Expression and function of osteopontin variants in HCV-related liver disease and hepatocellular carcinoma

Renée Jade Phillips, B. Biotech (Hons)

THE DEPARTMENT OF MICROBIOLOGY AND IMMUNOLOGY

THE SCHOOL OF MOLECULAR AND BIOMEDICAL SCIENCE

THE UNIVERSITY OF ADELAIDE

A dissertation submitted to The University of Adelaide

In candidature for the degree of

Doctor of Philosophy in the Faculty of Science

April 2010

CHAPTER 1: Introduction

1.1 Hepatitis C Virus (HCV)

1.1.1 Natural history and classification

Hepatitis C virus (HCV) is a significant infectious agent and one of the leading causes of chronic liver disease worldwide. During the 1970s, Alter and colleagues discovered that up to 90% of post-transfusion hepatitis cases were not caused by either hepatitis A (HAV) or hepatitis B (HBV) virus (Alter, H. J. et al., 1975; Feinstone et al., 1975; Seeff et al., 1975), the only known hepatitis viruses at the time. This hepatitis virus was termed non-A non-B hepatitis (NANBH) until molecular cloning by recombinant DNA technology in 1989 identified it as a new hepatitis virus, which we now know as HCV (Choo et al., 1989).

HCV is the sole member of the hepacivirus genus of the *Flaviviridae* family (Robertson et al., 1998), which also includes flaviviruses such as yellow fever virus, Japanese encephalitis virus and dengue virus, and pestiviruses such as bovine virus diarrhea and classical swine fever. HCV has 6 genotypes (or "clades", numbered 1-6) and approximately 50 subtypes which are grouped by genomic variation in the NS5B region (Simmonds et al., 1993; Bukh et al., 1995). The 6 genotypes differ in genomic sequence by 30-35% (Heinz et al., 2000) and their geographical distributions vary greatly (Figure 1.1). Genotypes 1, 2, and 3 are found worldwide, whilst genotypes 4 (Egypt, Middle East), 5 (South Africa) and 6 (South East Asia) are confined to smaller areas (Zein, N. N. 2000; Nguyen and Keeffe 2005). In Australia, genotypes 1 and 3 predominate; subtypes 1a and 1b comprise 55% of all HCV infections and genotype 3 another 33% (Baker et al., 1996; McCaw et al., 1997; Mison et al., 1997; Kaba et al., 1998). All

Figure 1.1 Geographical distribution of HCV genotypes worldwide (adapted from Webster et al., 2000).

genotypes generate similar clinical outcomes and severity of illness but vary in their response to treatment (Bukh et al., 1995; Lau, J. Y. et al., 1996).

1.1.2 Epidemiology of HCV infection worldwide and in Australia

The most recent World Health Organisation (WHO) estimates put the number of people chronically infected with HCV worldwide at approximately 170 million, or 3% of the world's population, with another 3 to 4 million new infections every year. HCV infection occurs in every corner of the globe, however higher prevalence rates are strongly associated with low socioeconomic and developing areas (Figure 1.2). In Australia, there are an estimated 264,000 people currently exposed to HCV, with 9,700 new infections occuring annually mainly through injecting drug use (IDU) (Law et al., 2003; Law et al., 2006). The highest incidence occurs in the 20-39 year old age group (Figure 1.3), which correlates with the peak IDU incidence group (Dore et al., 2003). The number of new HCV infections in Australia increased steadily from 1961-2001 (Law et al., 2003) and continues to rise, posing a major health problem in this country. The widespread nature and increasing number of cases both in Australia and throughout the world mean that hepatitis C infection remains a costly burden on the world's healthcare systems.

1.1.3 Routes of transmission

Transmission of HCV readily occurs through the exchange of blood and bodily fluids. Prior to the availability of blood screening technology, HCV was primarily transmitted during blood transfusions (Tobler and Busch 1997) and organ transplantations (Eggen and Nordbo 1992; Pereira et al., 1992) from infected donors. Following the identification of the virus in 1989, the number of newly acquired post-transfusion hepatitis cases attributed to HCV declined dramatically due to screening of donors for anti-HCV antibodies (Donahue et al., 1992; Schreiber et al., 1996; Tobler and Busch 1997). Although transmission via blood transfusion and organ

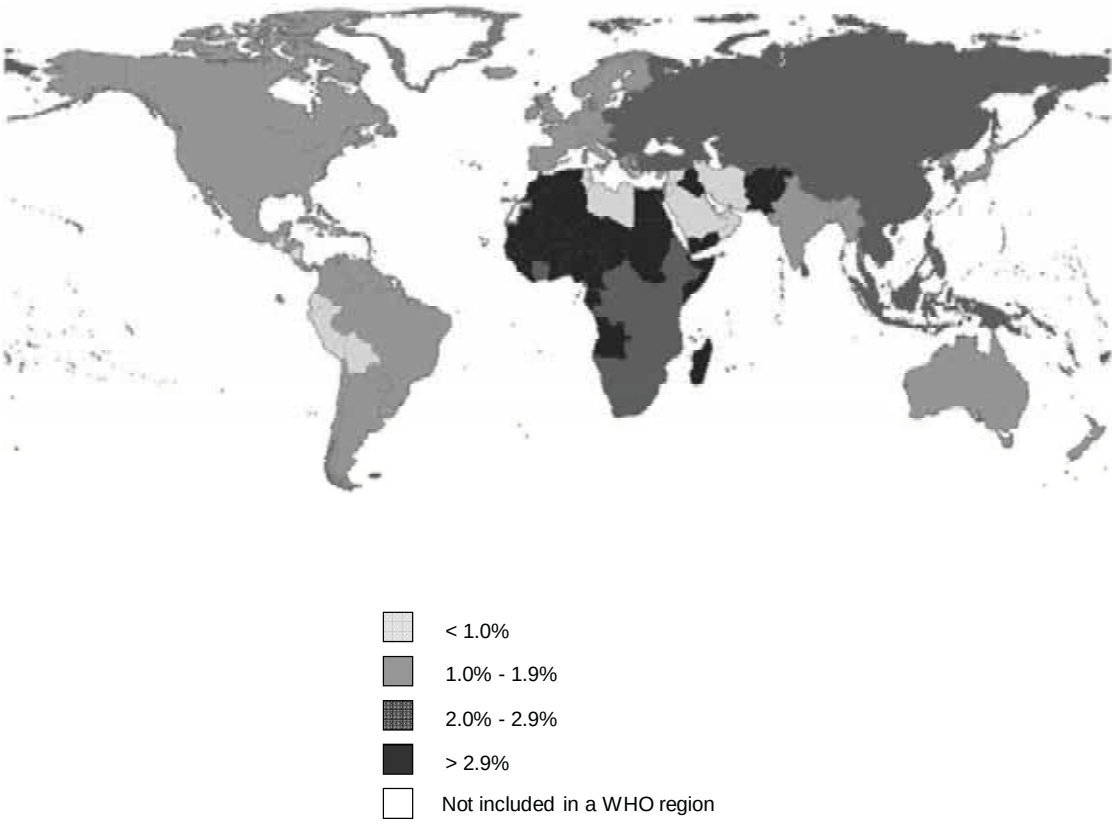


Figure 1.2 Worldwide prevalence of HCV (adapted from Global Burden of Disease).

Figure 1.3 Rate of Australian HCV notifications per 100,000 population in 2000 (NCHECR 2001).

donation rarely occurs today, a large number of people currently living with chronic hepatitis C were infected in this manner.

Today HCV is most commonly transmitted by the sharing of needles and other equipment between injecting drug users, especially amongst young adults (Wasley and Alter 2000; Dore et al., 2003). IDU transmission accounts for 68% of all current HCV infections in the US, and 80% in England and Australia (Ramsay et al., 1998; Alter, M. J. 2002; Dore et al., 2003). Similarly, the virus can be transferred by the improper sterilisation of equipment used during medical and surgical procedures, acupuncture, tattooing, and body-piercing (Mele et al., 1995; Sun, D. X. et al., 1996; Thomas 2000; Shin et al., 2002; de Nishioka et al., 2003; Haley and Fischer 2003; Howe et al., 2005; Prati 2006). Rare cases have been documented via other means, including the sharing of toothbrushes and razors (Tumminelli et al., 1995; Kelly 2000), sexual activity (Rooney and Gilson 1998; Wejstal 1999; Thomas 2000), vertical transmission (from infected mother to baby during childbirth) (Dore et al., 1997; Zanetti et al., 1999), and occupational needlestick injuries to healthcare workers (Ridzon et al., 1997; Thomas 2000).

1.1.4 Types and outcomes of HCV infection

1.1.4.1 Acute and chronic infection

Infection with HCV can result in one of two outcomes: an acute infection that is readily resolved, or one that develops into a chronic infection that, without treatment, persists for life. 15 - 20% of individuals will spontaneously resolve their acute infection, however chronic infection is the more common outcome, occurring in up to 85% of patients. Whilst in the majority of cases a resolved acute infection results in no damage to the liver, chronic infection can lead to cirrhosis, liver failure and hepatocellular carcinoma (HCC; Figure 1.4). The reasons why some people but not others progress to chronicity remain unclear, however there are many factors thought to be

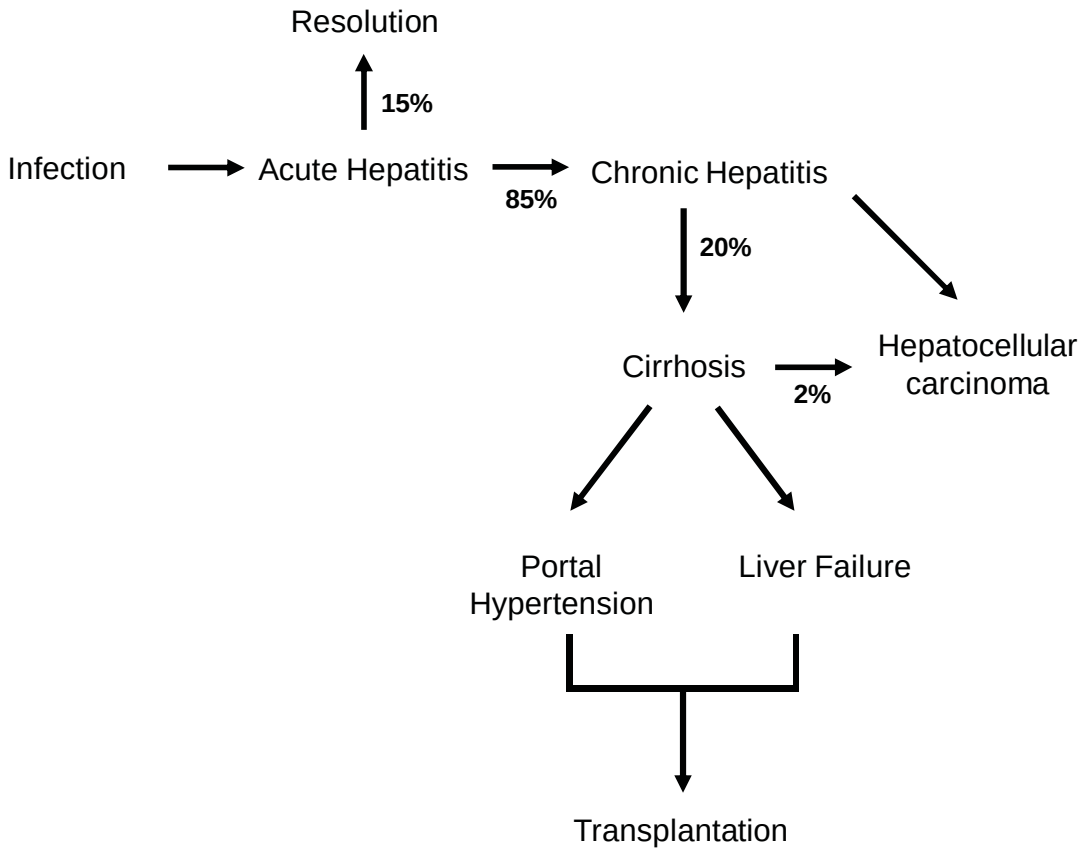


Figure 1.4 Outcomes of HCV infection.

positively correlated with a tendency towards developing chronicity, including age older than 20 years at infection (Alter, M. J. et al., 1999; Bellentani and Tiribelli 2001), being of African-American heritage (Alter, M. J. et al., 1999; Howell et al., 2000), and an absence of jaundice during the early stages of infection (Wiese et al., 2000; Gerlach et al., 2001).

