ground level generating energy in form of light (fig. 1.16).

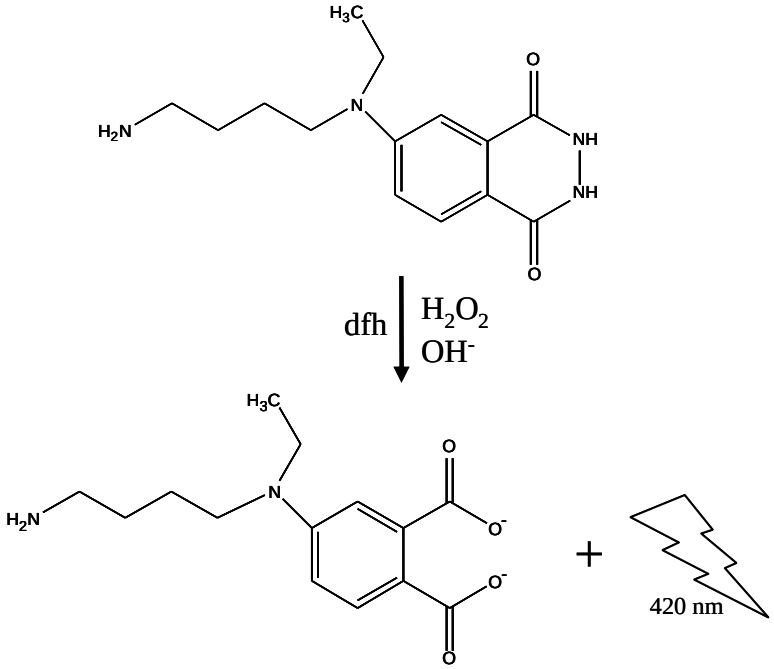


Figure 1.16: Chemiluminescence reaction of ABEI

The emission of light is recorded by the photomutiplier tube for an interval of just 3 seconds ("flash" chemiluminescence) and the signal is integrated over this interval (fig. 1.17). The final result is expressed in RLUs (Relative Light Units).
Using chemiluminesce is a great improvement over enzymatic signal generation of classic ELISA format assays. Sensitivity is highly increased and a greater dynamic range can be achieved.

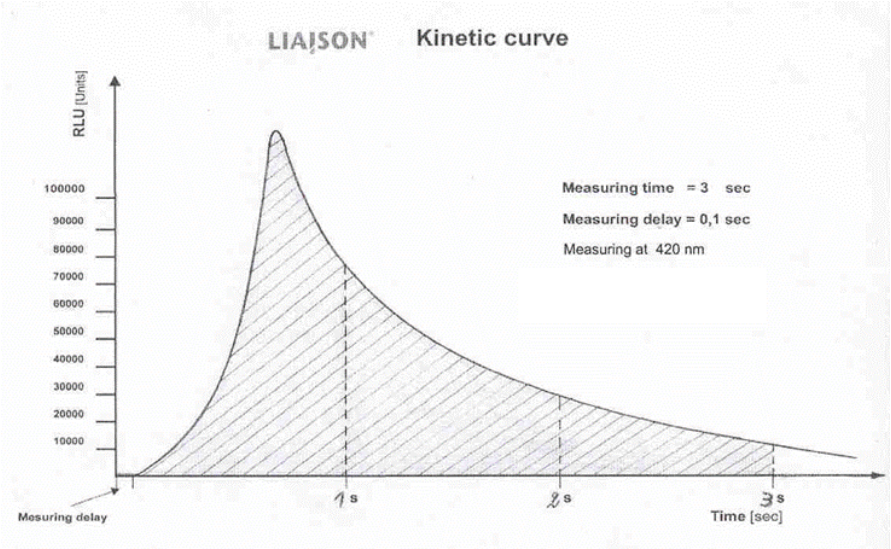


Figure 1.17: "Flash" chemiluminescence signal integration of the Liaison® system

Lower molecular weight and steric hindrance of ABEI compared to horseradish peroxidase allow conjugation of more signal generating molecules per tracer molecule. Moreover, generation and recording of signal is completed in a very short time (3 seconds), with a sensible throughput increase.

2. Aim of the work

Almost every anti-HCV antibodies diagnostic kit currently on the market makes use of the NS3 antigen, in different variants. Some of the NS3 antigen forms included in commercial diagnostic kit are shown in fig. 2.1.

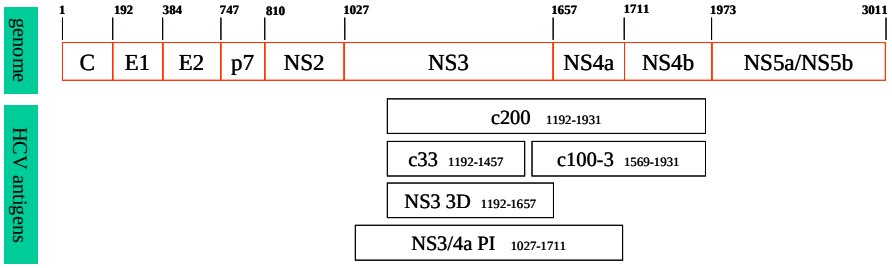


Figure 2.1: NS3-derived antigens used in various commercial HCV immunoassays

ETI-AB-HCVK-4 is the anti-HCV EIA kit sold by DiaSorin. It makes use of highly antigenic determinants of both the structural and non-structural regions of HCV, including NS3, in the form of c33 antigen, to detect IgG directed against HCV. Large shares of the market, especially in most developed countries, are shifting from the traditional ELISA immunoassay towards more automated systems, for reasons dictated by throughput and detection performances. Therefore, the necessity to fit the kit to the automated Liaison® platform is clear.
Adaptation of an EIA immunoassay to the CLIA format usually implies a development process that, in most simple scenario, only involves optimization of reaction variables (incubation times, washing steps, signal acquisition, concentration of the reagents). If these operations are not enough to obtain a kit with satisfying performances, it can be necessary to rethink the kit structure in a more dramatic way, such as modifying antigens and/or assay format.
Therefore, the first attempt to fit the ETI-AB-HCVK-4 kit on the Liaison® platform was made with the most simple approach, that is using the same reagents of the EIA kit (antigens and solutions), with solid phase support being the only sensible difference (paramagnetic microbeads instead of the ELISA microwells). This work was made by the Diasorin laboratories in Saluggia (Vercelli, Italy).
The assay format adopted for the CLIA assay prototype is the indirect, two-step format (figure 2.2). This method requires a solid phase, where one or more antigen are immobilized, and a labeled secondary antibody for detection. First, the serum or plasma specimen is added to the suspension of antigen-coated microbeads. An incubation at 37°C follows, during which

antibodies (IgG being the main fraction) directed against the immobilized antigens, if present in the sample, are captured on the microbeads. Then a magnetic field is applied and all the unbound material is washed away. The magnetic field is then released and an ABEI-conjugated monoclonal IgG specific for the constant portion of human IgG is added. Another incubation follows, during which the ABEI-labeled antibody can bind IgG captured on the microbeads by the solid-phase antigens. Another washing step follows, then the chemiluminescent reaction is primed with the addition of hydrogen peroxide, sodium hydroxide and deuteroferriheme. Finally, the emission of light at 420 nm is measured.

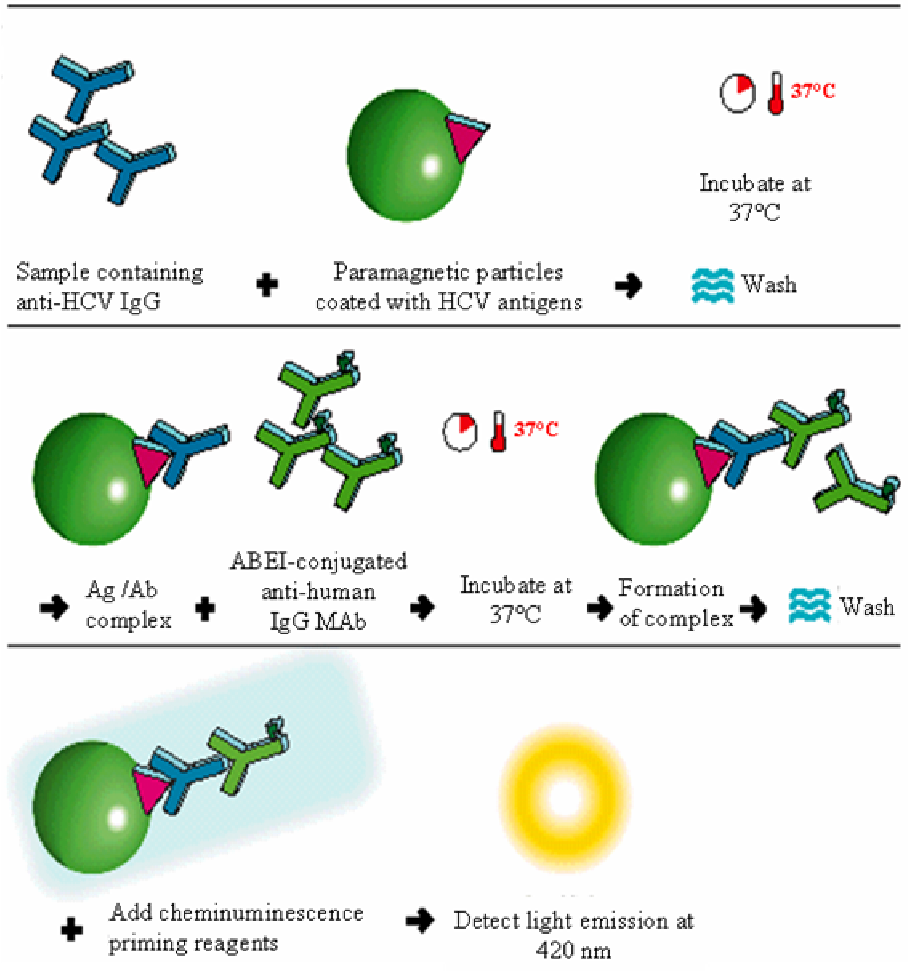


Figure 2.2: Mechanism of the indirect, two-step immunoassay format

The preliminary evaluation of this prototype on human sera from HCV-positive subjects, negative controls and seroconversion panels showed a

substantial equivalence of the CLIA prototype to the EIA kit in terms of sensitivity and specificity. Unfortunately, the prototype also showed extreme instability of its performances over time, specially after thermal stress at 30°C or 37°C. In fact, after only a few hours of storage at a temperature above 4°C, the chemiluminescence signal intensity obtained testing the same sample is greatly reduced. The chemiluminescent signal reduction is such that samples that were dosed as highly positive before stressing the kit, after 2/3 days of thermal stress became undistinguishable from negative samples. This is a particularly critical issue, because thermal stress stability is a stringent requirement for prototype acceptance and final commercialization of the kit. Even if storage of the kit at 4°C by the end-user is always recommended on the package and in the user manual, exposure of the kit to thermal stress during shipping is a possible event (i.e. waiting on airstrip during summertime or in warm countries). Therefore, at least some kind of resistance to thermal stress is required for industrializaton of the kit. In most cases, an acceptable threshold is maintaining at least 70-80% of the original RLUs after 3 days of storage at 37°C.
Separate stability studies performed on each antigen of the ETI-AB-HCVK-4 kit (c33, core and NS4) demonstrated that the thermal instability is caused almost exclusively by the c33 antigen.
The aim of this work was to solve or reduce the thermal instability problem of the c33 antigen in the environment provided by the Liaison® platform, bringing the thermal stability of the prototype to acceptable levels.
The most desirable scenery would have been finding a solution to the thermal instability problem not implying a change in the antigen primary sequence. This would have allowed to maintain the existing procedures and process already developed and consolidated for the ELISA c33 antigen.
Thus, the first possible solutions to be examined were the ones that could be applied to the existing c33 antigen without generating new constructs. Failure of this approach switched the focus to the realization of new antigens based on the NS3 protein.

Disclaimer:
All the antigens disclosed in this work were realized for research use only, and are not part of any device currently manufactured and/or sold by Diasorin S.p.A.

3. Materials and methods

Strains and vectors
E. coli XL1-Blue strain (Stratagene, Agilent Technologies, Santa Clara, USA) was used for cloning operations and plasmid maintenance.
E. coli BL21(DE3) strain (Novagen, Merck Chemicals Ltd, Beeston, UK) was used for protein expression. *E. coli* OverExpress C41, C43, C41pLys and C43pLys are from Lucigen (Middleton, USA).
Vectors pET-24b(+) and pET-30b(+) (Novagen, Merck Chemicals, Darmstadt, Germany) were used in this work.
Sequence span of the cloned antigens is reported in table 3.1. Aminoacids numbers refer to the HCV polyprotein reported in GenBank M62321.1 entry. SlyD and FKBP12 sequences are as reported in GenBank entries CAQ33668.1 and AAA58472.1

Antigen	Vector	AA	Cloning
c33	pET-30	1188-1463	BglII-XhoI
SlyD-c33	pET-24	1188-1463	EcoRI-HindIII
c33-7aa	pET-30	1188-1456	BglII-XhoI
FKBP12-c33	pET-24	1188-1463	EcoRI-HindIII
c33eu	pET-30	1188-1456 + 1479-1507	BglII-XhoI
c33 2D	pET-30	1188-1507	BglII-XhoI
NS3 3D	pET-30	1188-1654	BglII-XhoI
NS3 3D w/o extra seq.	pET-30	1188-1654	NdeI-XhoI

Table 3.1: Description of the vectors used in this work

Synthetic DNA sequences were purchased from GeneArt (Regensburg, Germany)

Protein expression and purification
For flask protein expression, 20 ml of overnight cultures were diluted in 1 l of LB medium with the appropriate antibiotic, grown at 30°C under constant shaking and induced with 1 mM IPTG upon reaching an optical density of 0.6 OD_{600}/ml. Samples corresponding to 0.6 OD_{600} were collected each hour and after 3 hours of induction cells were harvested by centrifugation.
All proteins were purified by native fast protein liquid chromatography (FPLC) on an AKTA Explorer 10 system (GE Healthcare, Uppsala, Sweden).

Frozen cell pellets were thawed, resuspended in PBS buffer and lysed by sonication. Cell lysates were clarified by centrifugation and surnatant were loaded on a Chelating Sepharose FF column, conditioned with buffer A. The column was washed with 4 column volumes (CV) of buffer A, then protein was eluted with a gradient of buffer B (100% of buffer B in 3 CV). Fractions containing the protein of interest were pooled and diluted with buffer C, in order to bring the final concentration of ammonium sulphate to a value of 0,8 M, then loaded on a Phenyl Sepharose HP column equilibrated with buffer D. After elution with buffer E, fractions containing protein were pooled. Final polishing was performed on a Superdex 200 gel filtration column in buffer F. Protein concentration was measured by measurement of absorbance at 280 nm.
All the resins were from GE Healthcare (Uppsala, Sweden), and the various reagents were bought from Sigma-Aldrich (St. Louis, USA).

Circular dichroism
CD measurement were done with a Jasco (Easton) J-500A spectropolarimeter both in the far (250–190 nm) UV, by use of 0.2-cm path length cuvettes. Sample were at a concentration of 0.2 mg/ml in 20 mM phosphate buffer. Temperature of the sample was increased by steps of 5°C from 25°C to 80°C with a Haake thermostatic bath. After every step, the CD measurement was repeated.

Formaldehyde fixation (see 4.2)
c33 protein at 1 mg/ml was fixed by addition 37% formaldehyde to a final concentration of 1.8%. After 3 hrs of reaction at 4°C, the reaction mix was dialyzed against buffer F to remove formaldehyde excess. NH_4HCO_3 to a final concentration of 0.1 M were added to quench the reaction. The mix was incubated for 45’ at 4°C, then dialyzed again against buffer F.

Addition of cofactors
AMP-PNP (adenylyl-imidodiphosphate, tetralithium salt, Sigma-Aldrich) was added to the storage solution at a concentration of 150µM. $(dU)_{12}$ oligonucleotide (Primm, Milano, Italy) was added to the storage buffer at a concentration of 1.5 µM.

HPLC analyses
High performance liquid chromatography was performed with a Beckman Coulter System Gold 126 chromatograph (Beckman Coulter, Fullerton, USA) and a TSK-GEL G3000SW column (Tosoh Biosciences, Stuttgart, Germany), in PBS buffer.

Liaison® immunoassay

Dynabeads® M-280 Tosylactivated paramagnetic microbeads (Invitrogen, Carlsbad, US) were coated with the various NS3 antigens according to the manufacturer's instructions, at an antigen concentration of 1.5 µM in coating buffer (see 4.5.1). NS4 and core antigen were not used to reduce complexity of the system and isolate NS3-related reactivity. Blocking of the coated beads was performed with 0.1% casein, then the beads were washed and resuspended in storage buffer. Immunoassay is performed as described in chapter 2, with a tracer concentration (ABEI-Anti human IgG MoAb, DiaSorin, Italy) of 150 ng/ml and a sample volume of 20 µl. HCV positive sera (cited in this work as POS #) were bought from Trina Bioreactives (Nanikon, Switzerland) and diluted 1:50 in negative serum prior to use. Negative sera (cited in this work as POS #) and the open population of sera described in 4.5.2 were obtained from the hospital of Chivasso (Italy). Seroconversion points (from panels 6213, 6214 and PHV915) were purchased from ZeptoMetrix Corporation (Buffalo, USA). Interfering sera (cited as INTERF #) were obtained from the collection of sera of the Diasorin facility in Saluggia (Italy).

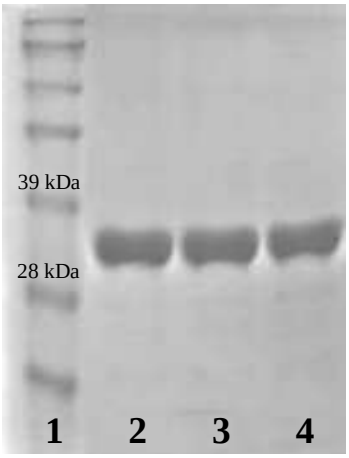
4. Results

4.1. Characterization of the instability phenomenon

The observed instability of c33 antigen reactivity during storage could be explained by many different factors. Several hypotheses were made regarding the factors involved in the process, which may comprise protein degradation, aggregation (in solution and/or on the beads) and conformational changes of the protein. First of all, some preliminary analyses were made to better understand the nature of the phenomenon causing c33 thermal instability.

4.1.1. Assessment of proteolytic degradation

The first approach was to investigate if the reactivity loss was due to proteolytic degradation of the protein, as this would have probably been the easiest problem to solve. Proteolytic degradation could destroy important antigenic regions of the protein, impairing its ability to bind antibodies.
Unfortunately, analysis of protein integrity in its most realistic environment (immobilized on the beads) is of difficult feasibility and would require techniques and studies beyond the intentions of this work. For this reason, the assumption was made that the environment provided by the beads is not related to a possible proteolytic event affecting the protein. This assumption is reasonable as the paramagnetic beads used in the assay are coated with an inert polysaccharidic layer.



Lane	Sample
1	Invitrogen® SeeBlue Plus 2
2	c33, t0
3	c33, after 3 days at 4°C
4	c33, after 3 days at 37°C

Figure 4.1: SDS-page of c33 (3 µg per lane) before and after 3 days of storage at 4°C and 37°C

Incubation of c33 antigen in solution at 37°C for several days does not cause evident proteolytic degradation, as reported by SDS-page (fig. 4.1) and HPLC analysis (data not shown) of the samples.
This result allows to rule out proteolytic degradation as a possible cause for the loss of reactivity.

4.1.2. Circular dichroism of c33

Circular dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light which arise due to structural asymmetry. The absence of regular structure results in zero CD intensity, while an ordered structure results in a spectrum which can contain both positive and negative signals. Secondary structure can be determined by CD spectroscopy in the "far-uv" spectral region (190-250 nm). At these wavelengths the chromophore is the peptide bond, and the signal arises when it is located in a regular, folded environment. Alpha-helix, beta-sheet, and random coil structures each give rise to a characteristic shape and magnitude of CD spectrum.