The incubation period of acute HCV infection generally lasts from 2 to 12 weeks and is characterised by a transient increase in serum alanine transaminase (ALT) levels, which is indicative of liver damage (Figure 1.5) (Hoofnagle 2002). Acute infection is often asymptomatic, with clinical symptoms apparent in only 30% of cases (Aach et al., 1991; Koretz et al., 1993; Iwasaki et al., 2002). HCV RNA is detected in most patients 1 - 2 weeks after exposure (Farci et al., 1991; Thimme et al., 2001), rising rapidly to a peak of 10^5 - 10^7 IU/ml immediately before peak ALT levels and onset of symptoms. Anti-HCV antibodies are first detected shortly after symptoms arise, and are present in 80% of individuals within 15 weeks of initial exposure (Alter, M. J. et al., 1992). Anti-HCV antibodies are generally long-lasting after acute infection (Alter, H. J. and Seeff 2000). Fulminant liver failure is rare following resolution of an acute infection (Liang et al., 1993; Wright 1993). Up to 67% of symptomatic acute HCV infected individuals will spontaneously clear the virus within 3 - 4 months of symptom onset (Hofer et al., 2003), with the rest maintaining a low viral load. Those with stronger $CD4^+$ and $CD8^+$ T-cell responses are more likely to clear the virus (Hofer et al., 2003).

Chronic hepatitis C (CHC) infection is defined by the persistence of serum HCV RNA for more than 6 months after initial exposure (Figure 1.6) (Alter, M. J. et al., 1999). Elevated serum ALT levels are also associated with CHC, however one third of patients maintain normal values (Conry-Cantilena et al., 1996; Zoulim et al., 2003). Patients with consistently normal ALT levels generally have a milder and less progressive liver disease than those with elevated ALTs (Haber

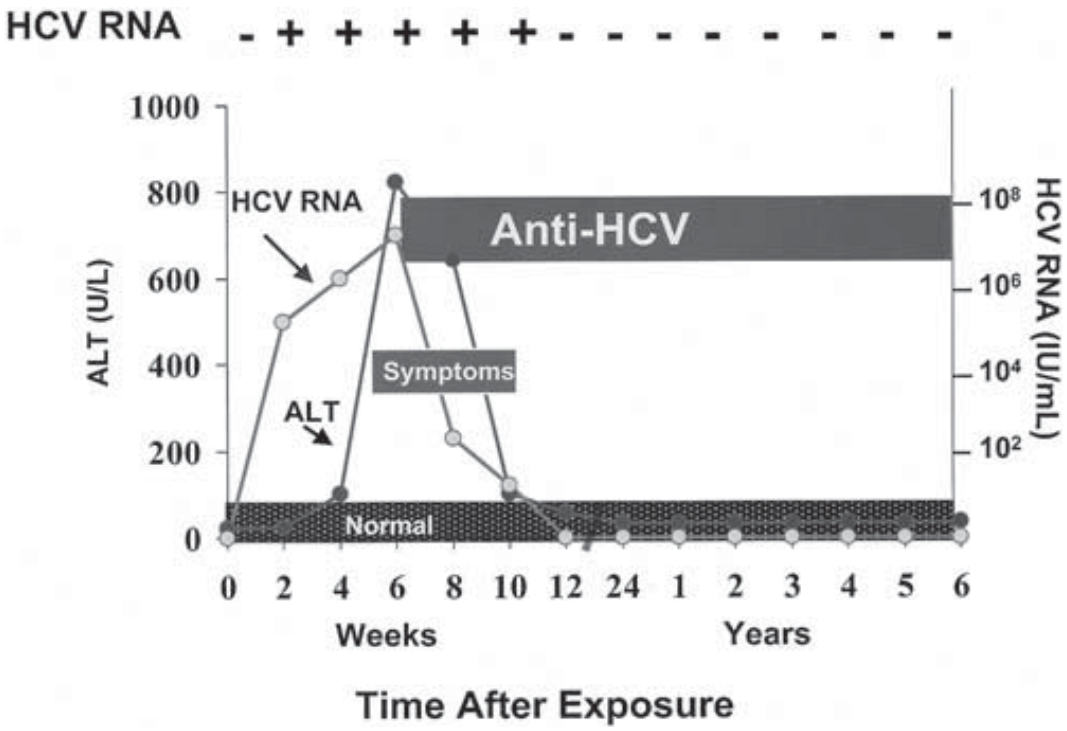


Figure 1.5 Course of acute, resolving HCV infection. Qualitative HCV RNA results are indicated as positive (+) or negative (-). Quantitative HCV RNA results are shown in IU/ml. Serum ALT levels are shown in U/L (Hoofnagle 2002).

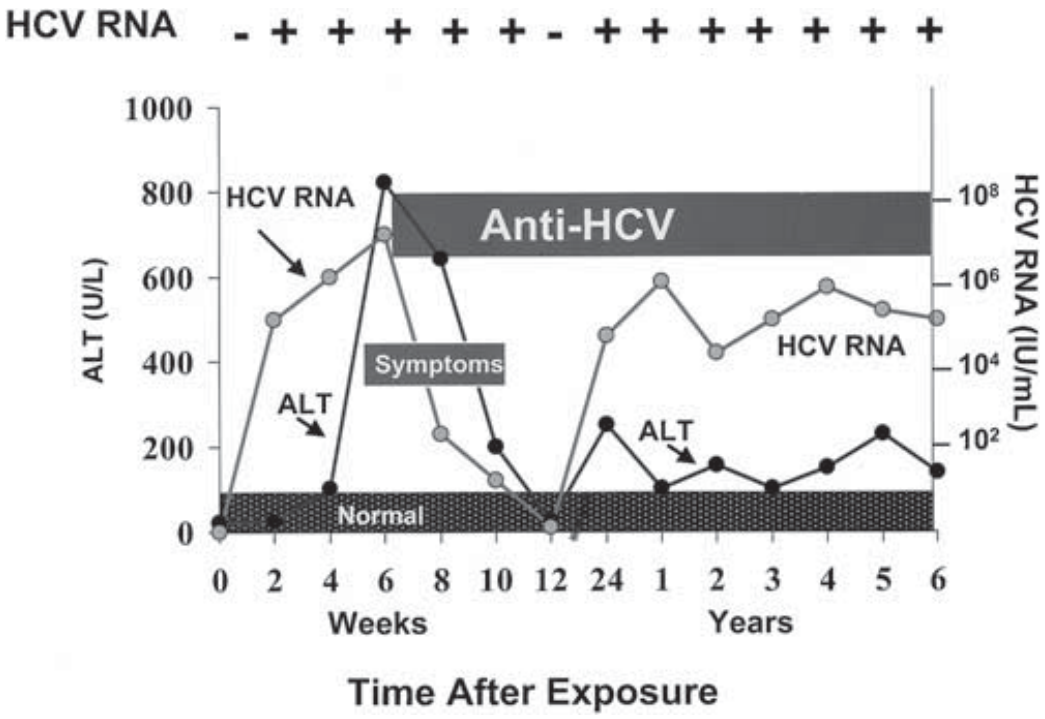


Figure 1.6 Course of acute HCV infection that evolves into a chronic infection. Qualitative HCV RNA results are indicated as positive (+) or negative (-). Quantitative HCV RNA results are shown in IU/ml. Serum ALT levels are shown in U/L (Hoofnagle 2002).

et al., 1995; Shakil et al., 1995; Conry-Cantilena et al., 1996; Hoofnagle 1997; Martinot-Peignoux et al., 2001). Patients with CHC are less likely to develop symptoms and jaundice than those with a resolving acute infection (Alter, H. J. and Seeff 2000). During the evolution of an acute infection to a persistent one, HCV RNA and ALT levels can vary greatly (Alter, H. J. et al., 1995; Major et al., 1999; Thimme et al., 2001), however RNA levels stabilise once chronicity is reached (Yeo et al., 2001). Anti-HCV levels are generally higher and more likely to persist in a chronic infection than in an acute resolving one (Takaki et al., 2000).

CHC has been associated with a number of extrahepatic manifestations including mixed cryoglobulinemia, B-cell non-Hodgkin lymphoma, seronegative arthritis, thyroid abnormalities, glomerulonephritis, and neurological disorders (Agnello et al., 1992; Misiani et al., 1992; Johnson, R. J. et al., 1993; Pawlotsky et al., 1994; Silvestri et al., 1996; Zuckerman et al., 1997; Heckmann et al., 1999; Lunel and Cacoub 1999; Cacoub et al., 2000; Hausfater et al., 2001). These complications can have a negative effect on the quality of life of those affected by CHC infection.

1.1.4.2 Progression of CHC-related liver disease

Over the course of CHC infection, necroinflammatory processes in the liver result in progressive hepatic fibrosis, cirrhosis and loss of liver structure and function. As HCV itself is not directly cytopathic, the majority of liver disease is caused by the host inflammatory response directed against HCV-infected hepatocytes. The long-term effects of CHC infection include cirrhosis, end-stage liver disease and HCC, which in most cases take decades to progress and only develop in a small proportion of patients (Figure 1.7) (Alter, H. J. and Seeff 2000; Liang et al., 2000).

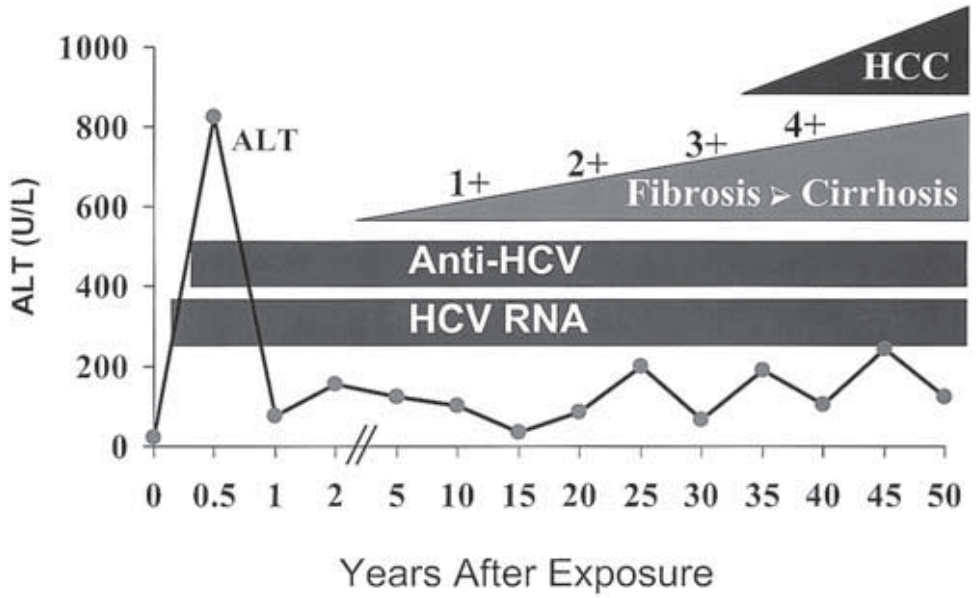


Figure 1.7 Natural history of chronic HCV infection over 50 years. Once chronicity is established, liver damage can progress through worsening stages of fibrosis (from none to 3+) to cirrhosis (4+) to HCC. HCC usually occurs in a patient with underlying cirrhosis (Hoofnagle 2002).

The progression of CHC-induced liver disease begins slowly, with few symptoms or physical signs in the first two decades following infection. Fibrosis, or scarring of the liver, develops gradually as chronic immune infiltration fights the viral infection of hepatocytes. Continual fibrotic injury develops into cirrhosis 20 - 25 years after initial infection in approximately 20% of chronically infected patients (Di Bisceglie 2000; NIH 2002; Rakela and Vargas 2002; Strader et al., 2004; Thomas and Seeff 2005), with cirrhosis being one of the leading indicators for liver transplantation (Zein, C. O. and Zein 2002; Szabo et al., 2003). The rate of progression of fibrosis to cirrhosis varies widely, with patients generally grouped into one of three categories: a rapid progressor who develops cirrhosis less than 20 years after initial infection, an intermediate progressor who develops cirrhosis 20 - 50 years after initial infection, or a slow or no progressor who develops cirrhosis more than 50 years after initial infection or does not develop cirrhosis at all (Poynard et al., 1997). 3 - 4% of those with cirrhosis will develop HCC every year (Davis et al., 2003). The Centres for Disease Control (CDC) estimate that there will be 30 - 40,000 deaths per year from end-stage liver disease by 2010 (Alter, M. J. 1997; Armstrong et al., 2000). There are many factors that can influence the outcome of CHC towards more serious liver disease, including older age at infection, male gender, African-American or Hispanic race, increased ALT levels, concurrent smoking or alcohol use, steatohepatitis (fatty liver), longer duration of infection and co-infection with human immunodeficiency virus (HIV) or HBV (Poynard et al., 1997; Zarski et al., 1998; Ramalho et al., 2000; Graham et al., 2001; Mohsen et al., 2002; Freeman et al., 2003; Poynard et al., 2003; WHO 2004).

1.1.4.3 Pathogenesis of CHC-related liver disease

The liver is the body's largest organ, and is responsible for many essential functions including carbohydrate, lipid and protein metabolism, red blood cell degradation, bile synthesis, vitamin storage, and detoxification. It is comprised of two lobes which contain an ordered structure of

hepatocytes, hepatic stellate cells (HSCs) and resident macrophages (Küpffer cells). The structural unit of the liver is the hepatic lobule, which consists of a roughly hexagonal arrangement of plates of hepatocytes radiating outwards from a central vein and separated by hepatic sinusoids. Situated at vertices between the lobules are portal triads, containing a bile duct and a terminal branch of both the hepatic artery and portal vein. When healthy, the liver acts as the body's main filtration system and keeps many of the body's systems in balance.

When an individual's liver becomes chronically infected with HCV, the host immune response acts to rid infected hepatocytes of the virus, setting off a chain of events that causes the liver great structural damage that can progress through various stages of fibrosis to cirrhosis, portal hypertension, liver failure and HCC. The development of liver fibrosis is a multicellular process involving paracrine signaling between the resident liver cells (hepatocytes and HSCs) and inflammatory cells (Friedman et al., 2000), and is a result of the healing response of the liver to repeated injury incorporating chronic liver damage and accumulation of extracellular matrix (ECM) proteins (Figure 1.8) (Friedman 2003). During a chronic HCV infection, virally infected hepatocytes release chemokines, resulting in a massive influx of mononuclear cells and large areas of inflammation. These mononuclear cells secrete pro-inflammatory cytokines such as IL-2, IFN-γ, TNF-α and IL-1β, IL-4 and IL-10 to promote inflammation (Gramenzi et al., 2005; Lau et al., 2005; Farinati et al., 2006), and apoptose infected cells in an attempt to control or clear the infection (Neuman et al., 2007).

Hepatocyte mass is regenerated to replace apoptotic cells resulting in mononuclear cell infiltration in portal and parenchyma areas (Desmet et al., 1994) and deposition of ECM proteins. Necroinflammatory disease is present in most cases, but the severity of disease (activity or "grade") and degree of structural damage (fibrosis or "stage") varies greatly. The degree of

Figure 1.8 Changes in hepatic architecture in a healthy liver **(A)** during progression of fibrosis in chronic HCV infection. During repeated injury to the liver during chronic HCV infection **(B)**, inflammatory lymphocytes infiltrate the liver parenchyma, some hepatocytes are apoptosed, and Kuppfer cells are activated to release fibrogenic mediators. Hepatic stellate cells are also activated and secrete large amounts of extracellular matrix proteins. Gaps between the sinusoidal epithelial cells become closed, constricting the sinusoidal lumen and restricting blood flow through the liver (Bataller et al., 2005).

necroinflammatory disease may be a good prognostic indicator of future worsening of fibrosis and development of cirrhosis (Ghany et al., 2000); worse necroinflammatory disease coupled with high serum ALTs may indicate a higher likelihood of fibrosis and cirrhosis.

If chronic hepatic injury continues, liver regeneration eventually fails and apoptotic and necrotic hepatocytes are replaced with ECM, including fibrous collagen. This ECM deposition, or fibrosis, is initially located around portal tracts and gradually extends into lobules towards the central vein (Scheuer et al., 1992; Lefkowitch et al., 1993; Goodman and Ishak 1995), disrupting the architecture of the liver by forming fibrous scars. HSCs have been identified as the main collagen-producing cells in the liver (Friedman et al., 1985; Gabele et al., 2003), and are activated following chronic liver insult (Milani et al., 1990; Marra 1999). Activated HSCs home to areas of cellular regeneration and secrete large amounts of ECM proteins. The degree of fibrosis in an infected liver is determined by performing a liver biopsy and analysing the activity grade and fibrosis stage using a number of scoring systems, including METAVIR (stages F0 - F4), Scheuer, HAI (stages 0 - 18), and Ishak (stages 1 - 6) (Marcellin et al., 2002). As liver disease progresses further, portal fibrous bands form septa which connect the bands together (Marcellin et al., 2002; Ramadori and Saile 2004), causing encapsulation of nodules of regenerating hepatocytes. At this point, the liver becomes severely compromised and blood is unable to circulate properly. This stage is referred to as cirrhosis and is generally thought to be irreversible, however recent data suggests that reversal may be possible (Wanless et al., 2000; Desmet and Roskams 2004). Cirrhosis represents the final stage of fibrosis, but further development of liver disease can lead to the development of primary liver cancer, or HCC.

1.1.5 Treatment of infection

The underlying goal of HCV treatment is to eliminate the virus from the liver, therefore slowing the progression of liver disease and preventing the development of cirrhosis and HCC. The standard treatment option for HCV infection comprises pegylated interferon-alpha (peg-IFN-α) and the nucleoside analogue ribavirin. This combination results in sustained virological response (SVR) rates of 40 - 50% for patients infected with genotype 1, 80% for those with genotypes 2 or 3, and 55 - 69% for those with genotype 4 (NIH 2002; El-Zayadi et al., 2005; Kamal et al., 2005). This treatment regime is associated with a number of side effects including flu-like symptoms, vomiting, depression, weight loss, insomnia, rashes, neutropenia, hypo- and hyper-thyroidism, itching, sinusitis and haemolytic anaemia (Russo and Fried 2003). Treatment is costly; a 48-week course costs US$25,000 - $40,000 (Reed 2003; Hoofnagle and Seeff 2006; Shah and Wong 2006), and one WHO study suggests that 90% of patients with HCV who require treatment cannot afford it (WHO 1999).

Alternative treatment strategies currently being investigated include inhibitors for HCV NS3-4A (serine protease) and NS5B (RNA polymerase), referred to as STAT-C (specifically targeted antiviral therapy for hepatitis C). The aim of these drugs is to inhibit specific phases of the HCV life cycle whilst giving improved SVR rates, shorter treatment time and fewer side effects than the standard combination treatment. There is currently no vaccine against HCV.

1.1.6 Viral genome and proteins

HCV is a small, enveloped virus of 55 - 65 nm in diameter (Kaito et al., 1994). Packaged within the viral envelope is a 33 - 40 nm icosahedral nucleocapsid comprised of HCV core protein (Ishida et al., 2001) and containing the 9.6 kilobase positive sense single stranded RNA genome (Kato et al., 1990; Choo et al., 1991). The genome encodes a single open reading frame (ORF)

flanked by 5' and 3' untranslated regions (UTR) which play important roles in translation and viral replication (Han et al., 1991; Tanaka et al., 1995; Ito et al., 1998). The ORF is translated into a solitary polyprotein of 3010 - 3030 amino acids (Choo et al., 1991; Okamoto et al., 1991; Takamizawa et al., 1991) which is cleaved co- and post-translationally into ten distinct protein products including structural (core, E1, E2, and p7) and non-structural (NS2, 3, 4A, 4B, 5A, and 5B) proteins (Figure 1.9) (Tan et al., 2002).

The structural proteins, cleaved by host proteases (Hijikata et al., 1991; Santolini et al., 1994), are the building blocks for virion assembly. Core protein is the major component of the viral nucleocapsid and possesses a wide range of functions (Kato et al., 1990). Glycoproteins E1 and E2 are type 1 transmembrane proteins and form non-covalent heterodimers that constitute the viral envelope (Deleersnyder et al., 1997) which facilitates attachment and entry into host cells, most likely through interactions of E2 with one or more receptors on the hepatocyte surface. At present the essential HCV entry factors include the tetraspanin CD81, the class B scavenger receptor SR-B1 and the tight junction proteins claudin-1 and occludin (Pileri et al., 1998; Scarselli et al., 2002; Evans et al., 2007; Ploss et al., 2009). The p7 protein is thought to form hexameric ion channels not unlike viroporins (Griffin et al., 2003; Pavlovic et al., 2003) and may be essential for virus infectivity (Sakai et al., 2003).