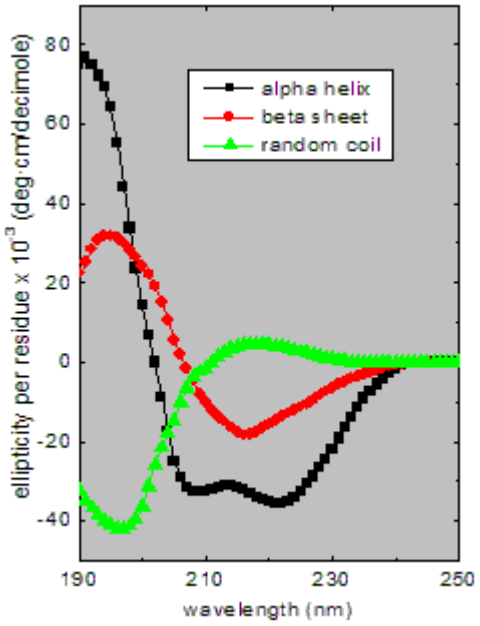


Figure 4.2: Characteristic CD curves of polypeptides with alpha-helix, beta-sheet and random coil secondary structure.

The approximate fraction of each secondary structure type that is present in any protein can thus be determined by analyzing its far-uv CD spectrum as a linear combination of such reference spectra for each structural type. Due to these characteristics, circular dichroism is an excellent spectroscopic

technique for following the unfolding and folding of proteins as a function of temperature. Secondary structure-specific signals can be measured over a wide range of temperatures to obtain informations on the structural rearrangements taking place during heating process.
Thus, a far-uv CD analysis of c33 was performed to better understand the rationale of its inactivation during thermal stress of the coated beads. The CD output clearly shows that a temperature-specific conformational change takes place, with a clear decrease of the β-sheet content to a more random coil structure. It is interesting to note that this transition happens over a narrow temperature interval between 30°C and 37°C. Over this temperature, no more structural changes are observed, but the CD signal at high wavelengths is not abolished, as should be in a completely unfolded protein; a signal characteristic of α-helical content is maintained up to the upper end of the tested temperature range (80°C). This suggests that this conformational change is not simply a generic thermal denaturation of the protein overall, but is more likely a structural rearrangement or unfolding of a specific portion of the protein. The transition T_m of 35°C is unsurprisingly very near to the temperature which the antigen inactivation happens at.

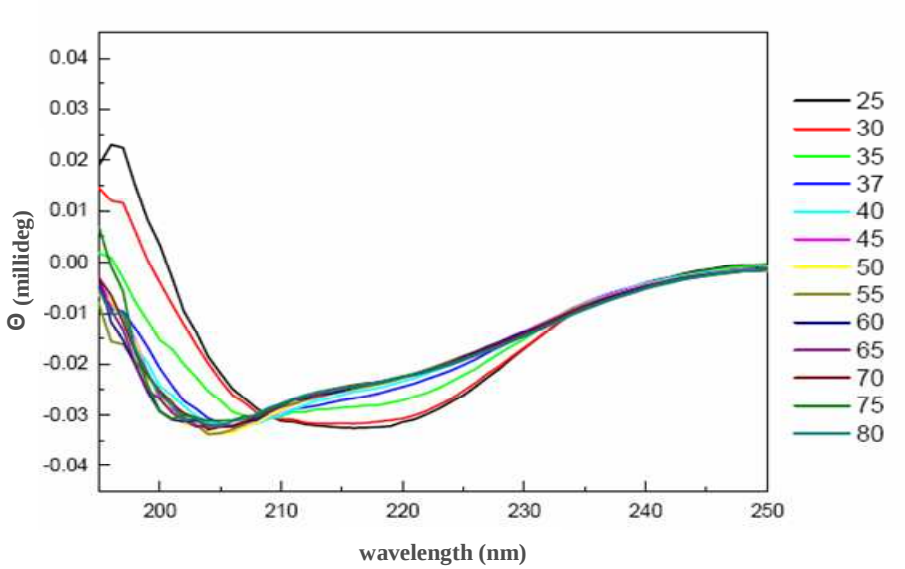


Figure 4.3: Far UV CD curves of c33 measured from 25°C to 80°C.

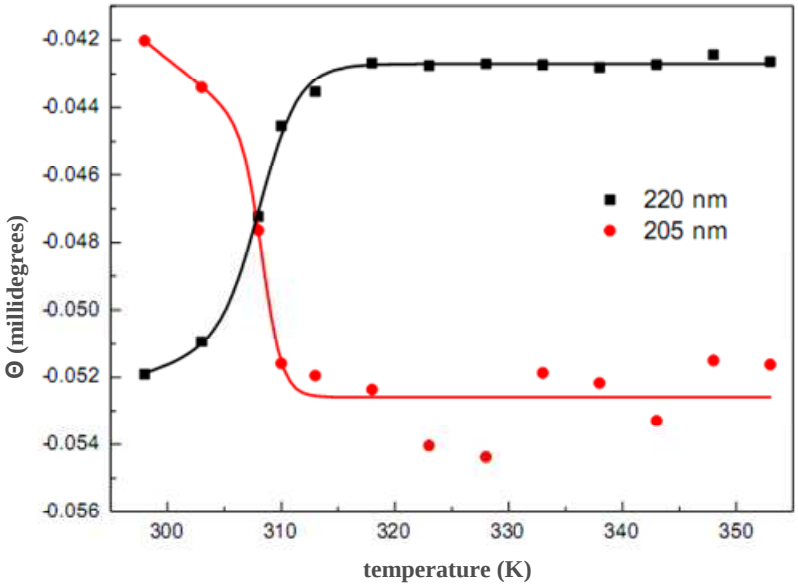


Figure 4.4: Extrapolation of ellipticity of c33 at 205 and 220 nm during the thermal scan reported in figure 4.3.

4.2. Formaldehyde fixation of c33

Formaldehyde fixation is a long-known method used in microscopy for specimen preservation and in vaccine production. One of the most common case is the preparation of *Clostridium difficile* toxoid, which is an inactivated toxin used as a vaccine [57].

Treatment of the toxin with formaldehyde leads to reaction with the N-terminal amino acid residues and the side chains of arginine, cysteine, histidine and lysine residues, and the subsequent formation of methylol groups and Schiff-bases as well as intramolecular methylene bridges with amine, phenol, imidazole or indole groups. The intramolecular crosslinking of toxin with formaldehyde leads to its complete inactivation with at least partial retention of its immunogenicity. In addition, the crosslinking may result in changes in the toxins' shape and overall stabilization of their structure [87].

This approach has been used on c33 antigen, with the aim of obtaining a substantial reduction of the degrees of freedom of the protein. This could prevent the putative structural rearrangement which causes the loss of reactivity on the beads.

First of all, a brief optimization of the reaction conditions was performed.

To control the possible formation of unwanted intermolecular aggregates, three fixation conditions were tested (1,8% formaldehyde for 3, 24 or 110

hours at 4°C), and aggregation status of the samples was analyzed on HPLC gel filtration (Fig. 4.5) .

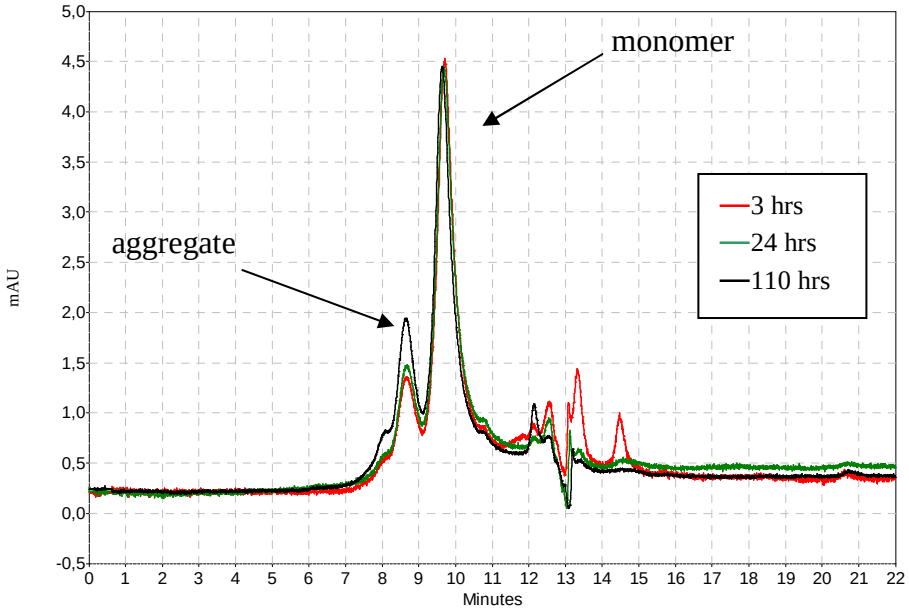
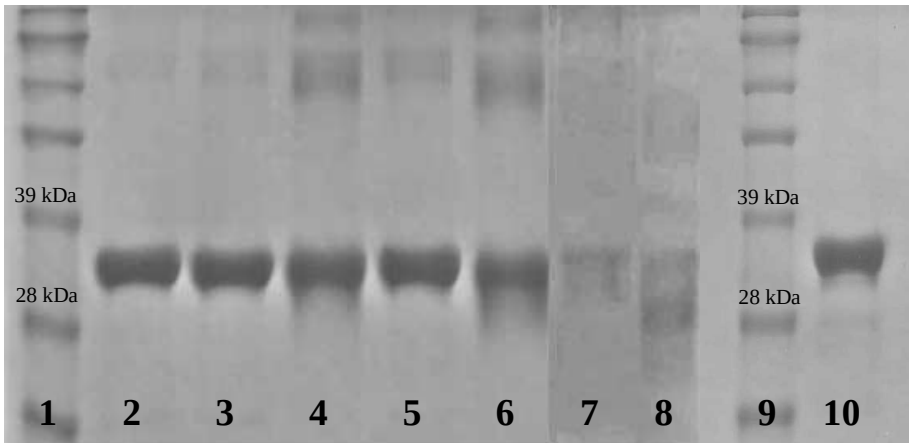


Figure 4.5: gel filtration HPLC profiles (A_{280}) of c33 fixed with 1.8% formaldehyde for different times

It can be seen that, upon formaldehyde fixation, some aggregates are formed, with longer fixation increasing the possibility of intermolecular crosslinks formation. The chromatographic profile the shows the accumulation of more aggregated forms with the increase of fixation time. Therefore, to minimize antigen alteration, the 3 hrs reaction was chosen.
To observe the behaviour of the formaldehyde treated protein during prolonged storage, formaldehyde-treated and untreated samples were stored at 4°C and 37°C for 24 h, 110 h and 30 days and analyzed with an SDS-page (Fig. 4.6).



Lane	Sample
1	Invitrogen® SeeBlue Plus 2
2	formaldehyde-treated c33, t0
3	formaldehyde-treated c33, 24 hrs at 4°C
4	formaldehyde-treated c33, 24 hrs at 37°C
5	formaldehyde-treated c33, 110 hrs at 4°C
6	formaldehyde-treated c33, 110 hrs at 37°C
7	formaldehyde-treated c33, 30 days at 4°C
8	formaldehyde-treated c33, 30 days at 37°C
9	Invitrogen® SeeBlue Plus 2
10	untreated c33, 110 hours at 37°C

Figure 4.6: SDS-page of formaldehyde-fixed c33 (3 μg) immediately after fixation and after storage at 4°C and 37°C. An untreated c33 sample stored at 37°C is reported in lane 10 for comparison.

The treated samples after storage at 37°C show a smeared band, due to different mobility of the heterogeneous intramolecular crosslinked forms generated. Also high molecular weight forms are generated, probably covalently linked dimers (according to the observed molecular weight) generated by intermolecular crosslinking. All these observation are in good correspondence with literature data about formaldehyde fixation of other proteins [74]

Both these process are more evident in samples stored at 37°C than those at 4°C.

It is interesting to note that both these processes happen during storage, after all unreacted formaldehyde has been removed. This suggest that some protein residues are in some way “activated” during the formaldehyde incubation and only after some time complete the crosslinking reaction, in a

temperature-dependent way. This is consistent with the nature of the formaldehyde fixation reaction, which is composed of two steps with different kinetics (Fig 4.7). It can be suggested that during incubation with formaldehyde only methylol groups and Schiff bases are formed on the protein (fast kinetic), and that the formation of stable methylene bridges (slow kinetic) requires a longer time and a higher temperature.

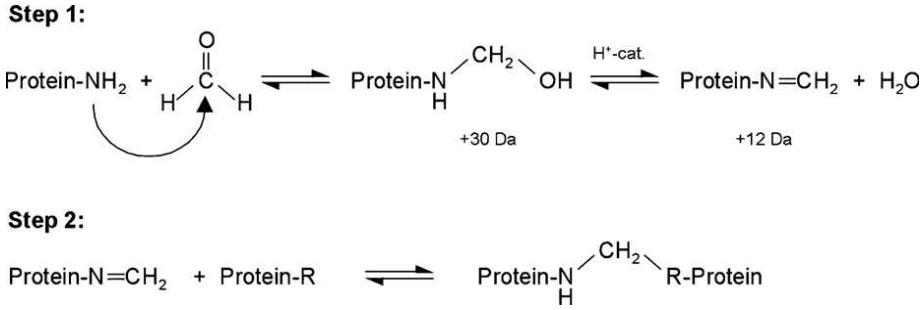
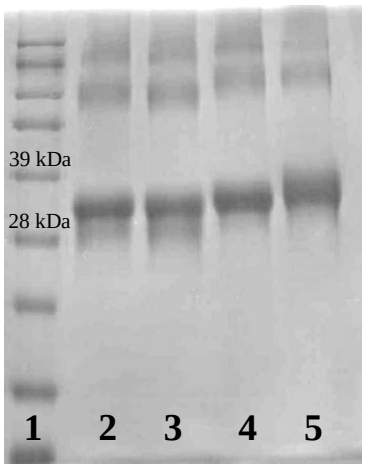


Figure 4.7: Two-step protein crosslinking reaction by formaldehyde. Taken from [107].

As this kind of uncontrolled crosslinking reaction can go on during storage, it could introduce an undesired variability in the sample during its shelf-life, posing a severe threat to the consistence and reproducibility of the diagnostic device. Therefore, a way to quench the reaction after the desired time is needed. In common histological practice this quenching is achieved by the addition of an excess of glycine [72] which stops the crosslinking reaction by inactivating the unreacted formaldehyde still loosely bound to the sample after washing and, probably, other reactive groups formed during formaldehyde incubation time. Such inactivation causes the addition of a glycyl group for every inactivated moiety, probably leading to a heavy modification of protein structure, which could in turn lead to a loss in antigenicity. To reduce this risk, other smaller nucleophilic agents were tested for their ability to quench the formaldehyde crosslinking reaction. Use of ammonium salts, particularly ammonium bicarbonate, is suggested in a patent by James and Hoag [43]. Sodium azide is another readily available nucleophilic reagent. To test these compounds, c33 protein was incubated with 1,8% formaldehyde for 3 hours at 37°, then unreacted formaldehyde was removed by dialysis and 0.1 M solutions of glycine, ammonium bicarbonate or sodium azide were used to quench crosslinking. In the negative control, water was added instead of the quencher solution. The quenched samples were then stored for 5 days at 37°C and loaded on a SDS-page. The result is shown in figure 4.8.



Lane	Sample
1	Invitrogen® SeeBlue Plus 2
2	formaldehyde-treated c33, not quenched
3	formaldehyde-treated c33, quenched with 0.1 M sodium azide
4	formaldehyde-treated c33, quenched with 0.1 M ammonium bicarbonate
5	formaldehyde-treated c33, quenched with 0.1 M glycine

Figure 4.8: SDS-page of c33 fixed with formaldehyde and quenched, respectively, with water, NaN_3, NH_4HCO_3 and glycine, after 5 days of storage at 37°C

Sodium azide showed to be ineffective in stopping the reaction, as no effect on electrophoretic mobility shift is seen upon addition of sodium azide. Ammonium bicarbonate and glycine supplemented samples contained protein with a lower electophoretic mobility than control sample. This probably reflects a lower level of crosslinking and/or a raise in the molecular weight of the protein due to the addition of amino or glycyl groups. This second possibility is suggested by the higher apparent molecular weight of the sample quenched with glycine compared to the one quenched with ammonium bicarbonate. Both compounds seem to be effective in stopping or slowing down the reaction, with ammonium bicarbonate producing a much sharper band, meaning a less heterogeneous sample than the one obtained with glycine.
Considering the lower steric hindrance of ammonium ions, which could mean a lower impact on the antigen surface, the sharper band, suggesting higher efficiency of ammonium bicarbonate quenching, and that blocking reaction intemediates with ammonium should introduce primary amino groups (useful for coating on the tosyl-activated beads), ammonium bicarbonate appears as the most attractive compound for crosslinking reaction quenching.
CD analysis of formaldehyde-fixed c33 (Fig. 4.9) shows very clearly that the β-sheet to random coil transition that was individuated in the first CD tests on c33 is completely abolished upon treatment with formaldehyde.

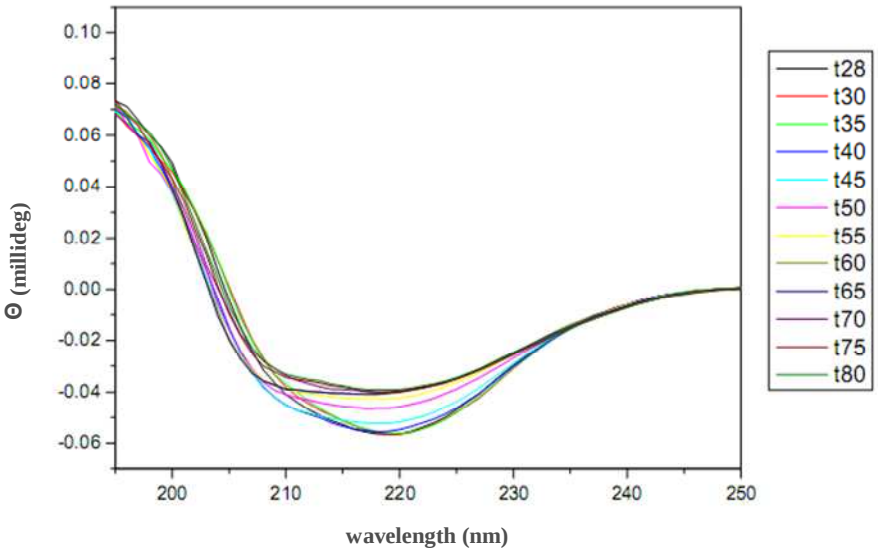


Figure 4.9: Far UV CD curves of formaldehyde-fixed c33 measured from 28°C to 80°C.