The non-structural proteins of HCV are self-cleaved by the NS2/3 and NS3/NS4A proteases and are responsible for polyprotein processing and viral replication. The carboxy terminus of NS2 and the amino terminus of NS3 form a metalloprotease to cleave NS2 from NS3 (Hijikata et al., 1991; Grakoui et al., 1993b). NS3 is a serine protease responsible for all downstream protein cleavages (Bartenschlager et al., 1993; Eckart et al., 1993; Grakoui et al., 1993a; Tomei et al., 1993; Manabe et al., 1994) and also acts as an RNA helicase (Kim, D. W. et al., 1995; Gwack et

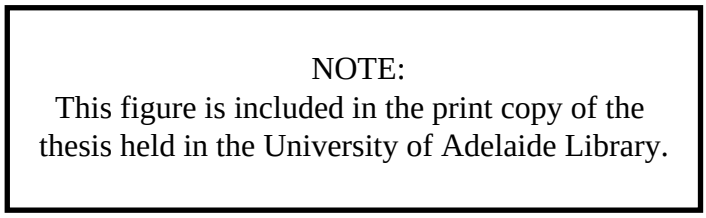


Figure 1.9 The HCV genome (Tan et al 2002).

al., 1996; Preugschat et al., 1996; Levin, M. K. et al., 2005). NS4A acts as a co-factor for NS3 to accelerate cleavage of non-structural proteins (Failla et al., 1994; Bartenschlager et al., 1995; Failla et al., 1995; Lin, C. and Rice 1995; Lin, C. et al., 1995; Tanji et al., 1995; Shimizu et al., 1996) and may be essential for hyperphosphorylation of NS5B (Tanji et al., 1995; Neddermann et al., 1999). NS4B is known to integrate into the endoplasmic reticulum (ER) membrane and may help induce formation of the viral replication complex (Hugle et al., 2001; Egger et al., 2002). NS5A is a multifunctional protein which has been shown to be an essential part of the replication complex (Gosert et al., 2003) and localises to lipid droplets where it plays a vital role in virion morphogenesis (Appel et al., 2008). The RNA dependant RNA polymerase NS5B is responsible for HCV replication (Koonin 1991; Behrens et al., 1996; Lohmann et al., 1997).

1.1.7 Viral replication

Replication of the HCV genome occurs primarily in hepatocytes (Lanford et al., 1995), however HCV RNA has been detected in B and T lymphocytes, monocytes and peripheral blood mononuclear cells (Roque-Afonso et al., 2005; MacParland et al., 2006; Pal et al., 2006). Due to the lack of suitable *in vitro* and *in vivo* model systems, the study of HCV replication has been severely hampered and therefore largely modeled on the replication cycles of related pestiviruses and flaviviruses. However, the recent development of an authentic HCV cell culture system based on a Japanese HCV isolate that caused fulminant hepatitis (JFH-1) will allow for more in-depth study of HCV replication in the future (Wakita et al., 2005). In brief, the HCV replication cycle begins with entry of the virus particle into the host cell and concludes with budding of virus particles from the cell (Figure 1.10) (Bartenschlager and Lohmann 2000; Suzuki, T. et al., 2007), and is summarised below.

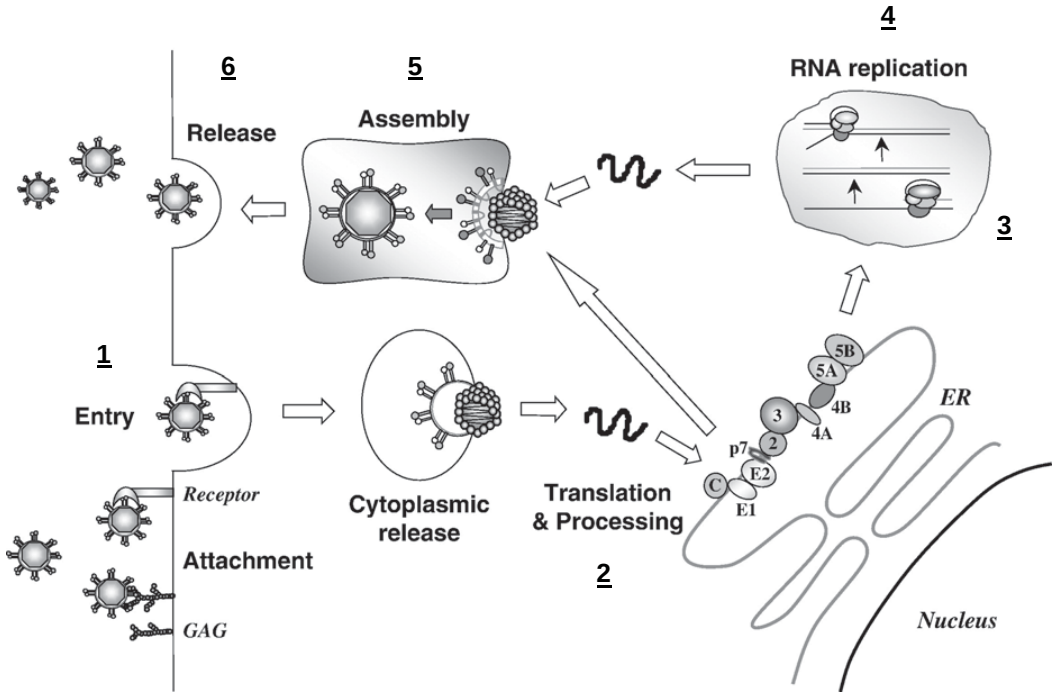


Figure 1.10 Proposed HCV life cycle (adapted from Suzuki et al., 2007).

1. *Attachment and Entry*

The virus enters the cell via clatherin-dependent endocytosis, thought to occur through interactions between the HCV E2 protein and one or more host receptors (as previously discussed in Section 1.1.6). Viral particles are uncoated (by acidification of the endosome) to release genomic RNA into the cytosol of the host cell.

2. *RNA translation and polyprotein processing*

The positive-sense single stranded RNA genome is directly translated by the internal ribosome entry site (IRES) at the ER membrane and the resultant polyprotein is cleaved. Cleavage products form a membraneous-like replication complex at the ER to facilitate production of viral RNA and proteins (Egger et al., 2002; Gosert et al., 2003).

3. *Synthesis of minus-strand RNA intermediate*

A negative strand RNA intermediate is synthesised from pregenomic RNA by the replication complex.

4. *Production of new positive-strand RNA*

The RNA intermediate is transcribed to provide a double-stranded template for production of positive sense genomic RNA, which can be packaged into a nucleocapsid core ready for release from the host cell, used as a template for further translation of viral proteins, or cycled back to the replication complex for continued production of negative strand intermediate.

5. *Virion packaging*

Nucleocapsids containing genomic RNA bud into the ER and acquire a viral envelope ready for release (Baumert et al., 1998; Ezelle et al., 2002; Murakami et al., 2006).

6. *Virus release from infected cell*

Mature virus particles move from the ER to the golgi apparatus where they are transported to the cell membrane and released (Sato, K. et al., 1993).

1.2 Hepatocellular Carcinoma (HCC)

1.2.1 Introduction

Hepatocellular carcinoma (HCC) accounts for 85% of all primary malignant liver tumors (Kew 2002), and represents the final, often fatal, stage of cirrhosis. Those with cirrhosis have a 1 - 6.5% chance per year of developing HCC (Chen, R. Y. et al., 2002), and approximately one third of individuals with cirrhosis will develop HCC over their lifetime (Sangiovanni et al., 2004). HCC is the fifth most common cancer worldwide and the third leading cause of cancer-related death (Okuda 2000; Parkin et al., 2001; Chen, J. G. et al., 2006). It is directly responsible for approximately 598,000 deaths per year (Parkin et al., 2005). WHO estimated that in 2002 cirrhosis and HCC caused 783,000 and 619,000 deaths respectively, collectively representing 2.5% of all deaths worldwide (WHO 2003).

1.2.2 Epidemiology

HCC is recognised as a global disease, however 80% of new cases occur in developing countries (Parkin et al., 2005; Wong, C. M. and Ng 2008). High prevalence rates are seen in East and South-East Asia and sub-Saharan Africa, whilst prevalence in Europe and North America is much lower (Wong, C. M. and Ng 2008). Yearly incidence (new cases per year per 100,000 population) follows a similar global pattern (Figure 1.11) (Leong and Leong 2005).

Whilst HCC incidence is low in North America and Europe, the age-adjusted incidence of HCC has risen 15 - 20% in Japan (Kew 2002) and doubled in the US (Bergsland and Venook 2000; El-Serag 2004) over the past 20 years. Liver cancer had the fastest growing mortality rate of all cancers in the US during recent years (Ries et al., 2008). It is widely thought that the recent increase in HCC incidence in western countries is due to rising prevalence of HCV infection, increased immigration from endemic HBV areas, and conditions such as non-alcoholic fatty liver

Figure 1.11 Yearly incidence (new cases per year per 100,000 population) of HCC worldwide (adapted from Leong and Leong 2005).

disease (non-alcoholic steatosis; NASH) being seen in increasing numbers due to the recent obesity epidemic (Kew 2002; Davis et al., 2003; But et al., 2008). HCV-related HCC has been estimated to peak in 2010 in low incidence areas (Wong, J. B. et al., 2000). Conversely, the high incidence rates in developing countries are slowly decreasing due to effective HBV vaccination programs and improved control of aflatoxin exposure (But et al., 2008).

1.2.3 Underlying causes of HCC tumor development

HCC tumors develop from prolonged and repeated liver damage caused by a number of disorders, including chronic HBV and HCV infection (Anzola 2004; Brechot 2004), chronic alcohol abuse (Stickel et al., 2002; Morgan et al., 2004), aflatoxin B_1 exposure (Yu, M. C. and Yuan 2004), NASH (Calle et al., 2003), Wilson's disease (Chu and Hung 1993) and inherited metabolic diseases such as haemochromatosis (Kowdley 2004). The majority of these disorders promote chronic inflammation of the liver which often progresses into cirrhosis and over time can develop into HCC. It is widely recognised that the presence of two or more risk factors in an individual can dramatically increase their risk of developing HCC (Smart et al., 1998; Kew 2002; Davis et al., 2003; Kew 2006). HBV and HCV infection are the two leading indicators for HCC, collectively responsible for approximately 80% of cases worldwide (Bosch et al., 2005; McGlynn and London 2005; Perz et al., 2006).

1.2.3.1 Virally-induced HCC

1.2.3.1.1 Hepatitis B Virus

HBV is the number one cause of HCC worldwide, and the geographical distributions of HBV infection and HCC incidence are closely matched (Bosch et al., 1999; Bosch et al., 2005). Chronic infection with HBV increases the risk of HCC development 200-fold over uninfected individuals (Leong and Leong 2005). The HBV viral genome integrates into the host genome and

it has been postulated that this integration may activate growth factors and endogenous genes such as SERCA1 and cyclin A, thus deregulating cell division and resulting in uncontrolled cellular growth (Wang, J. et al., 1990; Chami et al., 2000). Integration may also inactivate tumor suppressor genes and oncogenes responsible for controlling cell survival and death, leading to abnormal cell growth and tumor development (Leong and Leong 2005; Minami et al., 2005), and induce chromosomal instability through rearrangements and deletions, resulting in possible DNA mutations associated with tumor progression (Brechot et al., 1980; But et al., 2008). Certain HBV proteins, such as the x protein (HBx), may themselves possess oncogenic properties (Feitelson and Duan 1997). HBx inactivation of the p53 tumor suppressor gene represses p53-mediated apoptosis and is thought to play a major role in the malignant transformation of HBV-infected hepatocytes (Wang, X. W. et al., 1994). HBx also interferes with DNA repair and therefore results in a number of critical mutations capable of inducing carcinogenesis (Becker et al., 1998), however the role of the x protein remains controversial. Infection with HBV also indirectly causes continuous hepatic injury (via the host immune response to viral replication) and regeneration as observed in chronic HCV infections (see Section 1.1.4.3).

1.2.3.1.2 Hepatitis C Virus

HCV is the second most common cause of HCC worldwide. HCV-related cirrhosis confers a greater risk of HCC development on a patient than HBV-related cirrhosis (El-Serag 2001). In developed countries such as Italy, France, and Spain, 27 - 75% of HCC cases can be attributed to HCV infection (Deuffic et al., 1998; El-Serag 2001; Montalto et al., 2002; Bosch et al., 2005; El-Serag and Rudolph 2007). In contrast to most Asian countries (where HBV is the most common cause of HCC), HCV infection accounts for up to 90% of all reported HCC cases in Japan, with HCV-related HCC incidence tripling over the past 40 years (Bosch et al., 2005; El-Serag and Rudolph 2007). HCV infected individuals have a 17-fold greater chance of developing HCC than

uninfected individuals (Donato et al., 2002). HCC development in CHC patients without cirrhosis is rare (EASL 1999); those with cirrhosis have a 5.8% chance per year of developing HCC, compared to 0.5 - 2.6% for those with lesser degrees of fibrosis (Kiyosawa 2002). Unlike HBV the HCV genome does not integrate, and HCV is therefore thought to cause HCC through indirect means as described previously (see Section 1.1.4.3).

The exact molecular mechanisms of carcinogenesis in the HCV-infected liver have not been fully elucidated. Malignant transformation of hepatocytes is thought to arise via increased liver cell turnover induced by chronic liver insult and regeneration from inflammation, the host immune response, and oxidative DNA damage. This may induce genetic changes such as the over-expression of growth and angiogenic factors, activation of cellular oncogenes and deactivation of tumor suppressor genes, activation of telomerases, DNA mismatch repairs, impaired chromosomal segregation and loss of alleles (Chan et al., 1999; Ozturk 1999; Bergsland 2001; Thorgeirsson and Grisham 2002; Block et al., 2003; Mas et al., 2004; Suriawinata and Xu 2004). A proposed sequence of events leading to HCC in the liver is outlined in Figure 1.12.

A 2003 study of HCC arising from HBV and HCV infections and chronic alcohol use identified three major pathways that are affected during hepatocarcinogenesis, namely the Rb1, Wnt and p53 pathways (Edamoto et al., 2003). The most common genetic alterations were loss of Rb1 expression via methylation of its promoter, $p16^{INK4A}$ methylation and cyclin D1 amplification. The normal function of $p16^{INK4A}$ is to inhibit cyclin dependant kinases 4 and 6, which phosphorylate Rb1 and block cell cycle progression at the G1 phase (Feitelson et al., 2002); therefore, methylation of $p16^{INK4A}$ allows cell cycle progression to proceed unchecked. Mutations of the Wnt signaling protein β-Catenin are common in early HCC and affect expression of target genes including cyclin D1, c-myc, c-jun and fibronectin (Calvisi et al., 2001; Zimmers et al.,

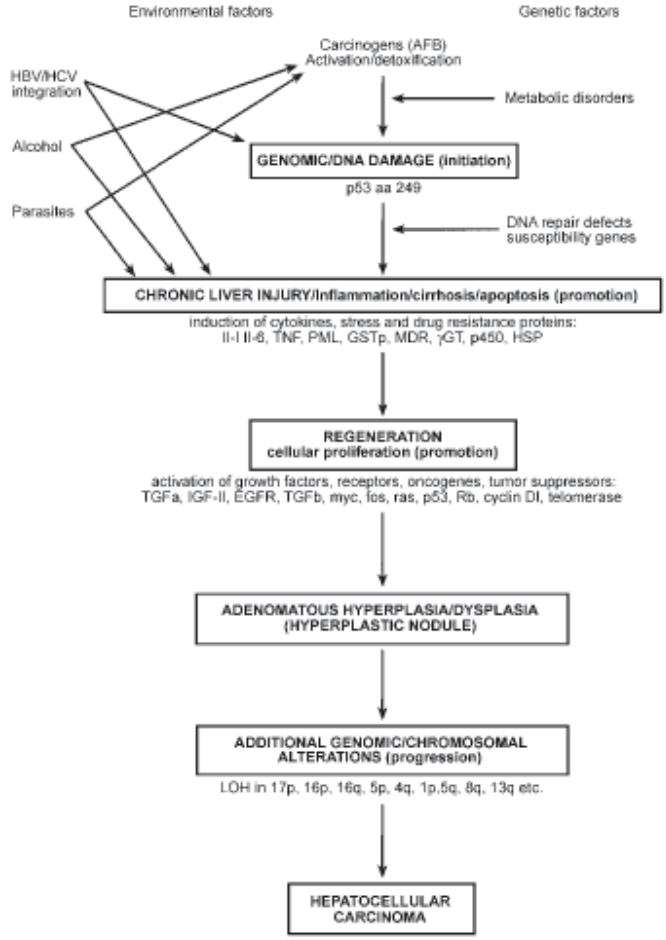


Figure 1.12 Proposed sequence of hepatocarcinogenesis, incorporating known major etiological agents, induction and activation of genes responsible for liver injury and regeneration and supposed genetic changes. LOH = loss of heterozygosity (Leong and Leong 2005).

2003). β-Catenin is important in cell-cell adhesion, and mutations disrupt normal cell-cell interactions and stimulate hepatocyte growth (Cadoret et al., 2001; Feitelson et al., 2002). p53 mutations have also been found in high numbers in HCC patients, particularly G to T mutations of the third base of codon 249 which results in a change from arginine to serine (Bressac et al., 1991; Hsu et al., 1991; Puisieux et al., 1991; Scorsone et al., 1992).

Certain HCV proteins have been demonstrated to play potential roles in the development of hepatocarcinogenesis, namely core, NS3 and NS5A. All three inhibit post-transcriptional expression of p21 (WAF1), a cyclin-dependant kinase inhibitor important in cell cycle control (Ghosh et al., 1999; Kwun et al., 2001; Lee, M. N. et al., 2002). NS3 and NS5A have also been shown to possess direct oncogenic potential as indicated by their ability to promote anchorage independent growth when expressed in fibroblasts and to induce tumor formation in nude mice (Sakamuro et al., 1995; Ghosh et al., 1999). HCV core protein is thought to be particularly important in HCC development. It acts as a transcriptional regulator for a number of host genes, including the proto-oncogene c-*myc* and the c-*fos* promoter, suggesting a role in the deregulation of normal cellular growth (Ray et al., 1995). Core protein can also inhibit a number of apoptosis activators, including Fas and tumor necrosis factor alpha (TNF-α) (Marusawa et al., 1999; Jin et al., 2006), potentially through constitutive activation of a MAPK/ERK signaling cascade (Hayashi et al., 2000). However, the *in vitro* nature of these experiments and over-expression of HCV proteins suggest that these conclusions should be interpreted with caution. A number of transgenic mouse studies have indicated that HCV core protein and possibly other HCV proteins are able to induce HCC in transgenic mice (Moriya et al., 1998; Koike et al., 2002; Lerat et al., 2002).

1.2.3.2 Non-viral causes of HCC

Whilst HBV and HCV are the leading indicators for HCC development, HCC can also be caused by non-viral factors such as ethanol consumption, NASH and aflatoxins. The link between chronic ethanol consumption and liver disease has been well established (Longnecker 1995; Lieber 1997; Leevy and Elbeshbeshy 2005; McKillop and Schrum 2005; Siegmund and Brenner 2005; Voigt 2005), with one recent review attributing 32% of cirrhosis and 25% of HCC worldwide to alcohol (Rehm et al., 2003). During alcoholic liver disease, it is metabolites and by-products of the ethanol metabolism process that are responsible for hepatocyte damage, rather than the ethanol itself (Crabb et al., 1987; Ekstrom and Ingelman-Sundberg 1989; Lu 2000; Wu, D. and Cederbaum 2003). NASH is the most severe form of non-alcoholic fatty liver disease (NAFLD). Biopsy features are identical to those of alcoholic liver disease, however NASH patients by definition have no history of excessive alcohol use (Adler and Schaffner 1979; Ludwig et al., 1980). NASH can remain asymptomatic for many years, but can also progress to cirrhosis and end-stage liver disease including HCC (Lee, R. G. 1989; Powell et al., 1990; Bacon et al., 1994; Younossi et al., 1998; Matteoni et al., 1999). Aflatoxins are mycotoxins produced by *Aspergillus flavus* and *Aspergillus parasiticus* and are the most powerful naturally occurring carcinogens known to man (Smela et al., 2001; Henry et al., 2002). *A. flavus* and *parasiticus* are widespread in nature, contaminating grain, nut and seed crops either before harvest or during storage. There are at least 13 types of aflatoxins, however aflatoxin B_1 (AFB_1) is the most prevalent and most toxic form. AFB_1 is converted to AFB_1-8,9-epoxide by cytochrome p450 in the liver, which can react with guanine nucleotides in hepatocyte DNA to form adducts leading to hepatocyte transformation (Miller 1978; Essigmann et al., 1983; Smela et al., 2001).

1.2.4 Treatment options and survival rates

HCC prognosis is typically poor due to the late presentation and diagnosis of patients with highly advanced cirrhosis, no matter what the causative agent (El-Serag 2002). Life expectancy is measured in weeks or months rather than years, and the mortality to incidence ratio is almost 1 (But et al., 2008). A recent study reports 1, 3 and 5-year survival rates of 66.1%, 39.7% and 32.5% respectively for HCC patients (Yang et al., 2003). Irrespective of cause or geographical location, patients at terminal stages of disease typically survive less than 6 months (Llovet et al., 1999). A number of treatments for HCC have been trialed with varying success, including both curative and palliative. The treatment regime for each individual is dependent on tumor stage at time of diagnosis (Llovet et al., 1999).