This suggests that formaldehyde fixation successfully addressed the structural rearrangement responsible of the CD spectra variation over thermal stress. Unfortunately, when the formaldehyde-fixed c33 protein was heat stressed and tested on the Liaison® system, it did not show a satisfying improvement over its non-treated counterpart (Table 4.1, graphical representation in Fig. 4.10)

Sample	**Antigen**			
	c33		formaldehyde-treated c33	
	t0	3 days at 37°C	t0	3 days at 37°C
POS 1	**457543**	**139652**	**380919**	**128667**
POS 2	**48829**	**16942**	**13953**	**6810**
POS 3	**19622**	**10417**	**8319**	**4636**
6214-10	**10246**	**2592**	**2178**	**933**
6214-11	**140772**	**34972**	**45909**	**10966**
6214-12	**145107**	**33574**	**56072**	**12840**
6214-13	**173156**	**37278**	**55947**	**12922**
PHV915-03	**5189**	**1670**	**3166**	**2633**
NEG 1	**1142**	**622**	**655**	**660**
NEG 2	**2134**	**1823**	**1103**	**1672**

Table 4.1: RLU values of c33 and formaldehyde-treated c33 before and after thermal stress

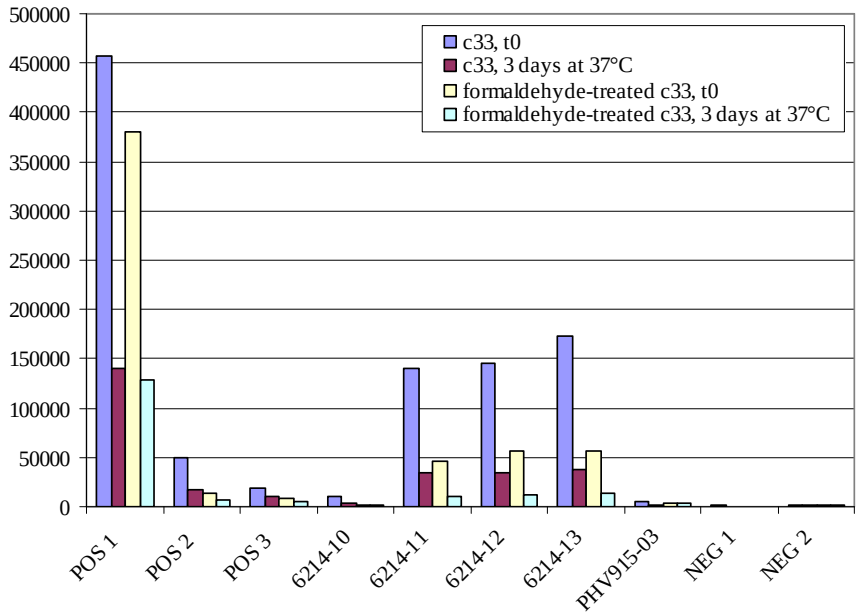


Figure 4.10: RLU values of c33 and formaldehyde-treated c33 before and after thermal stress

Signal loss was reduced, but not to a sufficient extent. This, in combination with a lower overall reactivity on positive samples and seroconversions (which could be expected due to the inevitable epitope alteration), led to RLU values after heat stress that were lower than the ones obtained with the untreated antigen.
Beside showing the failure of the formaldehyde fixation approach, these findings also suggest that the structural rearrangement highlighted by circular dichroism analysis, while evident and peculiar, is probably unrelated or only marginally linked to the heat instability phenomenon responsible of the signal loss on the Liaison® platform.

4.3. Adding cofactors

The interaction of proteins with small molecules, such as ligands and cofactors, often coincides with an increased stability of the protein due to the coupling of binding with the unfolding equilibrium. Thus, apart from their catalytic role, cofactors may also have a structural role. The common polymorphism C677T in methylenetetrahydrofolate reductase causing the single point mutation A222V reduces the affinity of the enzyme for the FAD cofactor, resulting in a lower thermal stability [104]. For the octameric protein vanillyl-alcohol oxidase, it was demonstrated that cofactor binding influences the quaternary architecture of the enzyme [100]. Similarly, for

lipoamide dehydrogenase, it was shown that FAD binding increases the protein melting temperature from 35 to 80 °C [121].
The NS3 helicase domain has at least two different binding sites for other molecules: one is the binding site for the ATP cofactor, necessary for the catalytic activity of the enzyme, and another is the binding site for nucleic acids. Binding of both these two ligands produces large movements into the structure of the helicase domain to allow its catalytic activity. Binding of polynucleotide by NS3 helicase in the absence of ATP leaves a large cleft between subdomains 1 and 2 (interdomain cleft). Binding of ATP through the β- and the γ-phosphate results in the closing of that interdomain cleft. After hydrolysis of ATP and release of AMP, the interdomain cleft is opened again [53].
Lack of ATP or substrate, which are so involved in the conformational status of the NS3 helicase domain, could be the reason for c33 instability to thermal stress. This, of course, with the assumption that c33, a truncated polypeptide, retains the binding capability and responsiveness of the wild-type NS3 helicase.
To check this possibility, the effect of the presence of cofactor and substrate in the storage buffer was evaluated. AMP-PNP, a non-hydrolyzable ATP analog, was added to the storage solution in an attempt to mimic the binding of ATP and (hopefully) block the protein in a status where the interdomain cleft is closed. Presence of substrate was mimicked by the addition of a polydeoxyuridine oligonucleotide, $(dU)_{12}$, that, according to literature data [64], had been selected among many tested polynucleotide as the one that most tightly bind the NS3 helicase domain of HCV.

Sample	**Antigen**			
	c33		c33 + AMP-PNP + poly(dU)	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 3	**19622**	**10417**	**20842**	**8745**
POS 4	**41587**	**15447**	**35876**	**14788**
6214-09	**3284**	**1773**	**4939**	**1302**
6214-10	**10246**	**2592**	**14930**	**1933**
6213-10	**4101**	**1589**	**3511**	**1487**
6213-11	**51847**	**23478**	**54210**	**19254**
NEG 1	**1142**	**622**	**1521**	**987**
NEG 2	**2134**	**1823**	**1989**	**2058**

Table 4.2: RLU values of c33 stored with or without AMP-PNP and poly$(dU)_{12}$ before and after thermal stress

Unfortunately, addition of AMP-PNP and poly(dU)$_{12}$ to c33 protein did not increase heat stability of c33, in terms of RLU values measured on the Liaison® system after thermal stress. (Table 4.2, graphical representation in Fig. 4.11)

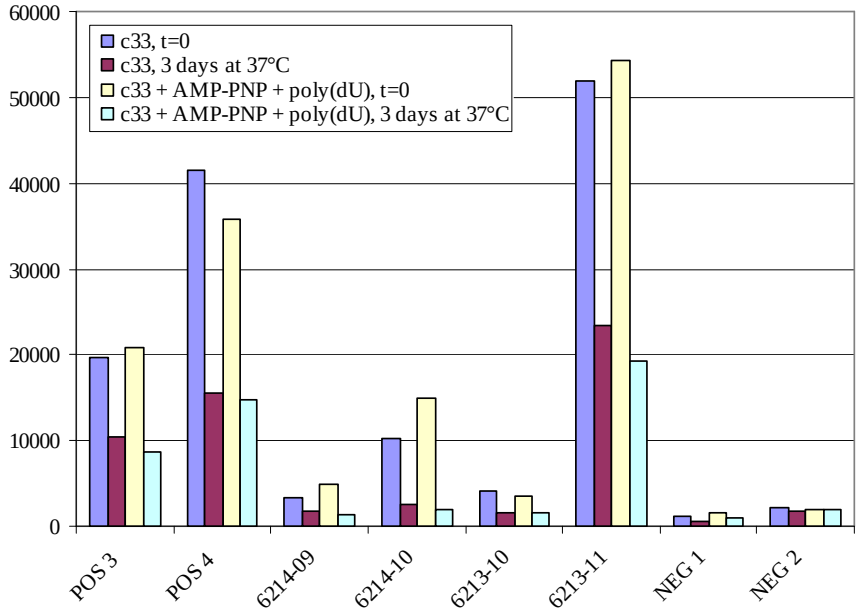


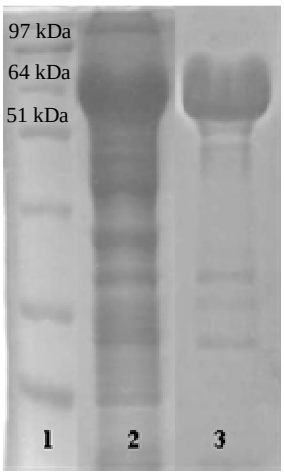
Figure 4.11: RLU values of c33 stored with or without AMP-PNP and poly(dU)$_{12}$ before and after thermal stress

4.4. Modification of primary sequence

As the previous attempts of improving thermal stability of c33 without modifying its primary structure did not have any success, designing new antigens based on the HCV NS3 helicase domain became inevitable. Producing a different antigen to be included in the anti-HCV assay would introduce new variables to be checked before industrialization of the kit; same or better diagnostic performances, absence of aspecific reactivity and feasibility of large-scale antigen production must be demonstrated. Nevertheless, adoption of a new antigen would be an acceptable solution, on condition that the thermal stability problem was solved. Starting from different hypotheses, several attempts to increase thermal stability by engineering the primary sequence of c33 were made and are now described.

4.4.1. Enhancing solubility: SlyD-c33

Aggregation of antigen during coating and/or aggregation of the coated beads were thought to be other possible factors for unstable reactivity of the c33 coated beads. This possibility was hinted by some previous HPLC profiles obtained during antigen development in Saluggia (data not shown) that showed a limited tendency of c33 antigen in solution to form soluble aggregates after a few days at 37°C. Aggregation of the coated beads during storage was excluded by direct observation of the beads with a common optical microscope. After 3 days at 37°C, the coated beads appeared monodisperse. Still, the possibility of aggregation of the protein during the overnight coating could not be excluded. To explore this possibility, a new antigen was made coding for c33 sequence fused in frame with two consecutive copies of the *E. coli* peptidyl-prolyl-isomerase SlyD, a molecular chaperone that has been shown to promote correct folding and solubility of aggregation-prone proteins even when expressed as a fusion partner [89][90].
Cloning in pET24 vector and expression in *E. coli* BL21(DE3) resulted in abundant production of SlyD-c33 protein (Fig 4.12, lane 2).



Lane	Sample
1	Invitrogen® SeeBlue Plus 2
2	Crude extract at the end of fermentation of *E. coli* BL21(DE3) pET24-SlyD-c33
3	Purified SlyD-c33

Figure 4.12: SDS-page of SlyD-c33 crude extract and purification product

Successful purification followed, with IMAC and GFC chromatography (Fig 4.12, lane 3). The retention volume of SlyD-c33 peak in the GFC profile corresponds to an apparent molecular weight of 161 kDa (Fig 4.13). This shows that in the experimental conditions SlyD-c33 is present in solution in form of a dimer. Purification yield was more than 17 mg per gram of wet cell weight.

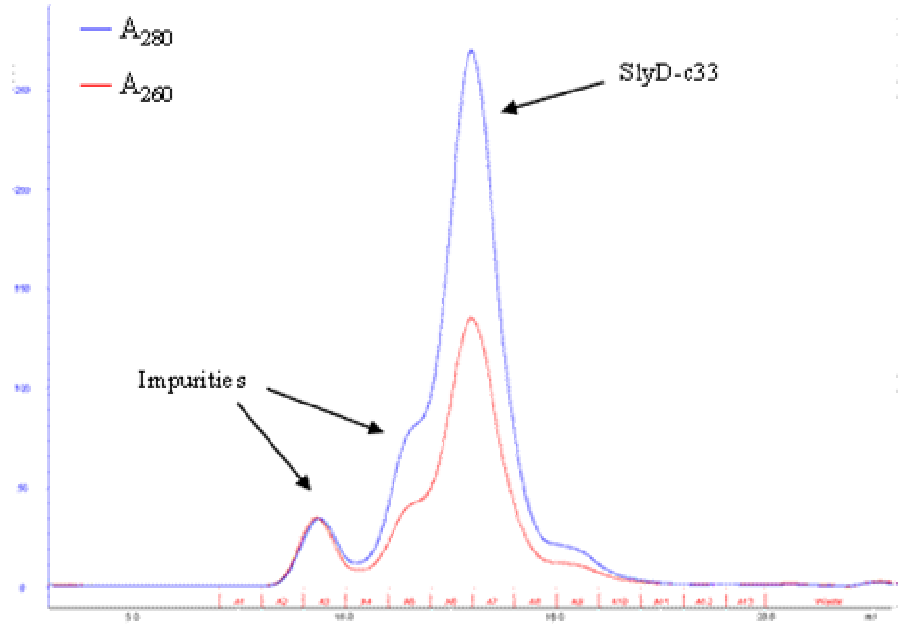


Figure 4.13: Gel filtration chromatography profile of SlyD-c33

SlyD has been previously shown to form dimers; therefore the observed dimeric structure could be due to the intrinsical dimerization property of SlyD. However, this dimeric structure is maintained over time without progressing further to more aggregated forms (the GFC profile is comparable after storage for 3 days at 37°C - data not shown). Nonetheless, when tested in the Liaison® immunoassay, the SlyD-c33 antigen showed no improvement in thermal stability compared to the original c33 antigen.

Sample	**Antigen**			
	c33		SlyD-c33	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 1	**457543**	**139652**	**285798**	**71652**
POS 5	**274674**	**82547**	**148954**	**35474**
POS 6	**18451**	**5554**	**20145**	**6447**
NEG 2	**1024**	**622**	**4829**	**5241**
NEG 3	**2134**	**1543**	**4407**	**4778**
NEG 4	**1823**	**989**	**4987**	**5320**

Table 4.3: RLU values of c33 and SlyD-c33 before and after thermal stress

Signal loss is of the same magnitude after incubation of the coated beads at 37°C. Also, RLU values of the negative samples are higher due to the probable presence of traces of anti-SlyD antibodies in human sera.

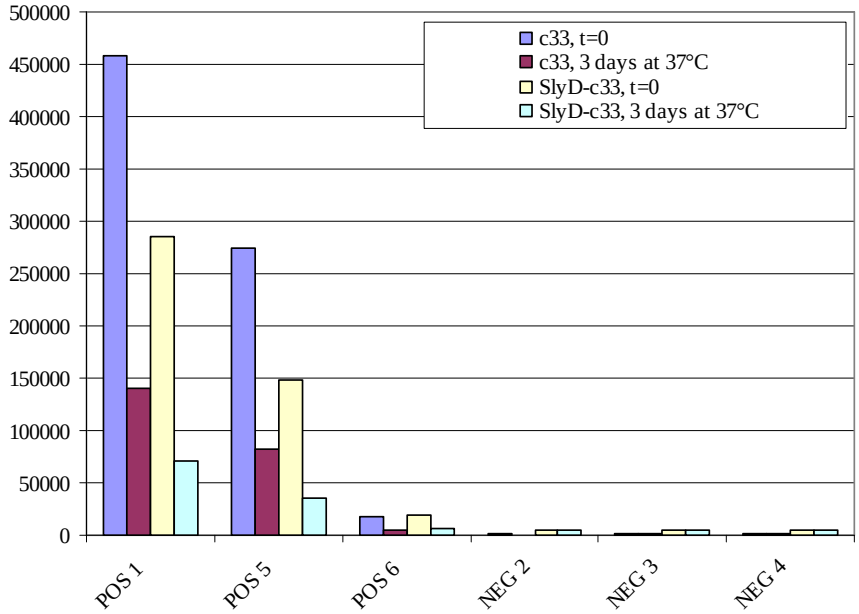


Figure 4.14: RLU values of c33 and SlyD-c33 before and after thermal stress

4.4.2. Avoiding loose ends: c33-7aa

Another set of hypotheses was made starting from the observation of the tertiary structure of the HCV NS3 helicase. By confronting c33 sequence with the whole NS3 protein, it can be seen that c33 sequence comprises the complete subdomain 1, the sequence acting as a linker between NS3 protease and helicase domains, and part of subdomain 2. In particular, the C-terminal end of c33 is located in the middle of the extended β-loop in subdomain 2 that, in wild-type NS3, protrudes towards subdomain 3. This β-loop of subdomain 2 embraces subdomain 3 forming a partially flexible hinge that allows the movement of NS3 along RNA strands.
In c33 antigen the extended β-loop is truncated to a 7 aminoacid stretch (of the original 20) and subdomain 3 is completely absent. It could be imagined that in these conditions, the seven remaining C-terminal aminoacids remain loose and solvent-exposed, possibly destabilizing the conformation of protein. The same could be told about the N-terminal end, which is composed by the flexible linker sequence and a short purification tag: the flexible linker, without the NS3 protease domain to connect to, could miss a

crucial component to adopt the correct configuration. Loose ends have previously been reported as a possible cause of low protein stability, and the removal or tying up of these ends is a common way to increase thermal and mechanical stability both in natural [19] and in engineered biological systems [111]. Removal of the complete extended β-loop has previously been reported for the *E. coli* expression of subdomain 2 of HCV NS3 helicase, in an effort to obtain a more soluble and stable product for NMR studies [66].

A new antigen named c33-7aa was designed whose sequence is identical to the c33 construct except for its C-terminal end, where the sequence is truncated at the beginning of the extended β-loop. This way, the remaining C-terminal seven aminoacids of the loop were removed, eliminating a putative loose end. The removed aminoacids are highlighted on NS3 helicase crystal structure in figure 4.15.

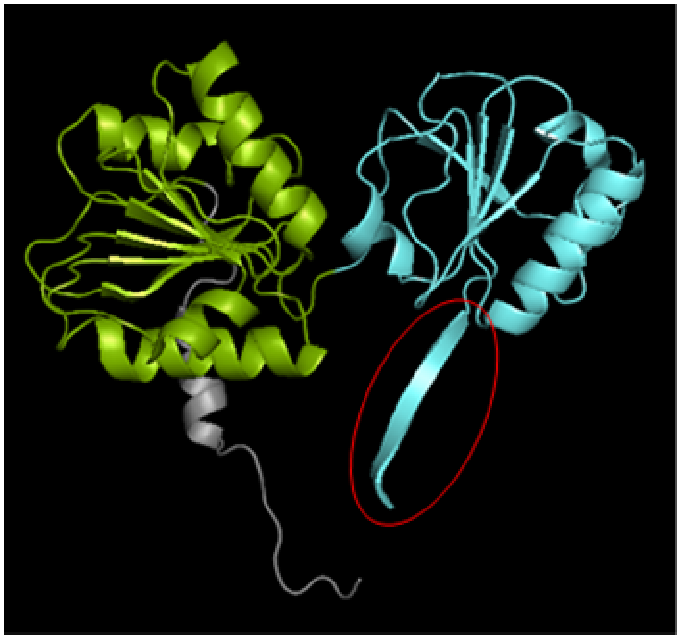
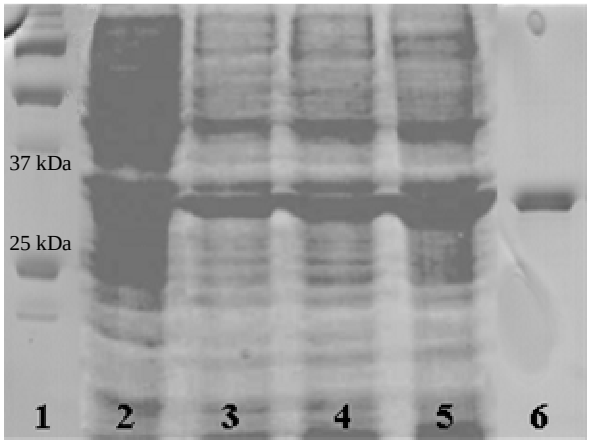


Figure 4.15: Aminoacids of c33 that have been removed in c33-7aa. The depicted structure is the portion of the NS3 helicase crystal structure (Protein Data Bank acc. no. 1A1V) comprised in c33 antigen

The construct was expressed in *E. coli* and purified with yield and purity comparable to the c33 antigen (Fig 4.16).
However, when tested on the Liaison® platform, c33-7aa showed no improvement over c33 antigen in terms of thermal stability (table 4.3 and figure 4.17).