Curative treatments for HCC carry greatly varying success rates. Surgical resection of HCC tumors is currently the method of choice and is thought to offer the best chance of a cure, however less than 20% of patients meet the criteria required for successful resection (Bismuth et al., 1999; Song, T. J. et al., 2004). Recurrence of tumors occurs in up to 70% of resected patients by 5 years post surgery (Kumada et al., 1997; Iizuka et al., 2003; Imamura et al., 2003). Liver transplantation is an attractive option for the treatment of HCC, however strict guidelines regarding tumor size and number, lobular distribution, and vascular invasion must be followed to increase the chances of a successful allograft (Mazzaferro et al., 1996; Wong, L. L. 2002). However, recurrence of HCV is extremely common, with most patients developing fibrosis to some degree within 5 years (Crosbie and Alexander 2000). Living donor transplantation is an alternative to deceased donor transplantation, however it is a highly complex procedure with a morbidity rate of 20 - 40% and a donor mortality rate of 0.3 - 0.5% (Cronin et al., 2001; Trotter et al., 2002; Lo et al., 2004). Tumor ablation is a minimally invasive procedure aimed at destroying tumor cells without surgery and is often the best option. Percutaneous ethanol injection (PEI) is

the most commonly used ablation technique and has been used to treat HCC since the mid 1980s (Shiina et al., 1993; Livraghi et al., 1995; Lencioni et al., 2003; Sala et al., 2004).

Chemotherapeutic drugs are limited although Sorafenib was approved by the FDA in 2005 to treat renal cell carcinoma and was granted approval in late 2007 for the treatment of inoperable HCC cases (Lang 2008). Results showed increased time to progression (5.5 months compared to 2.8 months) and increased overall survival (10.7 months compared to 7.9 months) for sorafenib-treated patients compared to placebo for advanced HCC (Llovet et al., 2007). Clearly new therapeutic options are required to treat HCC.

1.2.5 Diagnosis

1.2.5.1 Current diagnostic techniques

By the time clinical symptoms present, HCC tumors are generally well advanced and viable treatment options become limited. As most HCCs occur in patients with cirrhotic livers, screening for HCC in these individuals is recommended in an attempt to diagnose tumors at the earliest possible stage and increase the chance of a successful response to treatment. The main diagnostic techniques employed for detecting HCCs are ultrasound imaging and liver biopsy. The gold standard of diagnosis is liver biopsy, either by fine needle aspiration or core biopsy followed by histological analysis, however biopsy is increasingly being phased out of HCC diagnosis due to the pain experienced by patients undergoing the procedure, the risk of spread of the tumor from the biopsy needle track, and the technological advances in non-invasive diagnostic techniques. Clearly, a non-invasive marker of HCC, such as a reliable diagnostic biomarker, would significantly enhance the success of early HCC diagnosis.

1.2.5.2 Diagnostic biomarkers of HCC

Tumor biomarkers are utilised for a variety of purposes, including screening general populations or people at specific risk for tumor presence, localising tumor burden, monitoring the effect of treatment and detecting tumor recurrence post cure. Markers must provide reliable and reproducible results using simple techniques at minimal cost. Diagnostic biomarkers in particular are useful for identifying tumors in high risk individuals without the need for invasive techniques. They also allow for early detection of tumors before they progress to larger size, later stage and metastases; this is particularly important for HCCs as detection at an early stage heavily affects clinical outcome.

The serum glycoprotein α-fetoprotein (AFP) was first described as a serum biomarker for HCC in the 1960s. It is expressed at high levels in the liver during foetal development but expression decreases rapidly within 300 days of birth (Kashyap et al., 2001). Later increases in serum AFP levels are potentially indicative of malignant activity. HCC tumors can induce elevations in serum AFP to over 100,000 ng/ml (Colombo 2001; Koteish and Thuluvath 2002), however recent investigations have revealed its unreliability as a diagnostic marker. A high false-positive rate is seen in patients with active hepatitis but an absence of HCC (Johnson, P. J. 2001; Peng et al., 2004). Less than 50% of confirmed HCC patients possess AFP levels above that considered diagnostic, and normal AFP levels are seen in up to 30% of HCC patients at time of diagnosis which remain low over time (Colombo 2001). Using the diagnostic value of 400 ng/ml, AFP has a sensitivity of 100% but a specificity of only 45% (Gupta et al., 2003). Steps have been taken to improve the diagnostic accuracy of AFP, including the investigation of certain isoforms that may be specific to HCC such as lens culinaris agglutinin-reactive AFP (AFP-L3) (Taketa et al., 1993; Johnson, P. J. et al., 1997).

Des-γ-carboxy prothrombin (DCP) is a HCC marker that has been used widely in Japan since 1984. Also known as protein induced by vitamin K absence (PIVKA II), it is an abnormal form of prothrombin known to be highly specific for HCC (Liebman et al., 1984). DCP levels are significantly increased in 50 - 60% of total HCC patients but only 15 - 20% of HCC patients with tumors smaller than 3 cm (Weitz and Liebman 1993; Nomura et al., 1996). Using a cut-off value of 40 mAU/ml results in specificity of 74 - 98% and sensitivity of 48 - 86% (Tsai et al., 1990; Fujiyama et al., 1992; Takikawa et al., 1992; Kasahara et al., 1993; Mita et al., 1998; Nakagawa et al., 1999; Ishii et al., 2000; Lok et al., 2009), which are not significantly better than those obtained for AFP.

Intensive research is constantly uncovering novel tumor marker candidates, primarily through the use of high-throughput genomic and proteomic technologies, however many fall short of being useful in a clinical setting. Clearly with the increasing incidence of HCC worldwide, more reliable markers of disease are required to allow for early detection with greater treatment options. Osteopontin (OPN) is one such biomarker with HCC diagnostic potential. This thesis focuses on the role of OPN in HCV-related HCC and as such will be discussed below.

1.3 Osteopontin (OPN)

1.3.1 Introduction

Osteopontin (OPN) was first identified in 1979 as a 60 kDa transformation-specific phosphoprotein secreted from malignant epithelial cells (Senger et al., 1979). It is a highly phosphorylated sialoprotein and a member of the small integrin-binding ligand N-linked glycoprotein (SIBLING) family. The SIBLING family includes other proteins such as dentin matrix protein 1, dentin sialophosphoprotein, bone sialoprotein and matrix extracellular phosphoglycoprotein (Fisher, L. W. et al., 2001). Most SIBLING proteins are highly expressed in

a number of cancers, with expression level directly correlated with tumor grade (Fedarko et al., 2001). OPN has been independently discovered by numerous research groups, and referred to by various names, due to its many functions. Its identification as the major phosphoprotein secreted by cultured cells led to the term secreted phosphoprotein 1 (SPP1) (Wrana et al., 1989; Craig et al., 1990), and its cytokine activity when produced by activated lymphocytes and macrophages led to the term early T-cell activation factor (Eta-1) (Patarca et al., 1989). It has also, at one time or another, been known as bone sialoprotein 1, uropontin, and 2ar (Fisher, L. W. et al., 1987; Craig et al., 1988; Shiraga et al., 1992; Weber and Cantor 1996), however osteopontin is now the standard nomenclature used for all the aforementioned proteins.

1.3.2 Molecular structure

OPN has been identified in a number of species, including humans, mice, rats, chickens, pigs, cows, hamsters and chimpanzees, with a high level of inter-species sequence homology (Denhardt and Guo 1993; Patarca et al., 1993; Sodek et al., 2000). The single-copy gene is located on human chromosome 4q21-25 and encodes 7 exons extending over approximately 11 kilobases (Figure 1.13) (Hijiya et al., 1994; Sodek et al., 2000). Exons 2 - 7 are translated, giving a coding sequence of approximately 954 base pairs. Alternative splicing of OPN RNA generates three distinct splice variants (Young et al., 1990; Saitoh et al., 1995); these will be discussed in greater detail in Section 1.3.5. A protease hypersensitive site separates the CD44 and integrin binding domains; within this region sits a thrombin cleavage motif that contains a conserved arginine-serine-lysine (RSK) amino acid sequence (Bautista et al., 1994). Numerous regulatory sequences are found upstream of the coding sequence: a TATA-like sequence (Sodek et al., 2000), an interferon regulatory factor-1 (IRF-1) binding element (Korber et al., 1988), two vitamin D responsive elements (VDREs), numerous CCAAT-like sequences, and other possible

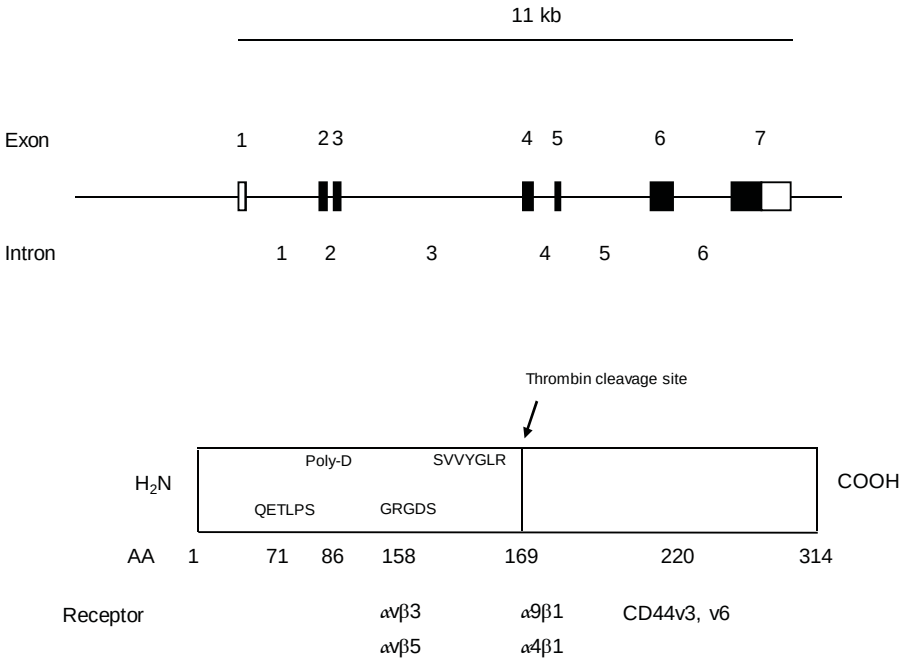


Figure 1.13 OPN gene structure and binding domains. Top: The full length gene product contains 7 exons and 6 introns; exon 1 and part of exon 7 are not translated (shown as unfilled boxes). Bottom: Various receptor binding and sequence domains are present. The thrombin cleavage site is located at amino acid 169. AA = amino acid position.

binding sites for various transcription factors such as AP2, EPA, Tcf, Myb, E2BP and Ets-1 (Sodek et al., 2000; Suzuki, K. 2005).