Lane	Sample
1	Biorad Precision Plus® protein standards
2	Crude extract of *E. coli* BL21(DE3) pET30-c33-7aa, before induction
3	Crude extract of *E. coli* BL21(DE3) pET30-c33-7aa, 1h after induction
4	Crude extract of *E. coli* BL21(DE3) pET30-c33-7aa, 2h after induction
5	Crude extract of *E. coli* BL21(DE3) pET30-c33-7aa, 3h after induction
6	Purified c33-7aa

Figure 4.16: SDS-page of c33-7aa: crude extracts before and after induction and final purified product

Sample	**Antigen**			
	c33		c33-7aa	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 6	**18905**	**8015**	**20782**	**8048**
POS 7	**41930**	**19357**	**41296**	**21102**
POS 8	**95765**	**40677**	**92985**	**32530**
POS 10	**50218**	**21857**	**56787**	**17986**
NEG 3	**1011**	**1183**	**537**	**521**
NEG 4	**3341**	**3249**	**2826**	**1936**

Table 4.4: RLU values of c33 and c33-7aa before and after thermal stress

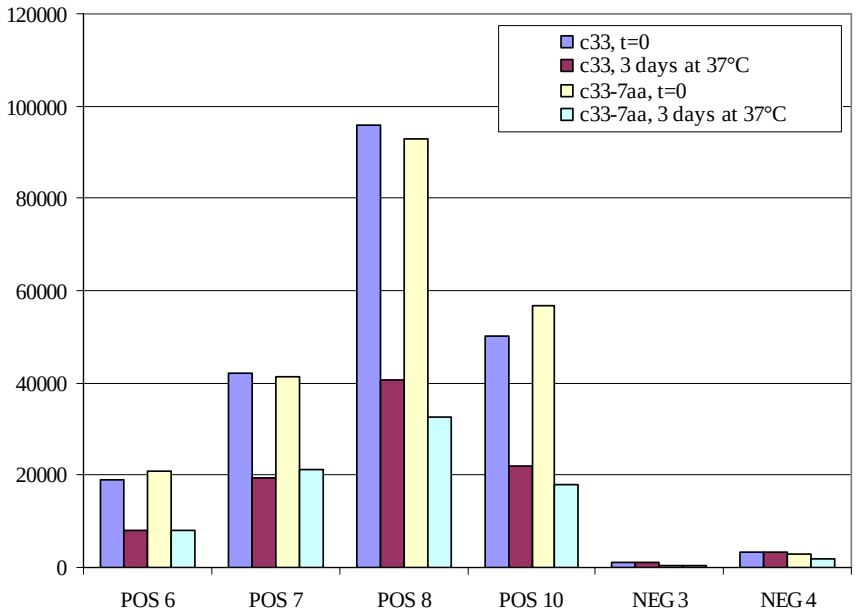


Figure 4.17: RLU values of c33 and c33-7aa before and after thermal stress

4.4.3. Limiting degrees of freedom: FKBP12-c33

FKBP12 is an abundantly expressed cytoplasmatic protein common to many eukaryotes. It is a 11.8-kDa immunophilin to which the immunosuppressant drugs FK506 (tacrolimus) and rapamycin (sirolimus) bind with high affinity. Like many of these immunophilins, FKBP12 possesses cis–trans peptidyl-prolyl isomerase or rotamase activity [46]. Besides its interest as a target for drug discovery, human FKBP12 has been studied for its prolyl isomerase activity and its role as a folding helper. FKBP12 possesses an exposed flexible loop ("flap") that has been shown to be tolerant to the insertion of entire domains without compromising the overall folding [55]. To exploit this property, human FKBP12 was used to act as a scaffold for c33 insertion, with the aim of reducing the conformational freedom of both the N-terminal and the C-terminal ends of c33. This way, both the ends of c33 should be constrained to a more ordered structure, avoiding the possible detrimental effects of loose ends on protein stability.

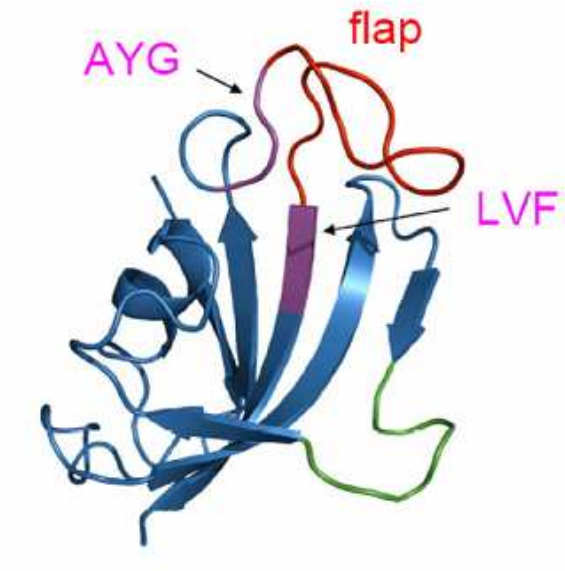
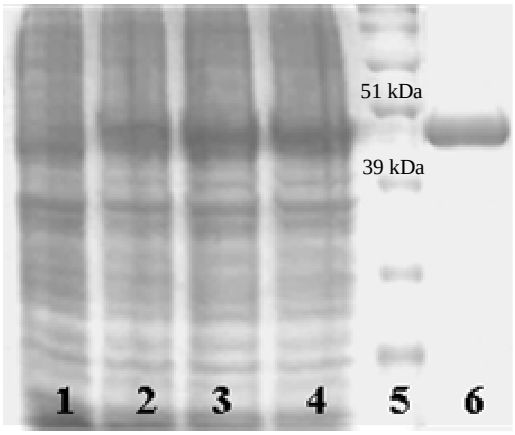


Figure 4.18: structure of human FKBP12 (taken from [55])

Therefore, c33 sequence was inserted in the same loop that was used in Knappe *et al*. [55] for the insertion of the IF domain of SlyD. Cloning was performed as described in materials and methods.. Expression of protein in *E. coli* resulted in a soluble and abundant product (figure 4.19) with a monomeric quaternary structure. Protein yield was 13 mg per gram of wet cell weight, more than two-fold compared to the original c33 antigen.



Lane	Sample
1	Biorad Precision Plus® protein standards
2	Crude extract of *E. coli* BL21(DE3) pET24-FKBP12-c33, before induction
3	Crude extract of *E. coli* BL21(DE3) pET24-FKBP12-c33, 1h after induction
4	Crude extract of *E. coli* BL21(DE3) pET24-FKBP12-c33, 2h after induction
5	Crude extract of *E. coli* BL21(DE3) pET24-FKBP12-c33, 3h after induction
6	Purified FKBP12-c33

Figure 4.19: SDS-page of FKBP12-c33: crude extracts before and after induction and final purified product

Unfortunately, antigen behaviour on the Liaison® platform (table 4.5; fig. 4.20) was similar to c33, exhibiting a sharp decrease in reactivity after storage at 37°C.

Sample	Antigen			
	c33		FKBP12-c33	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 11	**10879**	**1639**	**9136**	**1835**
POS 12	**26076**	**2620**	**23009**	**2451**
POS 13	**50129**	**4486**	**55996**	**3537**
POS 14	**19248**	**3266**	**25277**	**2900**
POS 15	**37745**	**2213**	**37420**	**2371**
NEG 5	**603**	**1400**	**797**	**1402**
NEG 6	**888**	**1310**	**828**	**1420**

Table 4.5: RLU values of c33 and FKBP12-c33 before and after thermal stress

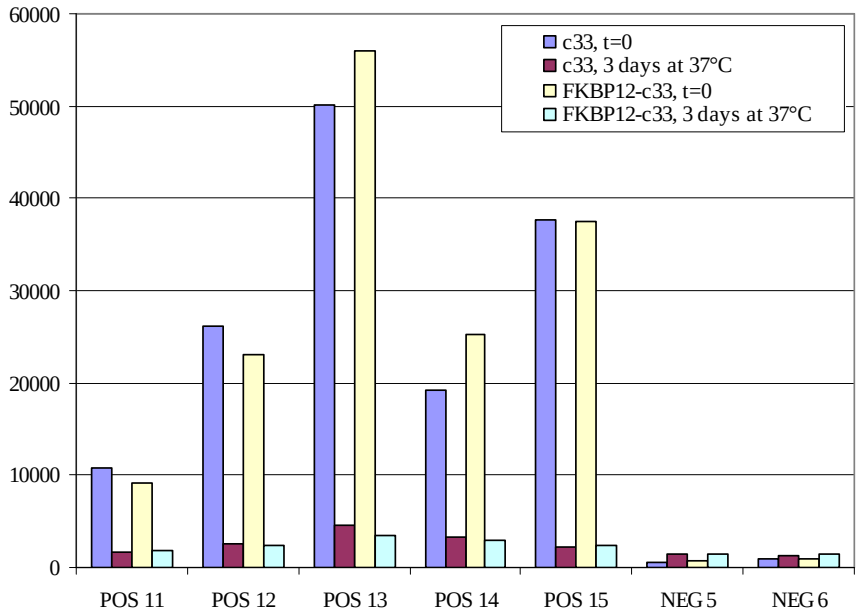


Figure 4.20: RLU values of c33 and FKBP12-c33 before and after thermal stress

4.4.4. Restoring subdomain integrity : c33eu

Another hypothesis on the causes of c33 instability was suggested by the sequence span of c33 antigen. As previously said, the sequence covers the complete subdomain 1 but only a minor part of subdomain 2. Thus, the organization of this truncated subdomain 2 into a folded structure (reflecting the fold of native subdomain 2) is not assured at all. Lack of a properly formed subdomain 2 could contribute to the overall protein stability and be responsible of the observed reactivity loss.

A new antigen, named c33eu was designed, with the same N-terminal end of c33, but with an extended sequence span to comprehend the entire subdomain 2. As in Gesell *et al.* [32], the extended β-loop was replaced with a short turn sequence (SDGK) to improve solubility. The SDGK sequence was chosen to mimic the most common β-turn motif of the NS3 protein, which had been found through a structural analysis. As a control, the corresponding sequence without deletion of the β-loop (which will be referred as c33 2D) was also cloned. The sequence span of c33 2D and c33eu, highlighted on the whole NS3 helicase structure (Protein Data Bank acc. no. 1A1V) is shown in figure 4.21 and 4.22.

The two proteins were expressed and successfully purified in the usual conditions (c33eu induction and purified product is shown in figure 4.23), but c33eu showed a higher solubility than c33 2D as expected.

Figure 4.21: Portion of the NS3 helicase crystal structure (Protein Data Bank acc. no. 1A1V) corresponding to the sequence span of c33 2D.

c33 2D in fact showed a marked tendency towards aggregation, and addition of at least 4M urea to the purification and storage buffers was necessary to avoid precipitation. As a result, c33 2D was discarded. Purification yield of c33eu was around 3 mg per gram of wet cell weight, a slightly low value compared to other constructs.

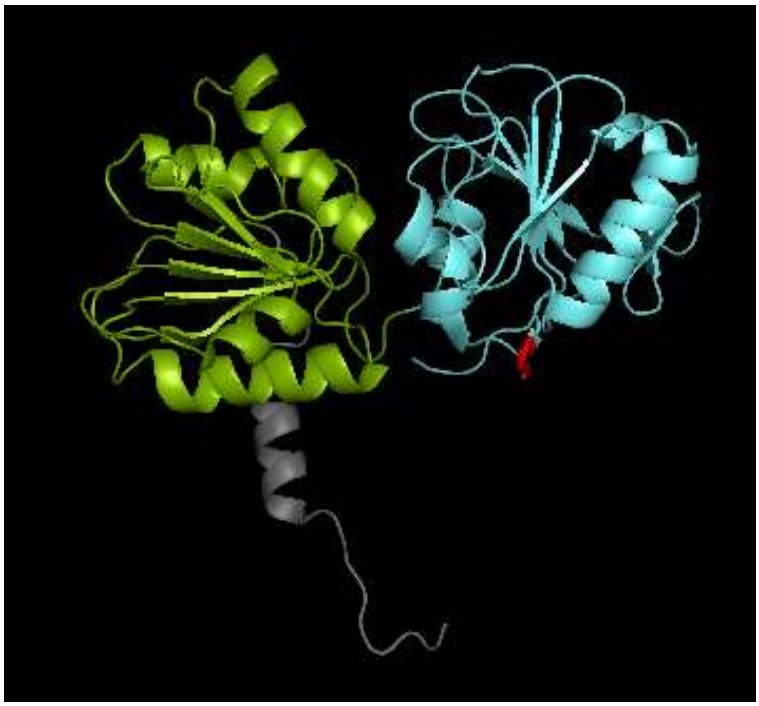
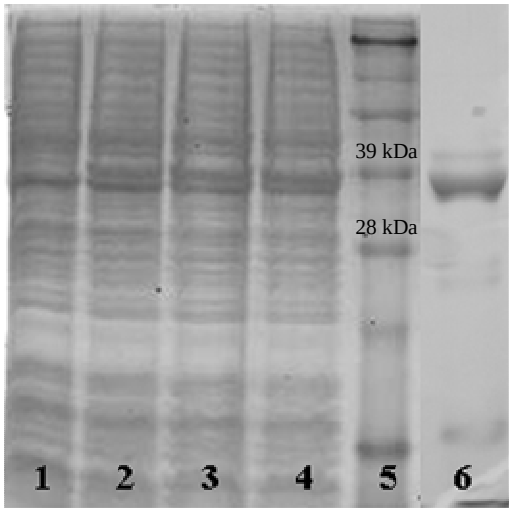


Figure 4.22: Portion of the NS3 helicase crystal structure (Protein Data Bank acc. no. 1A1V) corresponding to the sequence span of c33eu. The short turn replacing the extended β-loop is depicted in red.

Immunometric data obtained on Liaison® system (table 4.6; figure 4.24) showed for c33eu a slight increase in thermal stability, improving from an average 82% signal loss for c33 to an average 53% signal loss for c33eu after thermal stress on the tested positive samples.



Lane	Sample
1	Biorad Precision Plus® protein standards
2	Crude extract of *E. coli* BL21(DE3) pET30-c33eu, before induction
3	Crude extract of *E. coli* BL21(DE3) pET30-c33eu, 1h after induction
4	Crude extract of *E. coli* BL21(DE3) pET30-c33eu, 2h after induction
5	Crude extract of *E. coli* BL21(DE3) pET30-c33eu, 3h after induction
6	Purified c33eu

Figure 4.23: SDS-page of c33eu: crude extracts before and after induction and final purified product

Unfortunately, this lower instability of c33eu goes with a lower absolute signal level: the RLU values, while steadier, are substantially lower than the ones obtained with c33, especially for some samples. For example, positivity of PHV901-04 seroconversion point is completely missed.

Sample	Antigen			
	c33		c33eu	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 11	**105887**	**24324**	**68729**	**31796**
POS 12	**17132**	**4775**	**8733**	**4658**
PHV901-04	**32697**	**3902**	**4206**	**2996**
PHV901-07	**126747**	**22065**	**29731**	**5992**
PHV906-05	**57482**	**6420**	**8259**	**2126**
NEG 07	**652**	**716**	**705**	**674**
NEG 08	**1293**	**983**	**1120**	**1021**
NEG 09	**887**	**884**	**945**	**704**

Table 4.6: RLU values of c33 and c33eu before and after thermal stress

This lower reactivity could be attributed to the lack of the extended β-loop, which could contribute substantially to the immunogenicity of the molecule. However, these partial improvement hints that restoring the original domain integrity of NS3 helicase could be the right way to go to obtain a protein with a more stable antigenicity.

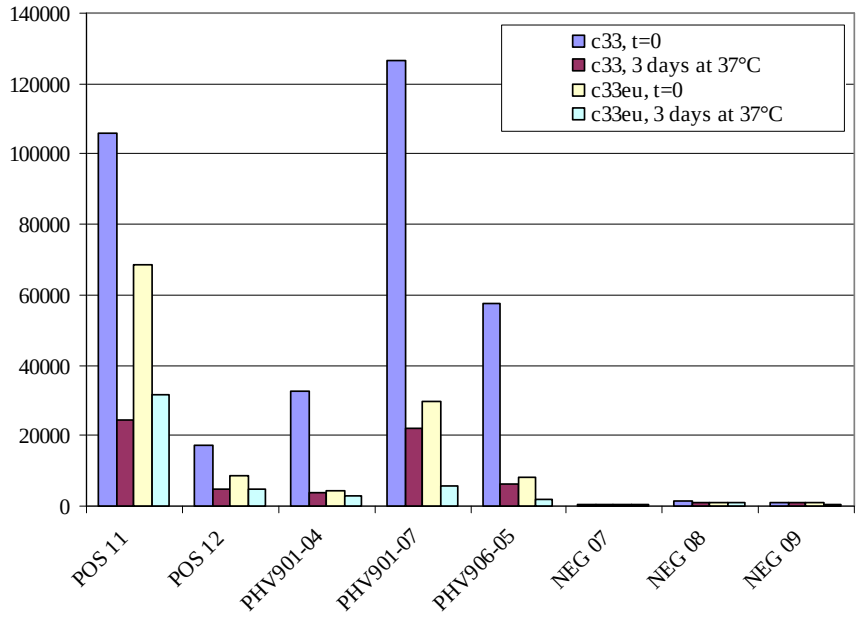


Figure 4.24: RLU values of c33 and c33eu before and after thermal stress

4.4.5. Restoring domain integrity: NS3 3D

Following the path traced with the c33eu antigen perspective, a new antigen was designed, with a markedly longer sequence, in an attempt to bring the c33 antigen back to a conformation better reflecting its natural form, applying to a greater extent the approach used for c33eu.
The NS3 3D antigen comprises all the NS3 protein except the protease domain. As a result, all 3 helicase subdomains are present in the complete form, hopefully recreating the situation happening in vivo where the helicase domain forms an almost self-sufficient region of NS3 protein, with minimal contact with the protease domain.
If plotted on the complete NS3 PDB structure, a rough representation of its structure could be the one depicted in fig. 4.25 (aminoacids of c33 on NS3 structure are reported on the right for comparison)

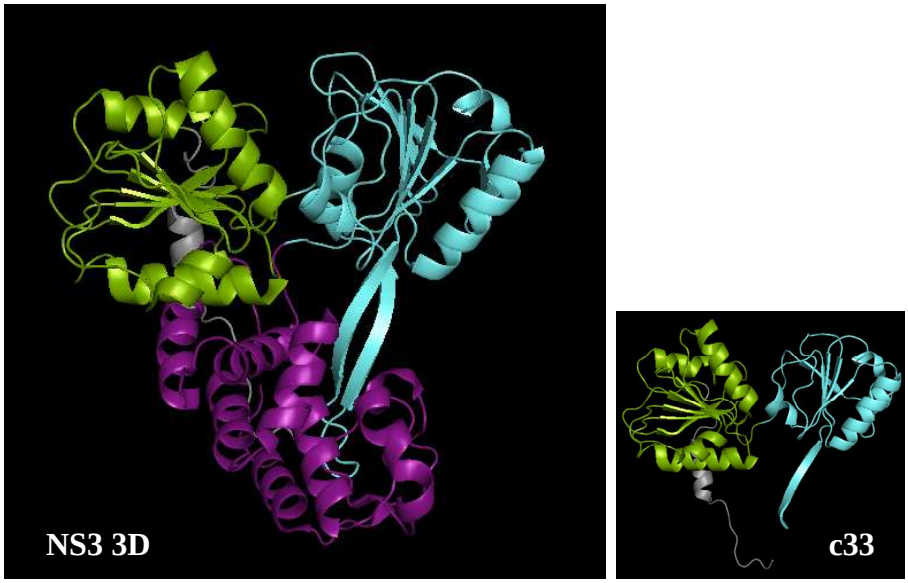


Figure 4.25: Portion of the NS3 helicase crystal structure (Protein Data Bank acc. no. 1A1V) corresponding to the sequence span of NS3 3D and c33.

In c33, due to the absence of subdomain 3, the solvent exposure of the β-loop, and the incomplete subdomain 2, the structural integrity of the protein is at least partially compromised, possibly accounting for the observed instability. In NS3 3D, the solvent exposure of internal regions should be greatly reduced, and the additional degrees of freedom caused by the incomplete domain 2 of c33 should be reduced as well.
Furthermore, literature data on HCV drug discovery showed that the complete helicase domain can be obtained as a soluble protein in *E. coli* pET expression system [44].

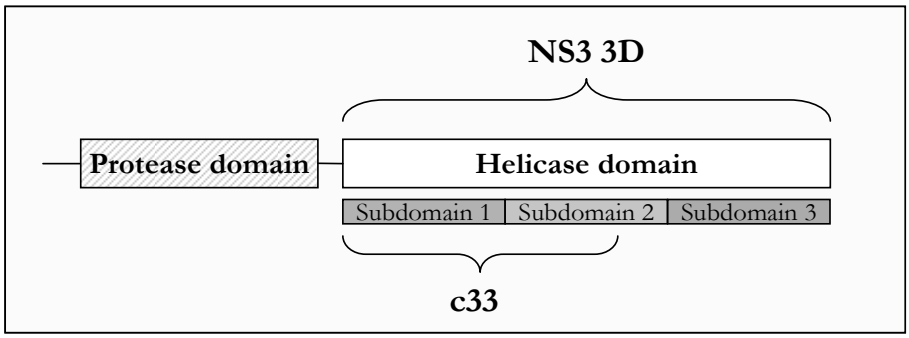
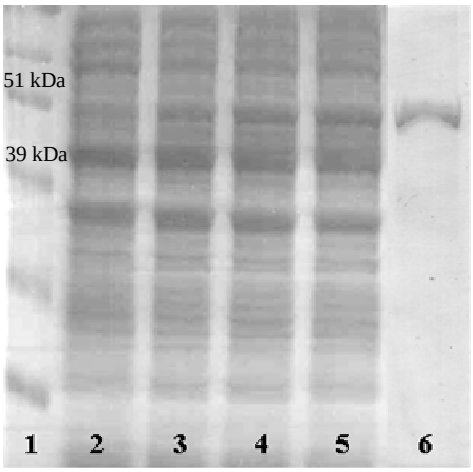


Figure 4.26: Schematic representation of the sequence span of NS3 3D construct in comparison with c33 and the entire NS3 protein.

The protein was successfully cloned, expressed and purified (figure 4.27) as reported in materials and methods, but with a substantially lower yield than the previous constructs, around 1.5 mg per gram of wet cell weight.



Lane	Sample
1	Biorad Precision Plus® protein standards
2	Crude extract of *E. coli* BL21(DE3) pET30-NS3 3D, before induction
3	Crude extract of *E. coli* BL21(DE3) pET30-NS3 3D, 1h after induction
4	Crude extract of *E. coli* BL21(DE3) pET30-NS3 3D, 2h after induction
5	Crude extract of *E. coli* BL21(DE3) pET30-NS3 3D, 3h after induction
6	Purified NS3 3D

Figure 4.27: SDS-page of NS3 3D: crude extracts before and after induction and final purified product

The first immunometric results obtained with NS3 3D on the Liaison® platform were encouraging (table 4.7, figure 4.28), as the initial RLU values measured before thermal stress were higher than the c33 values, showing an improvement in overall antigenicity of the construct. Thermal stress tolerance, though, was not satisfactory at all, as the initial gain in RLU is completely lost after incubation at 37°C.

Sample	**Antigen**			
	c33		NS3 3D	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 16	**105887**	**24324**	**122428**	**9487**
POS 17	**19880**	**3596**	**23892**	**1881**
POS 18	**32697**	**3902**	**43035**	**2280**
POS 19	**126747**	**22065**	**137077**	**11169**
NEG 2	**738**	**495**	**751**	**700**
NEG 5	**652**	**716**	**797**	**776**
NEG 6	**1293**	**983**	**1249**	**1196**

Table 4.7: RLU values of c33 and NS3 3D before and after thermal stress

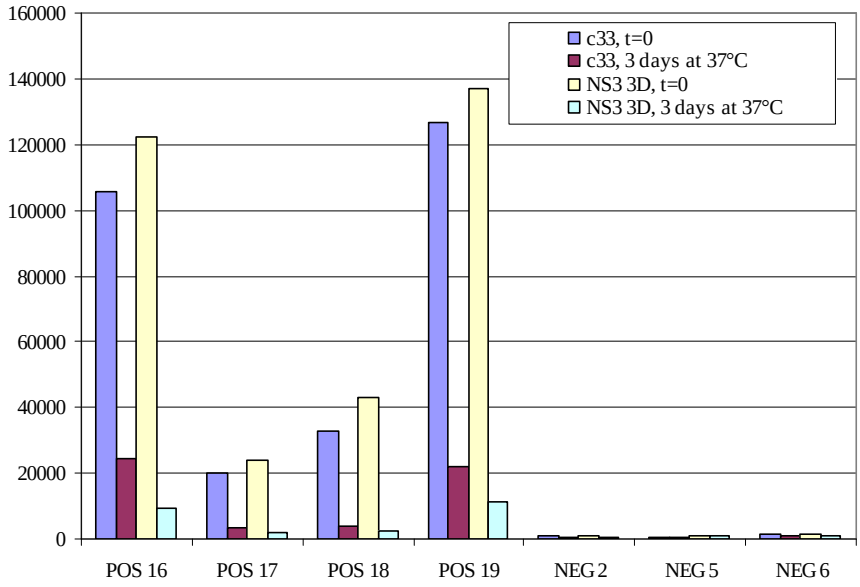


Figure 4.28: RLU values of c33 and NS3 3D before and after thermal stress

After this round of tests, c33eu and NS3 3D showed some kind of improvement over the original construct, c33eu for its slightly better thermal stability, NS3 3D because of its higher antigenicity. Therefore, continuation of research was focused on these two candidates.

With these two constructs, testing of a new coating protocol was made possible. The coating protocol recommended by the manufacturer of the beads is to be performed at 37°C for optimal coupling of the aminic groups of the ligand with the tosyl-activated beads. Coating at lower temperatures is allowed, provided that incubation time is increased enough, but still not the preferred condition. In previous experiments, coating was always performed at 4°C as a precautional measure because of the extreme thermal instability of the antigens. A high temperature during coating process could have led to a completely inactive antigen even before the thermal stress process. With the slight improvement in terms of total signal provided by NS3 3D and the improved thermal stability of c33eu, coating at 37°C could have become a viable alternative. For this reason, the performances of c33eu and NS3 3D were tested again after coating the paramagnetic beads overnight at 37°C instead of 4°C. c33 antigen (with coating at 37°C) was also tested as a control. The results are shown below (Table 4.8 and 4.9, fig. 4.29 and 4.30).

Sample	**Antigen**			
	c33		NS3 3D	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 11	**3744**	**1485**	**400015**	**65707**
POS 12	**1103**	**638**	**108276**	**11316**
POS 13	**2426**	**1419**	**369166**	**28156**
POS 14	**7867**	**2432**	**621315**	**124416**
POS 15	**664**	**481**	**842**	**708**
NEG 5	**702**	**538**	**1089**	**1259**
NEG 6	**1119**	**852**	**1259**	**1558**

Table 4.8: RLU values of c33 and NS3 3D, coated on the beads at 37°C, before and after thermal stress

Sample	**Antigen**			
	c33		c33eu	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 11	**3744**	**1485**	**16930**	**7802**
POS 12	**1103**	**638**	**33872**	**58466**
POS 13	**2426**	**1419**	**2272**	**3006**
POS 14	**7867**	**2432**	**6755**	**2396**
POS 15	**664**	**4801**	**517**	**510**
NEG 5	**702**	**538**	**1103**	**1091**
NEG 6	**1119**	**852**	**752**	**880**

Table 4.9: RLU values of c33 and c33eu, coated on the beads at 37°C, before and after thermal stress

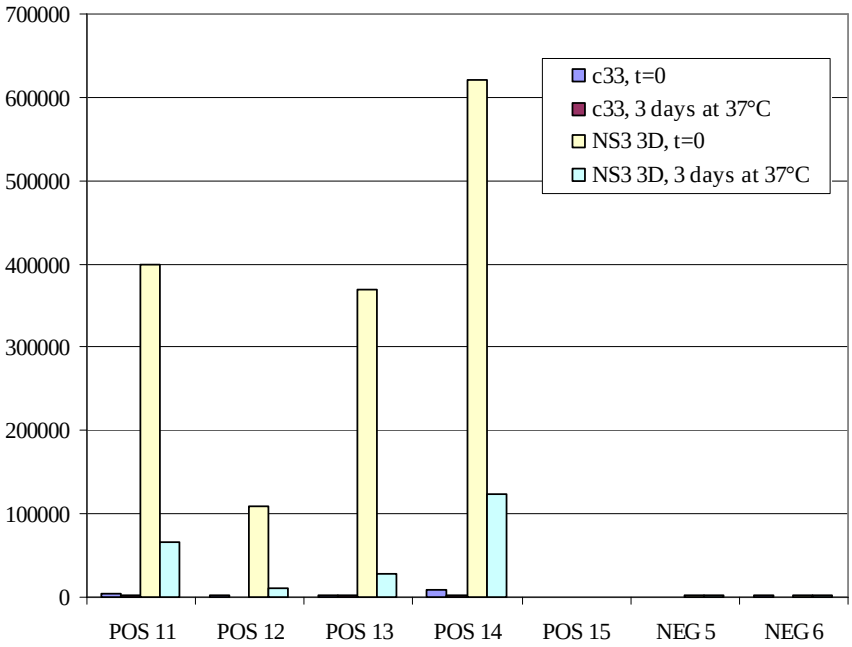


Figure 4.29 RLU values of c33 and NS3 3D, coated on the beads at 37°C, before and after thermal stress

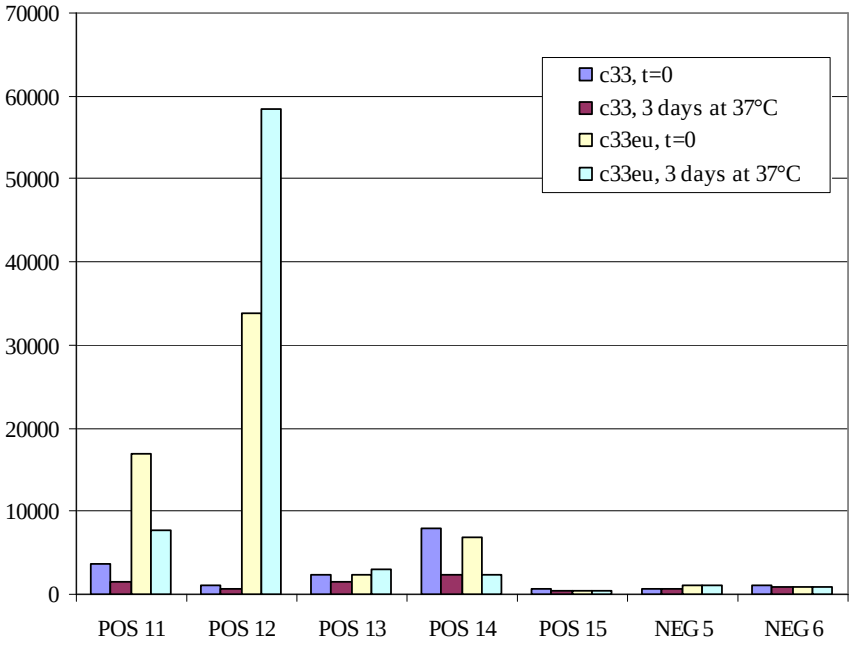


Figure 4.30: RLU values of c33 and c33eu, coated on the beads at 37°C, before and after thermal stress

Surprisingly, NS3 3D initial reactivity is vastly improved when the coating is performed at 37°C, while c33 reactivity is completely abolished. c33eu shows an intermediate behaviour, with a much lower signal on seroconversion points and discordant data on positive samples. The magnitude of signal increase of NS3 3D is at least 3-fold for all the tested samples. However, signal stability after thermal stress doesn't look affected by the coating temperature, with the RLU values dropping to intolerable values for all the three antigens.

The high RLUs detected on NS3 3D coated beads after coating at 37°C are in striking contrast to the complete inactivation of c33 antigen. If such a high signal can be detected after coating at 37°C, either the NS3 3D protein is stable enough to endure the overnight incubation at 37°C in the conditions provided by the coating solution, or the reactivity of NS3 3D is high enough to show high RLU values still after the drastic drop in reactivity that should happen during the overnight coating at 37°C.
It should be noted that after the coating process, the buffer in which the microbeads are suspended is exchanged with a storage buffer with a different composition. It could be supposed that NS3 3D coated microbeads are resistant to thermal stress when stored in the coating buffer, and that the drop in reactivity happens because of the addition of the storage buffer.
To verify this possibility, a new test was performed. Paramagnetic microbeads were coated with NS3 3D or c33 antigen overnight at 37°C, then washed and resuspended in fresh coating buffer instead of the usual storage buffer. Immunochemical data obtained on the Liaison® platform with these beads were absolutely surprising and are shown in table 4.10 and figure 4.31.

Sample	**Antigen**			
	NS3 3D (storage buffer)		NS3 3D (coating buffer)	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 21	**176133**	**27842**	**270718**	**228810**
POS 22	**36423**	**5073**	**63070**	**66677**
POS 23	**18219**	**3962**	**33950**	**37553**
POS 24	**7983**	**1545**	**12639**	**12923**
POS 25	**14195**	**7264**	**18261**	**18684**
NEG 7	**2360**	**1456**	**2197**	**1507**
NEG 8	**935**	**669**	**938**	**611**

Table 4.10: RLU values NS3 3D, coated on the beads at 37°C, before and after thermal stress, performed in the usual storage buffer or in fresh coating buffer

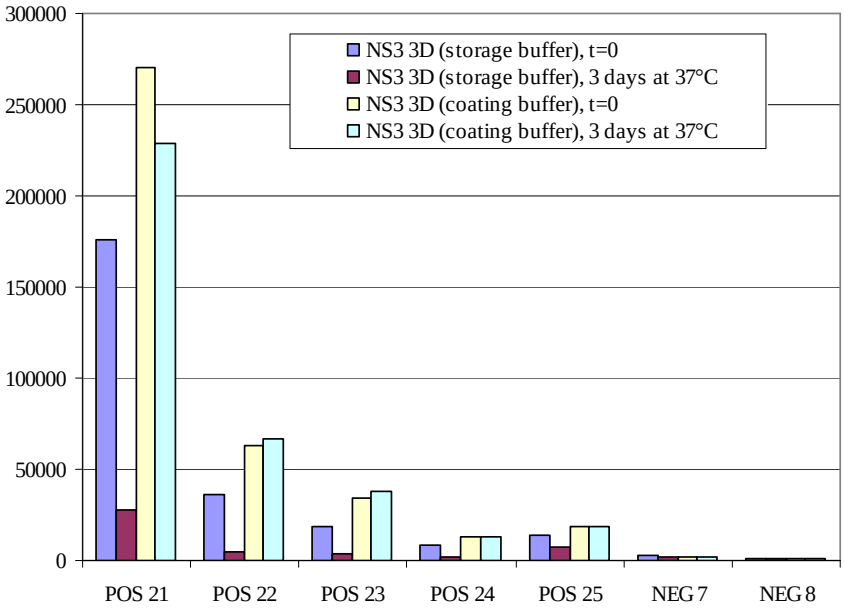


Figure 4.31: RLU values NS3 3D, coated on the beads at 37°C, before and after thermal stress, performed in the usual storage buffer or in fresh coating buffer

These data clearly show that the thermal stability of the NS3 3D-coated beads is clearly improved if the buffer used for coating is maintained also during storage and thermal stress. In these conditions, loss of reactivity after thermal stress is negligible. However, the stabilizing effect of coating buffer is limited to NS3 3D, as was shown in a parallel experiment where c33 antigen (coated at 4°C to maintain initial reactivity) was not stabilized in the same conditions (table 4.11, figure 4.32).

Sample	**Antigen**			
	c33 (storage buffer)		c33 (coating buffer)	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 26	**31790**	**1563**	**36881**	**2011**
POS 27	**5541**	**567**	**7617**	**1054**
POS 28	**4997**	**904**	**6808**	**877**
POS 29	**1276**	**1979**	**1542**	**1124**
NEG 9	**522**	**937**	**659**	**987**

Table 4.11: RLU values c33, coated on the beads at 4°C, before and after thermal stress, performed in the usual storage buffer or in fresh coating buffer

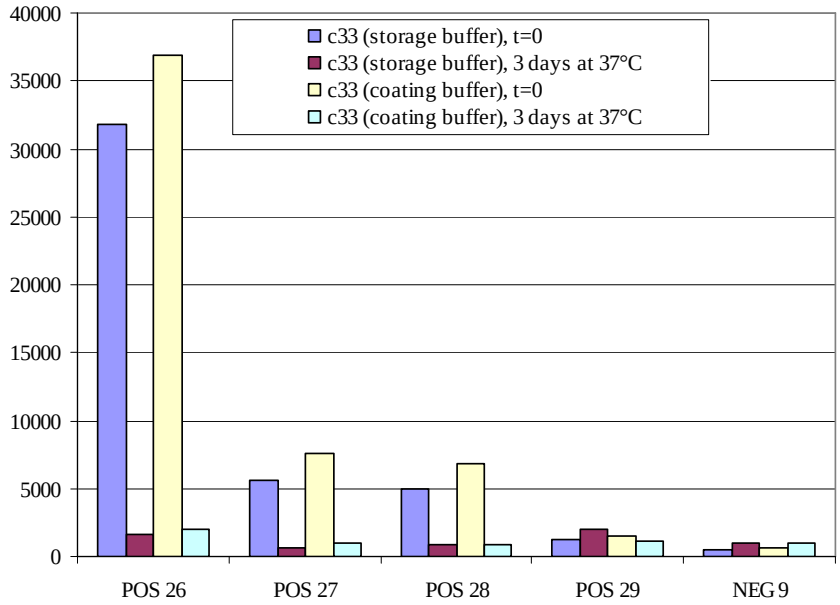


Figure 4.32: RLU values c33, coated on the beads at 4°C, before and after thermal stress, performed in the usual storage buffer or in fresh coating buffer

Taken altogether, these data allow to state that:

a) the NS3 3D antigen is resistant to thermal stress at 37°C, provided that the beads are maintained in coating buffer.
b) this stabilization effect by the coating buffer is specific for NS3 3D, as in the same conditions c33 is not stabilized
c) NS3 3D has an overall higher reactivity than c33 in every tested condition.
d) coating of NS3 3D at 37°C is not only tolerated by the antigen, but also markedly improves its performances.

These conclusions clearly show that NS3 3D is a neat improvement over c33 for application in the Liaison® platform. Therefore, NS3 3D was chosen among all the tested constructs for further development and optimization.

4.5. NS3 3D optimization

4.5.1. Fine tuning of the storage buffer

The critical effect of the coated microbeads storage buffer on thermal stability was discovered in the previous experiments. More precisely, coating buffer conferred thermal stress resistance to NS3 3D where the original storage buffer did not. Composition of the two buffers is confronted below (table 4.12).

Coating buffer	
Inorganic buffering agent	-
Casein	0,1 g/L
TCEP	1,43 g/L
Pluronic F-127	0,025%
EDTA	0,83 g/L
pH	6.4

Storage buffer	
MES	1.952 g/L
NaCl	11.7 g/L
Casein	0,1 g/L
TCEP	1,43 g/L
Glycerol	10 %
NaN_3	1 g/L
Tween 20	0,5 %
Ethylene Glycol	5%
pH	7

Table 4.12: comparison of composition of coating buffer and storage buffer

One of the most eye-catching differences in buffer compositions is the buffering agent, which is MES (2-(N-morpholino)ethanesulfonic acid) in the storage buffer and an undisclosed inorganic buffering agent in the coating buffer. A test was made to check if this discrepancy is responsible of the different effect of the two buffers on NS3 3D stability.
After an overnight coating of NS3 3D at 37°C, the microbeads were washed and resuspended in fresh coating buffer, as done before, or in another buffer, identical to the first except for the inorganic buffering agent, which was replaced with MES.

Comparison of the results obtained in the Liaison® system clearly show that the microbeads stored in a solution buffered with the inorganic compound maintain reactivity after thermal stress whereas the microbeads stored in a MES-buffered solution do not. The reactivity loss of the MES-buffered microbeads after thermal stress is absolutely comparable to the one happening when the beads are stressed in the original storage buffer.