The translated protein consists of a single chain of 264-333 amino acids rich in serine, sialic acid and glutamic acid residues depending on species (Sodek et al., 2000). The human protein is thought to comprise 314 amino acids (Young et al., 1990; Saitoh et al., 1995; O'Regan, A. and Berman 2000). It contains a glycine-arginine-glycine-aspartic acid-serine (GRGDS) signaling sequence which is an integrin-binding motif that can mediate cell attachment and is common to many extracellular matrix proteins (Sodek et al., 2000). OPN also contains a polyaspartic acid motif that can bind to hydroxyapatite and calcium ions, and a leucine-proline-valine (LPV) motif at the N terminus (Sodek et al., 2000; Suzuki, K. 2005). The predicted secondary structure of OPN possesses 8 α-helices and 6 β-sheets (Figure 1.14) (Denhardt and Guo 1993; Sodek et al., 2000). OPN undergoes extensive post-translational modifications such as glycosylation, sulphation, phosphorylation and enzymatic cleavage. Glycosylation is responsible for approximately a third of OPN's weight in most tissues (Oldberg et al., 1986), and sulphation occurs commonly to highly phosphorylated forms of the protein (Nagata et al., 1989). Phosphorylation occurs at up to 28 sites on the protein (Prince et al., 1987). OPN has a molecular weight of 41 - 75 kDa depending on which cell-specific and condition-specific post-translational modifications it undergoes (Denhardt and Guo 1993). A thrombin cleavage site is situated near the GRGDS sequence which is important for OPN function as thrombin-cleaved OPN promotes greater cell attachment and spreading than intact OPN (Sodek et al., 2000; Suzuki, K. 2005).

1.3.3 Expression, distribution and regulation of OPN

OPN is widely expressed in the human body, with protein detected in a range of tissues including brain, lungs, bone, heart, liver, gastrointestinal tract, kidneys and joints, and various bodily fluids

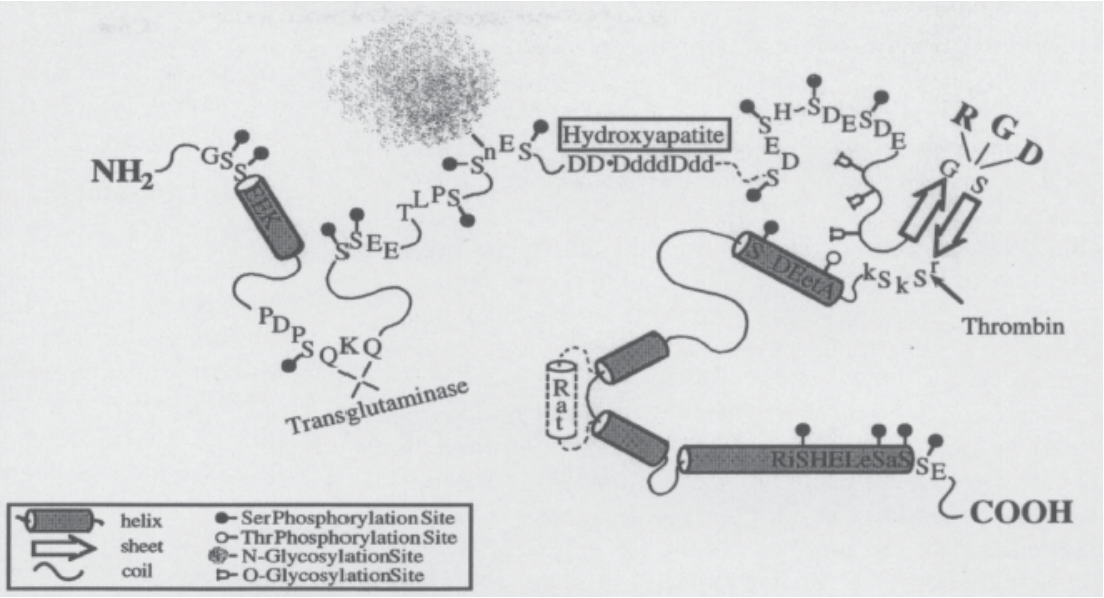


Figure 1.14 Proposed OPN secondary structure, based on the amino acid sequence and secondary structure predictions. Alpha helices, beta sheets and post-translational modifications (phosphorylation and glycosylation sites) are indicated (Sodek et al 2000).

including serum, plasma, urine, breast milk, bile and seminal fluid (Sodek et al., 2000). Expression is observed in a number of cell types, including epithelial cells, endothelial cells, osteoblasts, osteoclasts, nerve cells and vascular smooth muscle cells (Craig and Denhardt 1991; Denhardt and Guo 1993; Denhardt and Noda 1998; O'Regan, A. and Berman 2000; Weber 2001). Activated immune cells such as Küpffer cells, macrophages, natural killer T (NKT) cells, and T cells also express OPN. OPN expression increases in response to stress such as cell or tissue injury, and is heavily involved in related processes such as cell survival, tissue remodeling and inflammation (O'Regan, A. and Berman 2000; Denhardt, Noda et al., 2001).

High levels of OPN are expressed in a variety of human tumors, including breast, lung, oesophagus, stomach, prostate, oral cavity, central nervous system, colon, bladder, ovary, thyroid, pancreas, gastric and liver cancers (Senger et al., 1988; Brown et al., 1994; Chambers et al., 1996; Casson et al., 1997; Tuck et al., 1998; Ue et al., 1998; Devoll et al., 1999; Thalmann et al., 1999; Hirama et al., 2003), and has been shown to be important for tumor growth *in vivo* (Behrend et al., 1994; Gardner et al., 1994; Su et al., 1995), however expression was initially thought to occur exclusively in tumor infiltrating macrophages (Furger et al., 2001). Further studies have also visualised expression within the tumor cells themselves (Rittling and Chambers 2004), with OPN produced by macrophages and tumor cells appearing to regulate different signaling pathways (Crawford, H. C. et al., 1998). OPN has been detected in cancerous cells of the kidneys, colon, skin, breast, brain and liver (Brown et al., 1994; Oates et al., 1997; Takano et al., 2000; Gotoh et al., 2002; Zhou et al., 2005).

Cytokines, hormones (such as Vitamin D_3 and oestrogen) and growth factors can all regulate the expression of OPN (Sodek et al., 2000). Epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factor-β and various inflammatory cytokines such as

interleukin-1 (IL-1) and tumor necrosis factor alpha (TNF-α) all increase OPN expression (Denhardt and Noda 1998; Chatterjee 2001). The proto-oncogene Src regulates OPN expression in both neoplastic and non-neoplastic tissues (Levin, V. A. 2004). AP-1 stimulates expression of OPN by binding to an AP-1 cis-regulating site found within the OPN promoter (El-Tanani, M. et al., 2004), and it has also been recently found that OPN can induce AP-1 transcriptional activity through induction of c-Src kinase and subsequent phosphorylation of EGRF, resulting in a positive feedback loop whereby OPN can stimulate its own expression (Das et al., 2004). Other transcriptional regulators of OPN include Myc, TCF-4, OCT-1, TP53, USF, TGF-β/Smad-3, TGF-β/Hoxa-9, and v-Src/CBF-like factor (Wai and Kuo 2004).

1.3.4 OPN receptors

1.3.4.1 CD44

CD44 is a ubiquitously expressed transmembrane glycoprotein found on the cell surface that acts as a structural link between the cell cytoskeleton and the ECM (Georgolios et al., 2006). It also acts as a signaling molecule with roles in cellular adhesion and migration (Georgolios et al 2006). The gene contains 19 exons; 10 constant exons (c1-10), 5 each at the carboxy and amino termini, and 9 variant exons (v2-10) situated between exons c5 and c6. Variant exon v1 has been described in rats but is absent in the human homologue (Figure 1.15). CD44 undergoes extensive splicing and post-translational modification, giving rise to a number of variants containing all 10 constant exons and differing combinations of variant exons (Goodison et al., 1999). The mechanism controlling the alternative splicing of CD44 is unknown, however the *ras* oncogene has been reported to regulate the CD44 promoter (Hofmann et al., 1991; Gallagher 1992).

CD44 binds a number of ligands, including hyaluronic acid (its primary ligand), collagen, fibronectin, laminin, serglycin, chondroitin sulfate and OPN. OPN specifically interacts with

Figure 1.15 CD44 gene structure. Standard exons are illustrated in green filled boxes and variant exons in yellow filled boxes. Top: Rat CD44 has 10 standard exons (c1-10) and 10 variant exons (v1-10). Bottom: Human OPN has the same 10 standard exons but is lacking variant exon 1. Exons 1-17 are located on the extracellular side of the membrane, exon 18 represents the transmembrane region, and exons 19-20 are situated within the cell cytoplasm (adapted from Liu and Jiang 2006).

CD44 variants 6 and/or 7 (CD44v6/7) to induce chemotaxis in an RGD-independent manner, and may also bind CD44v3 via a heparin bridge (Weber et al., 1996; Katagiri et al., 1999; Ashkar et al., 2000). Studies have shown that CD44v7-OPN interaction mediates chemotaxis and adhesion of fibroblasts, bone marrow cells and T cells (Weber et al., 1996; Katagiri et al., 1999; Denhardt, Giachelli et al., 2001; Weber 2001), affects the host inflammatory response by down-regulating IL-10 expression in murine peritoneal macrophages (Ashkar et al., 2000), and confers metastatic behaviour on rat carcinoma cells (Gunthert et al., 1991; Rudy et al., 1993). These data suggest that OPN activates CD44v6 and/or v7-mediated pathways to promote cell homing, adhesion, survival and chemotaxis, with these pathways enhancing metastatic behaviour in tumors. Interactions between OPN and CD44 will be discussed in greater detail in Chapter 4.

1.3.4.2 Integrins

Integrins are heterodimeric cell surface receptors that interact with the ECM and mediate intracellular signaling. They are type I transmembrane glycoproteins, formed by non-covalent associations between α and β subunits, and incorporate a large extracellular domain and short cytoplasmic tail (Wai and Kuo 2004). In mammals there are 18 α and 8 β subunits, which associate in different combinations to give 24 distinct integrins with specific functions (Hynes 2002). The functions of all integrins come under one of two umbrellas: cell-cell or cell-ECM adhesion, or signaling from the ECM to the cell. Tumor cells that express integrins can induce constitutive activation of signaling pathways leading to increased growth (Wai and Kuo 2004).