Sample	Antigen			
	NS3 3D (coating buffer w/ MES)		NS3 3D (coating buffer)	
	t=0	3 days at 37°C	t=0	3 days at 37°C
POS 30	**201064**	**36117**	**197237**	**201589**
POS 31	**39716**	**6311**	**40103**	**46127**
POS 32	**23995**	**4548**	**26460**	**23670**
POS 33	**11576**	**5029**	**10074**	**6890**
POS 34	**163441**	**16824**	**150886**	**121701**
NEG 35	**693**	**669**	**701**	**567**
NEG 36	**944**	**854**	**895**	**762**

Table 4.13: effect of the replacement of inorganic buffering agent with MES in the coating buffer (used as a storage buffer): comparison of RLU values of NS3 3D, before and after thermal stress.

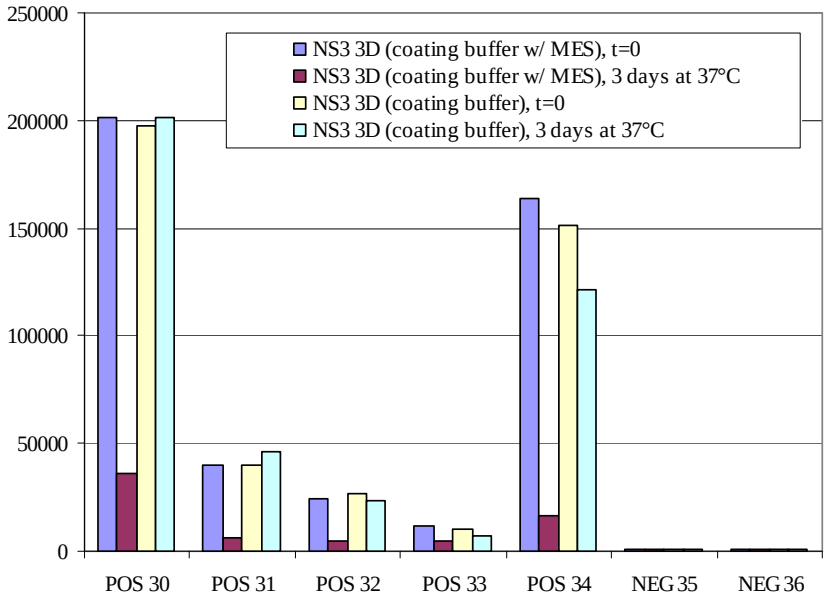


Figure 4.33: effect of the replacement of inorganic buffering agent with MES in the coating buffer (used as a storage buffer): comparison of RLU values of NS3 3D, before and after thermal stress.

This experiment clearly shows that the presence of the inorganic buffering agent instead of MES in the buffer where the NS3 3D coated microbeads are stored is the critical factor determining the different behaviour of the two analyzed buffers. As the coating buffer performed so well, it was decided to use it also as a storage buffer, with just some little adjustments such as adding standard preservatives and detergents that are not disclosed here.

These added components, however, were still confirmed to be ininfluent in terms of thermal stress stability (data not shown).

4.5.2. Crossreactivity analysis and S-tag

The addition of such a long aminoacidic sequence to the C-terminal end of c33 to obtain NS3 3D generated an antigen with definitely better performances in terms of reactivity and stability. However, it also raised concerns on the possible inclusion of crossreacting and/or interfering sequences that were previously not included in the well-characterized c33 antigen. A preliminary sequence similarity search with the BLAST algorithm was performed on c33 and NS3 3D sequences to find potential similar motifs that could lead to crossreactivity of the antigen with antibodies unrelated to HCV. Significant sequence similarity was found only in the GB virus C polyprotein, a fact that was already known for c33. This finding did not raise particular concerns, because the similarity is stronger in the N-terminal part of NS3 3D, which is common to c33, and weaker in the C-terminal part, exclusively owned by NS3 3D.

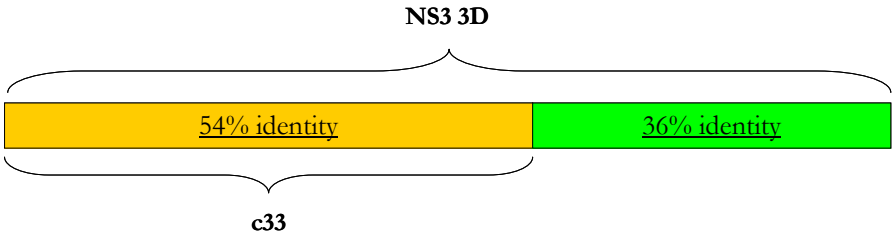


Figure 4.33: percentage of sequence identity between NS3 and GB virus C polyprotein in the NS3 helicase regions covered by c33 or NS3 3D

Nevertheless, NS3 3D was tested with the Liaison® system on an open population of 128 sera obtained from the hospital of Chivasso, looking for potential interfering reactivity. The related data can be found in Table 4.14.
Of these 128 sera, 5 gave RLU values above an arbitrary threshold set to 5000 (which is a low RLU value indeed for the Liaison® system), and clearly out of the general distribution. These 5 sera were retested with the Ortho HCV SAVe 3.0 ELISA kit and all tested positive for anti-HCV antibodies. Therefore, these five samples were defined as positive and separated from the collection. The remaining 122 were considered a population of negative samples. The RLU distribution of this population approximates a normal distribution and is reported in figure 4.34. The average value of this negative population was 924 RLU, with a standard deviation of 449 RLU. These numbers are well in line with the values normally recorded on other indirect tests of the Liaison® platform and should be considered very satisfying (positive clinical samples usually give

RLU values of several thousands). Neither unspecific reactivity on isolate samples, nor high background values were found in the analyzed population.

Sample	RLU	**Sample**	RLU	**Sample**	RLU
CHIVASSO18	519032	CHIVASSO61	645	CHIVASSO104	742
CHIVASSO19	835	CHIVASSO62	636	CHIVASSO105	3044
CHIVASSO20	950	CHIVASSO63	572	CHIVASSO106	738
CHIVASSO21	762	CHIVASSO64	729	CHIVASSO107	738
CHIVASSO22	999	CHIVASSO65	630	CHIVASSO108	805
CHIVASSO23	535	CHIVASSO66	673	CHIVASSO109	1362
CHIVASSO24	597	CHIVASSO67	1600	CHIVASSO110	892
CHIVASSO25	2386	CHIVASSO68	726	CHIVASSO111	1060
CHIVASSO26	724	CHIVASSO69	1608	CHIVASSO112	1818
CHIVASSO27	811	CHIVASSO70	757	CHIVASSO113	820
CHIVASSO28	214996	CHIVASSO71	2306	CHIVASSO114	675
CHIVASSO29	850	CHIVASSO72	577	CHIVASSO115	1166
CHIVASSO30	994	CHIVASSO73	793	CHIVASSO116	1010
CHIVASSO31	525	CHIVASSO74	556	CHIVASSO117	742
CHIVASSO32	709	CHIVASSO75	652	CHIVASSO118	999
CHIVASSO33	736	CHIVASSO76	919	CHIVASSO119	882
CHIVASSO34	512	CHIVASSO77	838	CHIVASSO120	1201
CHIVASSO35	579	CHIVASSO78	865	CHIVASSO121	777
CHIVASSO36	1141	CHIVASSO79	888	CHIVASSO122	762
CHIVASSO37	812	CHIVASSO80	1264	CHIVASSO123	678
CHIVASSO38	550	CHIVASSO81	1334	CHIVASSO124	748
CHIVASSO39	553	CHIVASSO82	604	CHIVASSO125	2779
CHIVASSO40	938	CHIVASSO83	895	CHIVASSO126	1326
CHIVASSO41	1035	CHIVASSO84	843	CHIVASSO127	1244
CHIVASSO42	196088	CHIVASSO85	12825	CHIVASSO128	1080
CHIVASSO43	832	CHIVASSO86	795	CHIVASSO129	828
CHIVASSO44	907	CHIVASSO87	793	CHIVASSO130	1356
CHIVASSO45	520	CHIVASSO88	631	CHIVASSO131	3030
CHIVASSO46	747	CHIVASSO89	793	CHIVASSO132	1456

CHIVASSO47	568	CHIVASSO90	777	CHIVASSO133	1304
CHIVASSO48	1485	CHIVASSO91	993	CHIVASSO134	915
CHIVASSO49	754	CHIVASSO92	714	CHIVASSO135	847
CHIVASSO50	588	CHIVASSO93	614	CHIVASSO136	700
CHIVASSO51	673	CHIVASSO94	849	CHIVASSO137	774
CHIVASSO52	810	CHIVASSO95	1024	CHIVASSO138	1024
CHIVASSO53	652	CHIVASSO96	686	CHIVASSO139	756
CHIVASSO54	754	CHIVASSO97	658	CHIVASSO140	673
CHIVASSO55	386637	CHIVASSO98	820	CHIVASSO141	961
CHIVASSO56	772	CHIVASSO99	746	CHIVASSO142	718
CHIVASSO57	586	CHIVASSO100	1264	CHIVASSO143	625
CHIVASSO58	740	CHIVASSO101	702	CHIVASSO144	730
CHIVASSO59	642	CHIVASSO102	775	CHIVASSO145	990
CHIVASSO60	867	CHIVASSO103	898		

Table 4.14: RLU values obtained with NS3 3D on an open population of 128 sera from Chivasso hospital. The five positive samples are marked in red

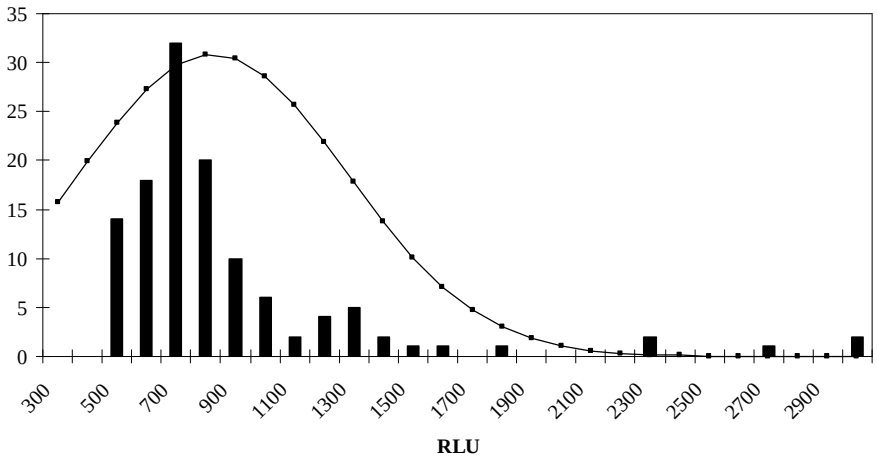


Figure 4.34: distribution of the RLU values obtained testing the 123 negative sera from Chivasso hospital with NS3 3D (bars). Theoretical normal distribution is shown with the dotted line

NS3 3D performances were also tested on a small group of interfering sera of HCV-negative with unspecific reactivity that had been isolated in the past during the development of other diagnostic kits. Assuming again the

arbitrary threshold of 5000 RLU, 5 of the 8 interfering sera (INTERF 1 to 5) showed high reactivity against NS3 3D (table 4.15).

Sample	RLU
INTERF 1	**24787**
INTERF 2	**208403**
INTERF 3	**1096296**
INTERF 4	**14632**
INTERF 5	**68381**
INTERF 6	**4478**
INTERF 7	**4375**
INTERF 8	**2731**

Table 4.15: RLU values obtained with NS3 3D on 8 known interfering HCV-negative sera

More experiments were planned to eliminate or reduce the nonspecific reactivity found against 5 of the 8 known interfering sera. The hypothesis was made that this unspecific reactivity was caused by the N-terminal extra sequence added by the vector pET30 and not belonging to the HCV proteome.
This sequence SSGLVPRGSGMKETAAAKFERQHMDSPDL consists mainly of the S-Tag and the thrombin recognition sequence. The S-Tag is a 15-aa peptide derived from bovine RNase A protein, but with complete sequence identity to its human counterpart. Formation of autoantibodies directed against endogenous RNase and DNase enzymes has been reported in individuals affected by autoimmune diseases[58]. It is possible that the unspecific reactivity of the 5 isolated interfering sera could be directed against the S-Tag. To verify this possibility, a version of NS3 3D without the extra N-terminal sequence was cloned and expressed (data not shown) as described in materials and methods and tested on the five interfering sera.

Sample	**Antigen**		
	NS3 3D	NS3 3D w/ scavenger	NS3 3D w/o extra sequence
INTERF 1	**24787**	**20174**	**12653**
INTERF 2	**208403**	**5401**	**2291**
INTERF 3	**1096296**	**7677**	**2699**
INTERF 4	**14632**	**11265**	**2282**
INTERF 5	**68381**	**3104**	**1083**
POS 40	**607524**	**617508**	**380477**

Table 4.16: RLU values obtained with NS3 3D, NS3 3D with a scavenger and NS3 3D without extra sequence on the 5 interfering HCV-negative sera and a positive control.

As an alternative, an unrelated protein (*E. coli* thioredoxin - as coded by an empty Novagen® pET32 vector) containing the same extra sequence of NS3 3D was spiked into the interfering samples to act as a scavenger.
If the unspecific reactivity had been directed against the N-terminal extra sequence, the putative antibodies responsible of it should have bound the excess of scavenger protein, resulting in the elimination of their ability to bind the immobilized NS3 3D.
The results obtained on the Liaison® system in the three different conditions are reported in table 4.15 and figure 4.35.

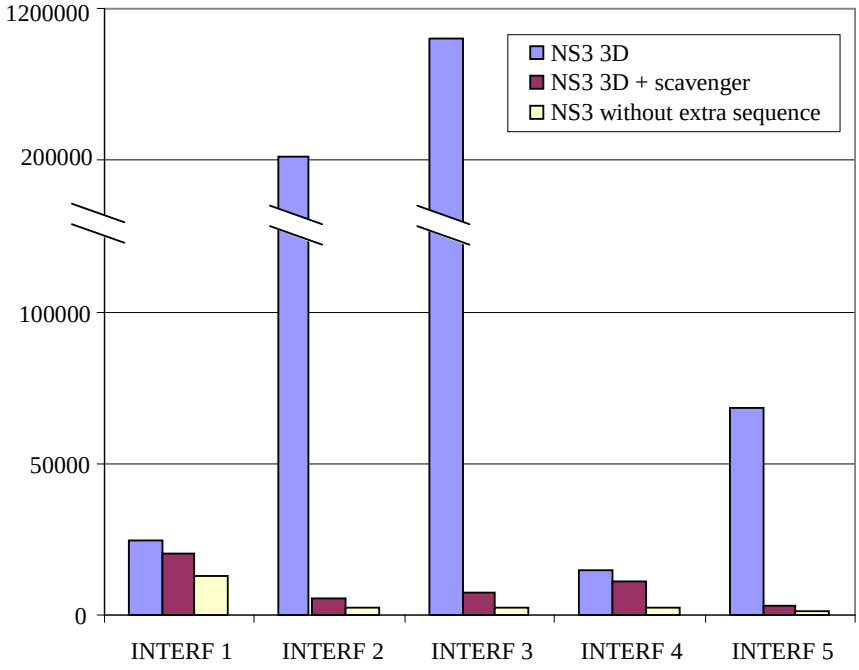


Figure 4.35: RLU values obtained with NS3 3D, NS3 3D with a scavenger and NS3 3D without extra sequence on the 5 interfering HCV-negative sera.

Both eliminating the extra sequence and adding a scavenger were effective in reducing the unspecific reactivity of the five interfering sera against NS3 3D. The signal reduction was dramatic in all but one case (INTERF 1), eliminating almost completely the concern caused by the interfering sera and confirming that the unspecific reactivity of these 5 sera is caused by the presence of antibodies directed against the extra N-terminal sequence. The positive control shows that while adding a scavenger does not reduce the RLU signal against a true positive, removing the extra sequence generates an antigen with lower reactivity also against positive samples. The causes of this phenomenon were not investigated but could be tentatively attributed to a lower solubility of the protein without the extra sequence. Due to these

factors, the addition of a scavenger protein to the sample diluent in the diagnostic kit should be the easiest way to cope with these interference problems.

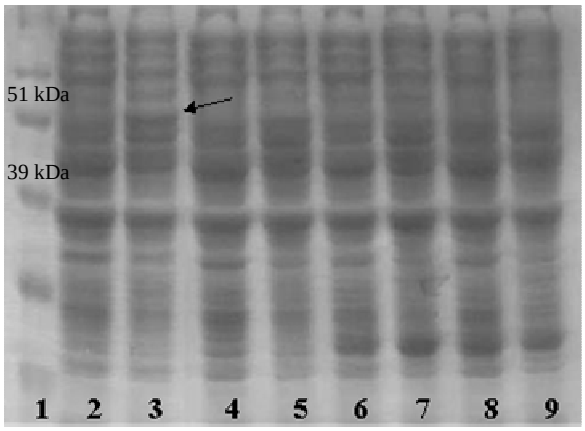
4.5.3. Codon adaptation and expression strains

NS3 3D was demonstrated to be a clear improvement over c33 for its higher reactivity and thermal stability. However, the protein yields obtained in the used model conditions (flask culture, growth at 30°C, induction for 3 hours with 1 mM IPTG) were not exciting. Yield of NS3 3D in two separate expression and purification batches was 0.4 and 0.6 mg of protein per gram of wet cell weight.

Optimization of the fermentation process would have required more time and resources and was not in the scope of this work. To investigate the possibility of raising the expression level of NS3 3D, experiments on different expression strains and codon adaptation were performed. The results will be now briefly reviewed.

Meddling with the expression strain was not productive; *E. coli* OverExpress™ C41, C43, C41pLys and C43pLys strains from Lucigen, isolated specifically for production of difficult to express proteins[69] were considered as an alternative to the *E. coli* BL21(DE3) expression system used in this work. Unfortunately, expression of NS3 3D in these strains resulted in an even lower expression of the protein, which was hardly visible on SDS-page as can be seen in figure 4.36.

Better results were obtained by improving codon adaptation. Poor codon adaptation of heterologous sequences to the host translational apparatus has often been reported as a cause for low expression of foreign proteins[8][114]. The Codon Adaptation Index (CAI) is a value that defines the fitness of a nucleic acid sequence to the codon bias of its host. The CAI value is computed by giving a weight (representing relative adaptiveness) to each codon, calculated from its frequency within a chosen small pool of highly expressed genes, to form a codon usage table. Variations in the calculated CAI are dependent on the set of sequences used to generate the codon usage table. The original NS3 3D sequence, obtained by PCR from a larger template, was analyzed with different CAI determination tools (reported in table 4.17). All of them gave low CAI values, reported below.



Lane	Sample
1	Invitrogen® SeeBlue Plus 2
2	Crude extract of *E. coli* C41 pET30-NS3 3D, before induction
3	Crude extract of *E. coli* C41 pET30-NS3 3D, 3h after induction
4	Crude extract of *E. coli* C43 pET30-NS3 3D, before induction
5	Crude extract of *E. coli* C43 pET30-NS3 3D, 3h after induction
6	Crude extract of *E. coli* C41pLys pET30-NS3 3D, before induction
7	Crude extract of *E. coli* C41pLys pET30-NS3 3D, 3h after induction
8	Crude extract of *E. coli* C43pLys pET30-NS3 3D, before induction
9	Crude extract of *E. coli* C43pLys pET30-NS3 3D, 3h after induction

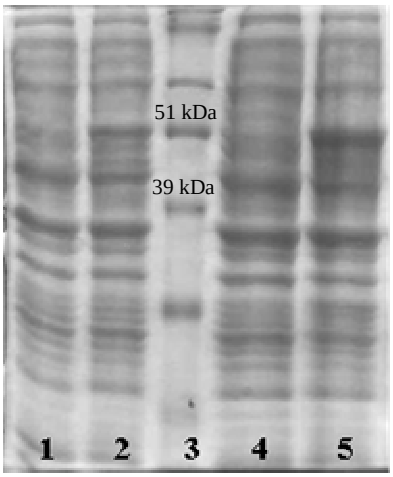
Figure 4.36: SDS-page of NS3 3D expression in Lucigen™ strains: crude extracts before and after induction.

Thus, a synthetic NS3 3D gene optimized with the GeneOptimizer® software tool was ordered to GENEART AG. This optimization software is designed to improve codon adaptation and remove sequence motifs (repetitive sequences, RNA destabilizing elements) that could negatively impact on the protein expression process. The CAI of this optimized gene, calculated with the same tools is reported for comparison (table 4.17).

Algorithm	**CAI** (pre-optimization)	**CAI** (post-optimization)
EMBOSS.cai[84]	**0,442**	**0,766**
OPTIMIZER[81]	**0,439**	**0,759**
Jcat[39]	**0,351**	**0,589**
GeneOptimizer[82]	**N/A**	**1,000**

Table 4.17: Codon Adaptation Index of the NS3 3D coding sequence before and after optimization with GeneOptimizer[82] (GENEART AG, Regensburg, Germany)

BL21(DE3) cells were transformed with the optimized gene and expression and purification were performed in the same conditions used for the original NS3 3D gene. As can be clearly seen in the SDS-page, the strain carrying the optimized NS3 3D sequence accumulates more NS3 3D protein than the original one during fermentation (figure 4.37).



Lane	Sample
1	Crude extract of BL21(DE3) pET30-NS3 3D (non-optimized), before induction
2	Crude extract of BL21(DE3) pET30-NS3 3D (non-optimized), 3h after induction
3	Invitrogen® SeeBlue Plus 2
4	Crude extract of BL21(DE3) pET30-NS3 3D (optimized), before induction
5	Crude extract of BL21(DE3) pET30-NS3 3D (optimized), 3h after induction

Figure 4.37: SDS-page of NS3 3D expression in Lucigen™ strains: crude extracts before and after induction.

The final yield of the two expression and purification batches performed with the optimized NS3 3D gene were 1.6 and 1.7 mg of protein per gram of wet cell weight, with a more than three-fold improvement in protein production.

5. Discussion

High prevalence of the hepatitis C virus and severe manifestations of infection clearly explain the need of reliable diagnostic instruments for HCV detection. This is particularly true in developing countries, where prevalence of HCV in the population is even higher. Due to the massive presence of various bioreagents, often diagnostic kits must be considered heat-labile products. For these kinds of products, conditions of shipping and delivery are a delicate issue and must be considered with attention, lest inactivation of the product occurs. Transportation at controlled temperature is not always possible, due to logistic and economic issues, and even when shipping at controlled temperature, it is sometimes difficult to obtain complete tracking of the cold chain, especially in countries where logistics is challenging and climatic conditions are adverse (warm countries). Therefore, when designing diagnostic devices, particular care should be taken in reducing susceptibility of the product to heat stress. A kit with high resistance to thermal stress can tolerate accidental exposure to non-controlled temperatures and is easier to store for the end-user. Moreover, it requires a less expensive shipping and should also exhibit a longer shelf-life.

Most of the automated kits for anti-HCV antibodies detection currently on the market must be stored in a refrigerated environment, as recommended by the manufacturer. As an example, Prism Anti-HCV by Abbott, Elecsys Anti-HCV by Roche and Access HCV Ab PLUS by Bio-Rad must all be stored (at least some components) at a temperature of 2-8°C, according to the related user manuals. Thus, an anti-HCV diagnostic kit able to stand transportation and storage at non-controlled temperature could have a competitive edge over the more heat-labile competitors. In this scenery, the results described in this thesis can represent a valuable contribution towards the realization of an anti-HCV assay with improved thermal stability.

The process of adaptation of c33 antigen from ELISA to CLIA immunoassay format turned essentially into solving an emerging thermal stability issue. A new antigen, NS3 3D, was developed with improved thermal stability. This antigen showed good tolerance to thermal stress, as its diagnostic performances were unaffected by a 3-days incubation at 37°C, meeting the target set at the beginning of the project. Further tests performed in our labs explored various stress conditions and determined that the thermal tolerance on NS3 3D still allows clear discrimination of negative and positive samples after 1 day at 45°C, 7 days at 37°, or one month at 4°C. As a nice collateral effect, NS3 3D is also characterized by a higher reactivity against HCV positive samples than c33, improving detection of low positive samples and early seroconversion points. Behaviour of the NS3 3D antigen in the final anti-HCV kit prototype is still to be determined; interaction with the other components of the kit should be investigated; still, NS3 3D seems to be a

more powerful instrument than c33 to detect anti-NS3 antibodies. Expression yield of NS3 3D in standard conditions is quite low, in comparison to c33. However, a limited and rudimental optimization, involving only expression strain and codon usage, raised the expression yield at least three-fold, bringing the level of production to more acceptable values. Optimization of culture conditions and fermentation process development, which are not covered in this work, are expected to further improve NS3 3D productivity. Furthermore, higher sensitivity of NS3 3D antigen will probably allow using a lower quantity of protein per kit, reducing antigen consumption in the manufacturing process.

Besides the added value of NS3 3D over c33 for the industrial application, this study also indirectly gave some more general insights on the HCV NS3 helicase domain and its antigenic properties. First of all, inclusion of the C-terminal part of NS3 helicase domain (completing subdomain 2 and adding subdomain 3) into the antigen sequence definitely improved the antigenicity of the construct. This is in striking contrast with data reported in literature, where influence of the C-terminal part of the NS3 helicase domain on overall antigenicity has often been considered negligible; the general consensus is that the immunodominant epitopes of NS3 are contained in the c33 region [77][99], but at least part of the NS3 helicase immunogenicity is carried by conformational-dependent epitopes [14]. Some of these conformational epitopes have been experimentally identified [17][70] or predicted through artificial neural networks [62], but none of them is located outside the c33 region. The discrepancy between these data and the findings of this thesis could be interpreted in two different ways: either the sequence included in NS3 3D but not in c33 contains one or more unknown conformational epitopes; or the presence of complete subdomains 2 and 3 improves the general folding of the *E. coli*-expressed NS3 3D, making it more similar to its native counterpart than c33 alone and in turn enhancing the recognition of epitopes contained in the c33 region. However, whatever the correct interpretation of the phenomenon may be (they could also be both true), implication of conformational epitopes is corroborated by the phenomenon of thermal susceptibility of c33 antigenicity; in fact, absence of antigen degradation or aggregation has been demonstrated, and consumption of possible critical cofactors during storage has been ruled out. Therefore, the drop in antigenicity of c33 observed upon thermal stress should be caused by some conformational epitopes that are destroyed or made unaccessible during some kind of remodeling (or local irreversible unfolding) of the protein. Similarly, NS3 3D thermal resistance is strictly dependent from storage buffer composition. The extended sequence of NS3 3D is more antigenic than c33, for the reasons discussed above; nevertheless, this higher antigenicity is quickly lost if the protein is not kept in the correct buffer. Presence of an inorganic buffer holds a crucial function in maintaining the reactive configuration of NS3 3D, as its replacement with

MES (or TRIS, as demonstrated by other tests, not reported in this work) completely accounts for the antigen inactivation recorded after thermal stress. Circular dichroism analysis showed that c33 undergoes a rapid structure transition with a midpoint temperature around 35°C, characterized by a sharp decrease of the β-sheet content, but not associated to a complete unfolding of the protein, as high ellipticity is retained even at 80°C. β-sheets of NS3 helicase are located in the core of subdomains 1 and 2, and in the extended β-loop connecting subdomains 2 and 3. In c33, probably only the β-sheet of subdomain 1 and part of the β-sheet of subdomain 2 are correctly structured: decrease of the β-sheet content upon increase of the temperature could reflect a collapse of the core of subdomain 1 and 2. This structural rearrangement, however, is completed in a few minutes (the time of a wavelength scan during CD measurement), while a sensible reduction of antigenicity happens on a longer time scale (at least 24 hours). Furthermore, intramolecular crosslinking of c33 with formaldehyde abolishes the structural transition but not the loss in antigenicity. This suggests that either the two phenomena are not linked in a cause-effect relationship, or the loss of antigenicity is caused by another structural modification with slow kinetics, that is somehow primed by the fast rearrangement detected with circular dichroism.

To block this putative rearrangement, many attempts were made in this work, aimed to increase protein rigidity: fixation with formaldehyde, that has been discussed before; addition of putative cofactors or inhibitors, an approach often used in crystallography to stabilize transient conformations; removal or tying up of possible loose ends, caused by the artificial N- and C-terminal ends of c33; reconstruction of the structural integrity of the helicase domain, by the addition of complete subdomain 2 and 3. The last approach, combined with modification of the storage buffer (which had been optimized for the c33-based EIA), succeeded in stabilizing the antigen, meeting the primary target and at the same time improving the overall performances of the antigen. A more extensive characterization of the NS3 3D antigen will be required on larger population of sera before inclusion in the final commercial kit, however this study clearly shows that NS3 3D is a very promising candidate to replace c33 in the CLIA format anti-HCV immunoassay.

6. References

[1] Alter, M. J. (1993). The detection, transmission, and outcome of hepatitis C virus infection. *Infectious agents and disease, 2*, pp. 155-166.

[2] Alter, M. J.et al. (2003). Guidelines for laboratory testing and result reporting of antibody to hepatitis C virus. Centers for Disease Control and Prevention. *MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports / Centers for Disease Control, 52*, p. 1-13, 15; quiz CE1-4.

[3] Ansaldi, F.et al. (2006). Combination hepatitis C virus antigen and antibody immunoassay as a new tool for early diagnosis of infection. *Journal of viral hepatitis, 13*, pp. 5-10.

[4] Anzola, M.et al. (2003). Hepatocellular carcinoma: molecular interactions between hepatitis C virus and p53 in hepatocarcinogenesis. *Expert reviews in molecular medicine, 5*, pp. 1-16.

[5] Baalmann-Mangano, L.et al. (2003). Evaluating liver disease in chronic hepatitis C--the role of the liver biopsy. *MedGenMed : Medscape general medicine, 5*, p. 21.

[6] Bartenschlager, R.et al. (1993). Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *Journal of virology, 67*, pp. 3835-3844.

[7] Brown, E. A.et al. (1992). Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic acids research, 20*, pp. 5041-5045.

[8] Burgess-Brown, N. A.et al. (2008). Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study. *Protein expression and purification, 59*, pp. 94-102.

[9] Carey, W. (2003). Tests and screening strategies for the diagnosis of hepatitis C. *Cleveland Clinic journal of medicine, 70 Suppl 4*, p. S7-13.

[10] Carithers, R. L. J.et al. (2000). Diagnostic testing for hepatitis C. *Seminars in liver disease, 20*, pp. 159-171.

[11] Carrère-Kremer, S.et al. (2004). Regulation of hepatitis C virus polyprotein processing by signal peptidase involves structural determinants at the p7 sequence junctions. *The Journal of biological chemistry, 279*, pp. 41384-41392.

[12] Caruthers, J. M.et al. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, pp. 13080-13085.
[13] Cavazza, S.et al. (2005). Indeterminate third-generation hepatitis C recombinant immunoblot assay and HCV RNA analysis: isolated reactivity against NS5 associated with HCV viraemia in clinical patients but not blood donors. *Scandinavian journal of infectious diseases*, *37*, pp. 488-492.
[14] Chen, M.et al. (1998). Human and murine antibody recognition is focused on the ATPase/helicase, but not the protease domain of the hepatitis C virus nonstructural 3 protein. *Hepatology (Baltimore, Md.)*, *28*, pp. 219-224.
[15] Choo, Q. L.et al. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*, *244*, pp. 359-362.
[16] Choo, Q. L.et al. (1991). Genetic organization and diversity of the hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, *88*, pp. 2451-2455.
[17] Claeys, H.et al. (1995). Localization and reactivity of an immunodominant domain in the NS3 region of hepatitis C virus. *Journal of medical virology*, *45*, pp. 273-281.
[18] Collett, M. S.et al. (1989). Recent advances in pestivirus research. *The Journal of general virology*, *70 (Pt 2)*, pp. 253-266.
[19] Craik, D. J. (2006). Chemistry. Seamless proteins tie up their loose ends. *Science*, *311*, pp. 1563-1564.
[20] Davis, G. L. (2002). Monitoring of viral levels during therapy of hepatitis C. *Hepatology (Baltimore, Md.)*, *36*, p. S145-51.
[21] De Francesco, R.et al. (2003). Approaching a new era for hepatitis C virus therapy: inhibitors of the NS3-4A serine protease and the NS5B RNA-dependent RNA polymerase. *Antiviral research*, *58*, pp. 1-16.
[22] Di Bisceglie, A. M. (1997). Hepatitis C and hepatocellular carcinoma. *Hepatology (Baltimore, Md.)*, *26*, p. 34S-38S.
[23] Di Bisceglie, A. M. (1998). Hepatitis C. *Lancet*, *351*, pp. 351-355.
[24] Di Bisceglie, A. M.et al. (2002). Optimal therapy of hepatitis C. *Hepatology (Baltimore, Md.)*, *36*, p. S121-7.
[25] Failla, C.et al. (1994). Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. *Journal of virology*, *68*, pp. 3753-3760.

[26] Feinstone, S. M.et al. (1975). Transfusion-associated hepatitis not due to viral hepatitis type A or B. *The New England journal of medicine*, *292*, pp. 767-770.
[27] Feinstone, S. M.et al. (1983). Inactivation of hepatitis B virus and non-A, non-B hepatitis by chloroform. *Infection and immunity*, *41*, pp. 816-821.
[28] Ferraro, D.et al. (2008). Assessment of hepatitis C virus-RNA clearance under combination therapy for hepatitis C virus genotype 1: performance of the transcription-mediated amplification assay. *Journal of viral hepatitis*, *15*, pp. 66-70.
[29] Forns, X.et al. (1999). Quasispecies in viral persistence and pathogenesis of hepatitis C virus. *Trends in microbiology*, *7*, pp. 402-410.
[30] Frank, C.et al. (2000). The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. *Lancet*, *355*, pp. 887-891.
[31] Friebe, P.et al. (2002). Genetic analysis of sequences in the 3' nontranslated region of hepatitis C virus that are important for RNA replication. *Journal of virology*, *76*, pp. 5326-5338.
[32] Gesell, J. J.et al. (2001). Design, high-level expression, purification and characterization of soluble fragments of the hepatitis C virus NS3 RNA helicase suitable for NMR-based drug discovery methods and mechanistic studies. *Protein engineering*, *14*, pp. 573-582.
[33] Ghany, M. G.et al. (2009). Diagnosis, management, and treatment of hepatitis C: an update. *Hepatology (Baltimore, Md.)*, *49*, pp. 1335-1374.
[34] Gorbalenya, A. E.et al. (1989). N-terminal domains of putative helicases of flavi- and pestiviruses may be serine proteases. *Nucleic acids research*, *17*, pp. 3889-3897.
[35] Gozdek, A.et al. (2008). NS3 Peptide, a novel potent hepatitis C virus NS3 helicase inhibitor: its mechanism of action and antiviral activity in the replicon system. *Antimicrobial agents and chemotherapy*, *52*, pp. 393-401.
[36] Gremion, C.et al. (2005). Hepatitis C virus and the immune system: a concise review. *Reviews in medical virology*, *15*, pp. 235-268.
[37] Gretch, D. R. (1997). Use and interpretation of HCV diagnostic tests in the clinical setting. *Clinics in liver disease*, *1*, p. 543-57, vi.
[38] Gretch, D.et al. (1992). Use of aminotransferase, hepatitis C antibody, and hepatitis C polymerase chain reaction RNA assays to

establish the diagnosis of hepatitis C virus infection in a diagnostic virology laboratory. *Journal of clinical microbiology*, *30*, pp. 2145-2149.
[39] Grote, A.et al. (2005). JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic acids research*, *33*, p. W526-31.
[40] Gómez, J.et al. (1999). Hepatitis C viral quasispecies. *Journal of viral hepatitis*, *6*, pp. 3-16.
[41] Higuchi, M.et al. (2002). Epidemiology and clinical aspects on hepatitis C. *Japanese journal of infectious diseases*, *55*, pp. 69-77.
[42] Horsmans, Y.et al. (2005). Isatoribine, an agonist of TLR7, reduces plasma virus concentration in chronic hepatitis C infection. *Hepatology (Baltimore, Md.)*, *42*, pp. 724-731.
[43] James, W.et al. (1997). Method and composition for controlling formaldehyde fixation by delayed quenching. *International Patent Publication WO/1997/039330*, , .
[44] Jin, L.et al. (1995). Expression, isolation, and characterization of the hepatitis C virus ATPase/RNA helicase. *Archives of biochemistry and biophysics*, *323*, pp. 47-53.
[45] Kaito, M.et al. (1994). Hepatitis C virus particle detected by immunoelectron microscopic study. *The Journal of general virology*, *75 (Pt 7)*, pp. 1755-1760.
[46] Kang, C. B.et al. (2008). FKBP family proteins: immunophilins with versatile biological functions. *Neuro-Signals*, *16*, pp. 318-325.
[47] Kang, L. W.et al. (1998). Crystallization and preliminary X-ray crystallographic analysis of the helicase domain of hepatitis C virus NS3 protein. *Acta crystallographica. Section D, Biological crystallography*, *54*, pp. 121-123.
[48] Khu, Y. L.et al. (2001). Mutations that affect dimer formation and helicase activity of the hepatitis C virus helicase. *Journal of virology*, *75*, pp. 205-214.
[49] Kiiver, K.et al. (2006). Complex formation between hepatitis C virus NS2 and NS3 proteins. *Virus research*, *117*, pp. 264-272.
[50] Kim, D. W.et al. (1995). C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. *Biochemical and biophysical research communications*, *215*, pp. 160-166.
[51] Kim, D. W.et al. (1997). Towards defining a minimal functional domain for NTPase and RNA helicase activities of the hepatitis C virus NS3 protein. *Virus research*, *49*, pp. 17-25.

[52] Kim, J. L.et al. (1996). Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell*, *87*, pp. 343-355.
[53] Kim, J. L.et al. (1998). Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure (London, England : 1993)*, *6*, pp. 89-100.
[54] Kleinman, S.et al. (1992). Increased detection of hepatitis C virus (HCV)-infected blood donors by a multiple-antigen HCV enzyme immunoassay. *Transfusion*, *32*, pp. 805-813.
[55] Knappe, T. A.et al. (2007). Insertion of a chaperone domain converts FKBP12 into a powerful catalyst of protein folding. *Journal of molecular biology*, *368*, pp. 1458-1468.
[56] Kolarski, V. (1999). [The European consensus on hepatitis C]. *Vutreshni bolesti*, *31*, pp. 21-28.
[57] Kotloff, K. L.et al. (2001). Safety and immunogenicity of increasing doses of a Clostridium difficile toxoid vaccine administered to healthy adults. *Infection and immunity*, *69*, pp. 988-995.
[58] Krasnorutskii, M. A.et al. (2008). Antibodies against pancreatic ribonuclease A hydrolyze RNA and DNA. *International immunology*, *20*, pp. 1031-1040.
[59] Kwong, A. D.et al. (1998). Hepatitis C virus NS3/4A protease. *Antiviral research*, *40*, pp. 1-18.
[60] Kwong, A. D.et al. (2000). Structure and function of hepatitis C virus NS3 helicase. *Current topics in microbiology and immunology*, *242*, pp. 171-196.
[61] Lam, A. M. I.et al. (2006). Hepatitis C virus subgenomic replicon requires an active NS3 RNA helicase. *Journal of virology*, *80*, pp. 404-411.
[62] Lara, J.et al. (2008). Artificial neural network for prediction of antigenic activity for a major conformational epitope in the hepatitis C virus NS3 protein. *Bioinformatics (Oxford, England)*, *24*, pp. 1858-1864.
[63] Lavanchy, D. (1999). Hepatitis C: public health strategies. *Journal of hepatology*, *31 Suppl 1*, pp. 146-151.
[64] Levin, M. K.et al. (2002). Helicase from hepatitis C virus, energetics of DNA binding. *The Journal of biological chemistry*, *277*, pp. 29377-29385.
[65] Lissen, E.et al. (1993). Hepatitis C virus infection among sexually promiscuous groups and the heterosexual partners of hepatitis C virus infected index cases. *European journal of clinical*

microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology, *12*, pp. 827-831.
[66] Liu, D.et al. (2001). Solution structure and backbone dynamics of an engineered arginine-rich subdomain 2 of the hepatitis C virus NS3 RNA helicase. *Journal of molecular biology*, *314*, pp. 543-561.
[67] Maillard, P.et al. (2001). Nonenveloped nucleocapsids of hepatitis C virus in the serum of infected patients. *Journal of virology*, *75*, pp. 8240-8250.
[68] McHutchison, J. G.et al. (1999). Combination therapy with interferon plus ribavirin for the initial treatment of chronic hepatitis C. *Seminars in liver disease*, *19 Suppl 1*, pp. 57-65.
[69] Miroux, B.et al. (1996). Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of molecular biology*, *260*, pp. 289-298.
[70] Mondelli, M. U.et al. (1994). Significance of the immune response to a major, conformational B-cell epitope on the hepatitis C virus NS3 region defined by a human monoclonal antibody. *Journal of virology*, *68*, pp. 4829-4836.
[71] Nolte, F. S. (1998). Branched DNA signal amplification for direct quantitation of nucleic acid sequences in clinical specimens. *Advances in clinical chemistry*, *33*, pp. 201-235.
[72] Oliver, C. (1994). Pre-embedding labeling methods. *Methods in molecular biology (Clifton, N.J.)*, *34*, pp. 315-319.
[73] Op De Beeck, A.et al. (2004). Characterization of functional hepatitis C virus envelope glycoproteins. *Journal of virology*, *78*, pp. 2994-3002.
[74] Paliwal, R.et al. (1996). Comparison of the conformation, hydrophobicity, and model membrane interactions of diphtheria toxin to those of formaldehyde-treated toxin (diphtheria toxoid): formaldehyde stabilization of the native conformation inhibits changes that allow membrane insertion. *Biochemistry*, *35*, pp. 2374-2379.
[75] Pawlotsky, J. M. (1999). Diagnostic tests for hepatitis C. *Journal of hepatology*, *31 Suppl 1*, pp. 71-79.
[76] Pawlotsky, J. M.et al. (2000). Standardization of hepatitis C virus RNA quantification. *Hepatology (Baltimore, Md.)*, *32*, pp. 654-659.
[77] Pentón, N.et al. (2003). Antigenicity of a recombinant NS3 protein representative of ATPase/helicase domain from hepatitis C virus. *Clinical biochemistry*, *36*, pp. 41-49.

[78] Poole, T. L.et al. (1995). Pestivirus translation initiation occurs by internal ribosome entry. *Virology*, *206*, pp. 750-754.
[79] Poynard, T.et al. (1997). Natural history of liver fibrosis progression in patients with chronic hepatitis C. The OBSVIRC, METAVIR, CLINIVIR, and DOSVIRC groups. *Lancet*, *349*, pp. 825-832.
[80] Preugschat, F.et al. (1996). A steady-state and pre-steady-state kinetic analysis of the NTPase activity associated with the hepatitis C virus NS3 helicase domain. *The Journal of biological chemistry*, *271*, pp. 24449-24457.
[81] Puigbò, P.et al. (2007). OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic acids research*, *35*, p. W126-31.
[82] Raab, D.et al. (2004). Method and device for optimizing a nucleotide sequence for the purpose of expression of a protein. *International Patent Publication WO/2004/059556*, , .
[83] Randall, G.et al. (2004). Interfering with hepatitis C virus RNA replication. *Virus research*, *102*, pp. 19-25.
[84] Rice, P.et al. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends in genetics : TIG*, *16*, pp. 276-277.
[85] Robertson, B.et al. (1998). Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. International Committee on Virus Taxonomy. *Archives of virology*, *143*, pp. 2493-2503.
[86] Saldanha, J.et al. (1999). Establishment of the first international standard for nucleic acid amplification technology (NAT) assays for HCV RNA. WHO Collaborative Study Group. *Vox sanguinis*, *76*, pp. 149-158.
[87] Salnikova, M. S.et al. (2008). Physical characterization of clostridium difficile toxins and toxoids: effect of the formaldehyde crosslinking on thermal stability. *Journal of pharmaceutical sciences*, *97*, pp. 3735-3752.
[88] Sarrazin, C. (2002). Highly sensitive hepatitis C virus RNA detection methods: molecular backgrounds and clinical significance. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology*, *25 Suppl 3*, p. S23-9.
[89] Scholz, C.et al. (2005). Functional solubilization of aggregation-prone HIV envelope proteins by covalent fusion with chaperone modules. *Journal of molecular biology*, *345*, pp. 1229-1241.

[90] Scholz, C.et al. (2008). Chaperone-aided in vitro renaturation of an engineered E1 envelope protein for detection of anti-Rubella virus IgG antibodies. *Biochemistry*, *47*, pp. 4276-4287.
[91] Sen, G. C. (2001). Viruses and interferons. *Annual review of microbiology*, *55*, pp. 255-281.
[92] Serebrov, V.et al. (2004). Periodic cycles of RNA unwinding and pausing by hepatitis C virus NS3 helicase. *Nature*, *430*, pp. 476-480.
[93] Serebrov, V.et al. (2009). Establishing a mechanistic basis for the large kinetic steps of the NS3 helicase. *The Journal of biological chemistry*, *284*, pp. 2512-2521.
[94] Shah, D. O.et al. (2003). Combination HCV core antigen and antibody assay on a fully automated chemiluminescence analyzer. *Transfusion*, *43*, pp. 1067-1074.
[95] Shimotohno, K. (2000). Hepatitis C virus and its pathogenesis. *Seminars in cancer biology*, *10*, pp. 233-240.
[96] Simmonds, P. (1999). Viral heterogeneity of the hepatitis C virus. *Journal of hepatology*, *31 Suppl 1*, pp. 54-60.
[97] Subramanya, H. S.et al. (1996). Crystal structure of a DExx box DNA helicase. *Nature*, *384*, pp. 379-383.
[98] Sy, T.et al. (2006). Epidemiology of hepatitis C virus (HCV) infection. *International journal of medical sciences*, *3*, pp. 41-46.
[99] Sällberg, M.et al. (1996). Immunogenicity and antigenicity of the ATPase/helicase domain of the hepatitis C virus non-structural 3 protein. *The Journal of general virology*, *77 (Pt 11)*, pp. 2721-2728.
[100] Tahallah, N.et al. (2002). Cofactor-dependent assembly of the flavoenzyme vanillyl-alcohol oxidase. *The Journal of biological chemistry*, *277*, pp. 36425-36432.
[101] Tai, C. L.et al. (1996). The helicase activity associated with hepatitis C virus nonstructural protein 3 (NS3). *Journal of virology*, *70*, pp. 8477-8484.
[102] Tellinghuisen, T. L.et al. (2002). Interaction between hepatitis C virus proteins and host cell factors. *Current opinion in microbiology*, *5*, pp. 419-427.
[103] Thimme, R.et al. (2001). Determinants of viral clearance and persistence during acute hepatitis C virus infection. *The Journal of experimental medicine*, *194*, pp. 1395-1406.
[104] Thomas, P.et al. (2008). Methylenetetrahydrofolate reductase, common polymorphisms, and relation to disease. *Vitamins and hormones*, *79*, pp. 375-392.

[105] Thomson, M.et al. (2003). The clearance of hepatitis C virus infection in chimpanzees may not necessarily correlate with the appearance of acquired immunity. *Journal of virology*, *77*, pp. 862-870.
[106] Tobler, L. H.et al. (2005). Performance of ORTHO HCV core antigen and trak-C assays for detection of viraemia in pre-seroconversion plasma and whole blood donors. *Vox sanguinis*, *89*, pp. 201-207.
[107] Toews, J.et al. (2008). Mass spectrometric identification of formaldehyde-induced peptide modifications under in vivo protein cross-linking conditions. *Analytica chimica acta*, *618*, pp. 168-183.
[108] Tomei, L.et al. (1993). NS3 is a serine protease required for processing of hepatitis C virus polyprotein. *Journal of virology*, *67*, pp. 4017-4026.
[109] Tsukiyama-Kohara, K.et al. (1992). Internal ribosome entry site within hepatitis C virus RNA. *Journal of virology*, *66*, pp. 1476-1483.
[110] Uyttendaele, S.et al. (1994). Evaluation of third-generation screening and confirmatory assays for HCV antibodies. *Vox sanguinis*, *66*, pp. 122-129.
[111] Whitehead, T. A.et al. (2009). Tying up the loose ends: circular permutation decreases the proteolytic susceptibility of recombinant proteins. *Protein engineering, design & selection : PEDS*, *22*, pp. 607-613.
[112] Witthöft, T. (2008). Review of consensus interferon in the treatment of chronic hepatitis C. *Biologics*, *2*, pp. 635-643.
World Health Organization (2003). *Global Alert and Response (GAR) - Hepatitis C*. Retrieved from http://www.who.int/csr/disease/hepatitis/whocdscsrlyo2003/en/index.html.
[114] Wu, J.et al. (2009). Enhancement of recombinant human ADAM15 disintegrin domain expression level by releasing the rare codons and amino acids restriction. *Applied biochemistry and biotechnology*, *157*, pp. 299-310.
[115] Yano, M.et al. (1996). The long-term pathological evolution of chronic hepatitis C. *Hepatology (Baltimore, Md.)*, *23*, pp. 1334-1340.
[116] Yeung, L. T. F.et al. (2007). Spontaneous clearance of childhood hepatitis C virus infection. *Journal of viral hepatitis*, *14*, pp. 797-805.
[117] Yeung, L. T.et al. (2001). Mother-to-infant transmission of hepatitis C virus. *Hepatology (Baltimore, Md.)*, *34*, pp. 223-229.

[118] Yoshiba, M.et al. (1994). Genotype of hepatitis C virus in fulminant hepatitis C. *Digestive diseases and sciences*, *39*, pp. 220-221.
[119] Zeuzem, S. (2008). Interferon-based therapy for chronic hepatitis C: current and future perspectives. *Nature clinical practice. Gastroenterology & hepatology*, *5*, pp. 610-622.
[120] al Meshari, K.et al. (1995). Hepatitis C virus infection in hemodialysis patients: comparison of two new hepatitis C antibody assays with a second-generation assay. *Journal of the American Society of Nephrology : JASN*, *6*, pp. 1439-1444.
[121] van Berkel, W. J.et al. (1991). The conformational stability of the redox states of lipoamide dehydrogenase from Azotobacter vinelandii. *European journal of biochemistry / FEBS*, *202*, pp. 1049-1055.

7. Riassunto

Il virus dell'epatite C (HCV), isolato nel 1989 come agente eziologico della maggior parte delle epatiti non riconducibili ai virus dell'epatite A e B, infetta ad oggi circa 180 milioni di persone, ovvero il 3% della popolazione mondiale, secondo le stime del WHO. L'infezione diviene cronica in almeno il 70% dei casi, e può portare a gravi complicazioni quali cirrosi ed epatocarcinomi. Tali patologie possono insorgere anche a distanza di molti anni, durante i quali l'infezione risulta asintomatica nella maggior parte dei casi. La principale via di trasmissione del virus è lo scambio di sangue infetto, attraverso trasfusioni da donatori infetti, procedure medico-sanitarie in condizioni di igiene precaria o l'abuso di droghe intravenose. L'elevata diffusione del virus e la gravità delle sue complicazioni evidenziano la necessità di strumenti diagnostici affidabili e precisi per la rilevazione di infezioni da HCV. Il campo di applicazione di questi test comprende non solo il monitoraggio dei soggetti a rischio e la valutazione degli effetti del trattamento sui pazienti infetti, ma anche lo screening dei donatori di sangue e plasma. I test attualmente in commercio per la determinazione di infezioni da HCV sono di tipo immunologico o molecolare. Per via del facile impiego e del costo inferiore, i test immunologici sono i più usati, soprattutto in fase di screening di ampie popolazioni. L'ultima generazione di test per la rilevazione di anticorpi diretti contro proteine di HCV fa uso di antigeni ricombinanti appartenenti alle proteine virali immunodominanti NS3, NS4 e core. Lo sviluppo della tecnologia del saggio immunologico sta rapidamente modificando il metodo di esecuzione dei test da formati manuali o semi-automatici, come l'ELISA (Enzyme-Linked ImmunoSorbent Assay) o il RIA (Radio ImmunoAssay), verso formati più automatizzabili, caratterizzati da maggiore throughput e minore necessità di intervento dell'operatore. Liaison® è lo strumento sviluppato da Diasorin S.p.A. per l'esecuzione di saggi immunologici in formato CLIA (ChemiLuminescent ImmunoAssay) in modo completamente automatizzato. Lo strumento fa uso di kit che contengono i reagenti necessari per l'esecuzione del saggio e che vengono caricati sulla macchina all'occorrenza. Diasorin attualmente produce un test per la rilevazione di anticorpi diretti contro HCV in formato ELISA. L'adattamento di questo saggio al formato CLIA per l'inserimento nel menu prodotti Liaison® è un obbiettivo prioritario. Nella maggior

parte dei casi, l'adattamento di un test dal formato ELISA al formato CLIA richiede solo la modifica della molecola tracciante, non più coniugata ad una perossidasi ma al substrato chemiluminescente ABEI (N-(4-Amino-Butyl)-N-Ethyl-Isoluminol), ed un processo di ottimizzazione della composizione delle soluzioni in cui i reagenti vengono conservati. Sfortunatamente, durante l'adattamento del kit anti-HCV ELISA al formato CLIA è stata evidenziata la comparsa di una criticità relativa all'antigene c33, che si verifica solo all'interno del prototipo CLIA. c33 è un frammento immunodominante della proteasi/elicasi virale NS3, comprendente poco più della metà del dominio elicasico. Tale dominio è composto da tre sottodomini, dei quali il primo e parte del secondo sono compresi nella sequenza di c33. L'antigene c33 ha mostrato un'immunoreattività ottimale solo se conservato a temperatura controllata (2-8 °C). In caso di stress termico a 37°C, l'antigene perde gran parte della sua attività diagnostica in poche ore. Le possibilità che stress termici di questa entità si verifichino durante la distribuzione del prodotto finito, specialmente in aree del mondo caratterizzate da problemi logistici e temperature elevate, sono piuttosto elevate. Per questo motivo, la suscettibilità dell'antigene a stress di natura termica porterebbe a gravi problemi logistici e di affidabilità per la commercializzazione del prodotto finito. In questa tesi viene presentato il lavoro effettuato allo scopo di aumentare la stabilità dell'antigene c33 a stress termici, in modo da renderlo compatibile con la distribuzione su larga scala senza comprometterne le proprietà antigeniche. Il sistema modello applicato per testare la resistenza a stress termico è quello di un test immunologico di tipo indiretto, in cui l'antigene (c33 o derivati) è immobilizzato sulla fase solida, per catturare eventuali anticorpi specifici presenti nel campione di siero. Il legame di IgG alla fase solida è successivamente evidenziato tramite un anticorpo monoclonale anti-IgG umane, coniugato ad ABEI. Il segnale ottenuto subito dopo il coating della fase solida viene confrontato con quello ottenuto sugli stessi campioni dopo un'incubazione del kit a 37°C per tre giorni.

La caratterizzazione del fenomeno di instabilità termica ha evidenziato che non sussistono problemi di degradazione proteolitica, e che la proteina c33 presenta una ben precisa transizione di tipo conformazionale, caratterizzata da una diminuzione del contenuto in β-sheet a temperature superiori ai 35°C. Diversi approcci sono stati utilizzati per aumentare la stabilità termica della proteina. In primo luogo sono stati utilizzati metodi che avrebbero consentito di

mantenere inalterata la sequenza primaria dell'antigene c33, il che avrebbe rappresentato un vantaggio da un punto di vista produttivo. In tal senso è stato valutato l'effetto dell'aggiunta di cofattori e inibitori, potenzialmente in grado di bloccare la proteina in una struttura più "chiusa" e quindi più compatta e auspicabilmente più stabile; questo approccio, mutuato da tecniche usate in cristallografia, non ha portato risultati in termini di stabilizzazione. L'antigene c33 è stato successivamente sottoposto a trattamento con formaldeide, reagente in grado di creare legami intramolecolari nell'antigene. Analisi di dicroismo circolare hanno indicato che tale trattamento permette di annullare la transizione conformazionale a 35°C, ma la diminuzione di reattività dopo stress termico è risultata inalterata. Si è quindi passati a modificare la struttura primaria dell'antigene. Eventuali problemi di aggregazione dell'antigene durante lo stress termico sono stati esclusi esprimendo c33 come proteina di fusione con il chaperone molecolare SlyD, più volte utilizzato in letteratura per aumentare la solubilità di proteine poco solubili; anche in questo casi non si sono ottenuti miglioramenti nella stabilità termica dell'antigene. Un'altra possibile causa di instabilità che è stata vagliata è la presenza di regioni non strutturate al C- ed all'N-terminale dell'antigene; è stato rimosso un β-sheet incompleto all'estremità C-terminale di c33; parallelamente, c33 è stata inserita nella sequenza della proteina umana FKBP12 all'interno di un loop che è stato dimostrato essere tollerante all'inserzione di domini, mantenendo bloccate ambedue le estremità dell'inserto in regioni stabilmente strutturate. Nessuno dei due approcci ha sortito effetto in termini di stabilizzazione della reattività. La ricostituzione dell'intero dominio elicasico di NS3, ottenuta aggiungendo a c33 anche la sequenza mancante del secondo e terzo sottodominio, ha permesso di ottenere un antigene caratterizzato da una sensibilità molto più elevata. Inoltre, l'utilizzo di una soluzione ottimizzata di stoccaggio della fase solida ha permesso di ottenere una completa stabilizzazione allo stress termico di tale antigene (la reattività rimane inalterata fino a 7 giorni di incubazione a 37°C); questo non si verifica invece utilizzando la soluzione ottimizzata in congiunzione con l'antigene c33 originale. Particolarmente interessante è osservare che le sequenze aggiunte appartenenti ai sottodomini 2 e 3 non sono di per sé antigeniche, secondo quanto riportato in letteratura. Si può ipotizzare che esse compongano un determinante antigenico di natura conformazionale finora ignoto, o che la loro presenza influenzi la conformazione globale dell'antigene c33, rendendolo più simile alla forma naturale e quindi meglio

riconosciuta dagli anticorpi di soggetti infetti. Avendo individuato un antigene dotato delle caratteristiche richieste, si è proceduto all'ottimizzazione dei livelli di espressione migliorando l'adattamento dei codoni e testando diversi ceppi ospiti di *E. coli*. Tale ottimizzazione ha permesso di aumentare la produttività di almeno tre volte. Il nuovo antigene è quindi stato testato su una popolazione aperta di 128 sieri, sui quali si è ottenuta una distribuzione dei valori dei campioni negativi molto compatta ed una buona discriminazione dei campioni positivi. Infine, è stata individuata una sequenza, donata dal vettore di espressione, responsabile di alcune crossreattività aspecifiche individuate contro l'antigene; sia la rimozione di questa sequenza in fase di clonaggio, sia l'aggiunta di una proteina scavenger nel tampone di diluizione dei campioni hanno permesso di eliminare efficacemente le reattività aspecifiche.
L'antigene così ottenuto, grazie alla maggiore sensibilità ed alla stabilità allo stress termico rappresenta un candidato molto promettente per la sostituzione dell'antigene c33 nel saggio di tipo CLIA per il rilevamento di infezioni da HCV.