The integrin family is able to bind a wide range of ligands, including fibronectin, vibronectin, fibrinogen, thrombospondin and OPN (Gerber et al., 1996). Known integrin receptors for OPN include $\alpha_v\beta_1$, $\alpha_v\beta_3$, $\alpha_v\beta_5$, $\alpha_4\beta_1$, $\alpha_5\beta_1$, $\alpha_8\beta_1$ and $\alpha_9\beta_1$ (Hu et al., 1995; Liaw et al., 1995; Smith et al., 1996; Bayless et al., 1998; Denda et al., 1998; Barry et al., 2000); $\alpha_v\beta_3$ is thought to be the

primary receptor but can only bind OPN after thrombin cleavage and phosphorylation of the N-terminal fragment (Nau et al., 1999; Ashkar et al., 2000). Tumor cells capable of binding to OPN via $\alpha_v\beta_3$ integrin may possess a survival advantage over other tumor cells (Furger et al., 2001; Wai and Kuo 2004). Interactions between OPN and $\alpha_v\beta_3$ integrin instigate up-regulation of CD44v6 expression in an auto-stimulatory feedback loop (Gao et al., 2003), resulting in increased tumor cell survival, cellular adhesion (of platelets, osteoclasts, smooth muscle cells and B cells), homing, chemotaxis and metastasis (Reinholt et al., 1990; Liaw et al., 1994; Bennett et al., 1997; Georgolios et al., 2006). OPN-$\alpha_v\beta_3$ binding is thought to be heavily involved in tumor metastasis, responsible for directly mediating tumor cell migration and invasion (Senger and Perruzzi 1996; Angelucci et al., 2002; Furger et al., 2003), and altering host immunity by increasing macrophage IL-12 expression in a mouse model (Ashkar et al., 2000). This interaction is also responsible for regulating IFN-λ expression in NKT cells (Rabinowich et al., 1995), activating osteoclasts and thus representing a lytic mechanism in bone metastases (Engleman et al., 1997; Weber 2001), and inducing the formation of new blood vessels within the tumor by up-regulating endothelial cell migration, survival and lumen formation (Brooks et al., 1994; Senger et al., 1996; Scatena et al., 1998; Bayless et al., 2000). Vascular endothelial growth factor (VEGF) is known to induce OPN and $\alpha_v\beta_3$ expression in microvascular endothelial cells, which gives further credence to the role of OPN-$\alpha_v\beta_3$ interaction in angiogenesis (Senger et al., 1996; Denhardt, Noda et al., 2001). These studies indicate that interaction between $\alpha_v\beta_3$ and OPN is essential for numerous pathways involved in tumor growth, survival, angiogenesis, and metastasis.

1.3.5 Alternatively spliced variants of OPN

Not long after the OPN gene was cloned in 1989, an alternatively spliced form lacking 14 amino acids from residue 58 onwards was found in human bone and decidual cells (Kiefer et al., 1989;

Young et al., 1990). Another alternatively spliced variant was discovered in human glioma in 1995 (Saitoh et al., 1995), resulting in what are now recognised as the three variants of OPN: OPN-A, which encodes the entire protein, OPN-B which lacks the 42 base pair exon 5, and OPN-C which lacks the 81 base pair exon 4 (Figure 1.16). Very little investigation has been done into the differential roles of the two alternatively spliced variants and the significance of the missing exons to their structure and function; this is addressed in part in this thesis. He *et al.* (2006) reported that variant OPN-C strongly supported anchorage-independent growth of breast cancer cells and facilitated soft agar colony growth in a non-invasive breast cancer cell line that does not express OPN (MCF-7) better than the full length protein (OPN-A). Soft agar colony size of MCF-7 cells over-expressing OPN from transfected vectors was significantly larger for OPN-C than OPN-A, and both were significantly larger than those from cells transfected with empty vector. Interestingly, OPN-A but not OPN-C supported calcium-dependant adhesion of MCF-7 cells, and no difference in cellular proliferation was visualised between MCF-7 cells over-expressing OPN–A or OPN-C (He et al., 2006). This study allows some insight into differing functions of OPN-C in breast cancer resulting from the lack of exon 4, but does not suggest any alternative function for OPN-B. Further study is needed to fully elucidate the functions of OPN-B and OPN-C and how they differ from those already established for the full length protein (OPN-A) in different cell lines.

1.3.6 Functions of OPN and its role in tumorigenesis

OPN is involved in a wide variety of normal physiological functions, including inflammation and immunity (Denhardt, Noda et al., 2001; Diao et al., 2004), angiogenesis (Wu, Y. et al., 2000), bone remodeling (Ishijima et al., 2001; Chellaiah, M. A. et al., 2003; Shapses et al., 2003) and apoptosis inhibition (Fedarko et al., 2000; Lin, Y. H. and Yang-Yen 2001; Xie, Z. et al., 2003; Zohar et al., 2004). OPN is known to mediate cell adhesion and migration (El-Tanani, M. K. et

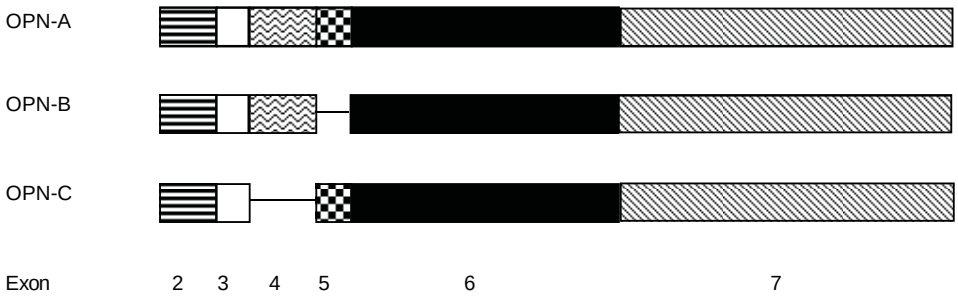


Figure 1.16 The three splice variants of OPN. OPN-A represents the full length protein, whilst OPN-B and OPN-C each lack an exon (5 and 4 respectively).

al., 2006), regulate cytokine production from macrophages, and act as a survival factor (Denhardt, Giachelli et al., 2001). It acts as an early cytokine by mediating cellular signaling through interactions with CD44 (chemotaxis) and integrins (haptotaxis) (Senger et al., 1994; Weber et al., 1996; Ashkar et al., 2000). The most widely researched functions of OPN however are its many roles in oncogenesis and cancer progression, including tumor angiogenesis, tumor cell proliferation and metastasis.

The formation of new blood vessels is imperative for tumor growth and metastasis, and OPN can promote neovascularisation through integrin-mediated endothelial cell migration, prevention of endothelial cell apoptosis, and vascular lumen formation (Brooks et al., 1994; Senger et al., 1996; Scatena et al., 1998; Bayless et al., 2000; Khan et al., 2002; Hirama et al., 2003; Leali et al., 2003). OPN and $\alpha_v\beta_3$ integrin are thought to work synergistically during angiogenesis as they are coordinately upregulated after vascular damage and are thought to be key regulators of repair and regeneration of blood vessels (Liaw et al., 1995). VEGF is also thought to contribute to angiogenesis by cooperating with OPN and $\alpha_v\beta_3$ to stimulate migration of endothelial cells (Senger et al., 1996).

The role of OPN in tumor cell proliferation is a controversial topic. Studies have shown a positive effect of OPN on the proliferation of human prostate cancer cells (Elgavish et al., 1998; Angelucci et al., 2004), colon cancer cells (Irby et al., 2004), murine melanoma cells (Philip et al., 2001), cultured rat vascular smooth muscle cells and human coronary artery smooth muscle cells (Panda et al., 1997), whilst inhibition of OPN with anti-sense sequences attenuated growth of breast cancer cells (Wu, Y. et al., 2000). Conversely, OPN has been shown to inhibit proliferation in a pre-osteoblastic cell line (Huang et al., 2004), and no correlation has been found between OPN expression and proliferation of fibroblast-like cells (Perrien et al., 2002) or breast

cancer cells (He et al., 2006) or the growth rate of primary tumors in mouse mammary tumor virus gene (MMTV)-c-myc/MMTV-v-*Ha-ras* transgenic mice (Feng and Rittling 2000). Structural differences between OPN products produced by distinct tumors or variances in tumor microenvironments may explain the contradictory roles of OPN in cellular proliferation. The role of OPN in cellular proliferation is the focus of Chapter 4 of this thesis, and will be discussed in more detail therein.

OPN's role in tumor metastasis is it's most widely researched and well-described function. A strong correlation between high OPN expression levels and malignant invasion has been seen in a number of human cancers (Senger et al., 1988; Brown et al., 1994; Bellahcene and Castronovo 1995; Singhal et al., 1997; Tuck et al., 1998; Gotoh et al., 2002). Plasma OPN levels are significantly higher in patients with metastatic disease than non-diseased individuals (Senger et al., 1988; Singhal et al., 1997; Fedarko et al., 2001). Functional studies have shown that OPN can induce a metastatic phenotype on non-metastatic or poorly metastatic cells in culture (Oates et al., 1996; Chen, H. et al., 1997; Yoneda et al., 1998), induce migration and invasion of human mammary epithelial cells (Tuck et al., 1999; Tuck et al., 2000), and induce lung metastases after injection of benign tumor-producing cells into mammary fat pads of Furth-Wistar rats (El-Tanani, M. et al., 2001). OPN has also been shown to mediate tumor invasion by regulating a number of ECM-degrading proteins to allow tumor cells to migrate and invade neighbouring tissues (Wai and Kuo 2004), including pro-matrix metalloprotease-2 (pro-MMP-2) and urokinase plasminogen activator (uPA) (Philip et al., 2001; Tuck et al., 2003).

1.3.7 Downstream signaling pathways regulated by OPN

OPN is involved in the stimulation and regulation of numerous signaling pathways, however its specific involvement is thought to be determined by the cellular source of OPN expression,

especially in regards to tumor pathophysiology. OPN is expressed within tumors by both resident tumor cells and infiltrating macrophages (Brown et al., 1994; Chambers et al., 1996; Yeatman and Chambers 2003) and appears to perform different functions depending on the cell type it is expressed from; macrophage-derived OPN recruits immune cells and suppresses tumor growth, whereas tumor cell-derived OPN inhibits macrophage function and enhances tumor growth (Crawford, H. C. et al., 1998).

A number of down-stream OPN signals are thought to be responsible for regulating tumor progression and invasion, including pathways involved in interrupting cell cycle progression, promoting cellular survival and preventing apoptosis (Figure 1.17) (Evan and Vousden 2001; Rangaswami et al., 2006). OPN has been shown to activate the phosphatidylinositol 3-kinase/protein kinase B (PI3K/Akt) pathway, thereby preventing apoptosis and conferring IL-3-mediated survival advantage on murine B cells (Lin, Y. H. and Yang-Yen 2001). The PI3K/Akt pathway is also a known mediator of osteoclast activation following interaction with OPN (Hruska et al., 1995; Chellaiah, M. and Hruska 1996). OPN can also induce PI3-kinase activity and PI3-kinase-dependant Akt-phosphorylation via $\alpha_v\beta_3$ interactions in breast cancer (Das et al., 2003). OPN is able to regulate the activity of the ECM degrading proteins pro-MMP-2 and uPA, both of which play roles in proteolysis and remodeling of the ECM, allowing cancerous cells to invade surrounding stroma. OPN up-regulates pro-MMP-2 expression in an NF-κB-dependant manner during ECM invasion; OPN-induced MMP expression is blocked by transfection with the NF-κB inhibitor IκBα, while anti-sense MMP-2 oligonucleotides reduce OPN-mediated migration and ECM invasion in murine melanoma cells (Philip et al., 2001). Cell invasiveness is also increased by stimulation of uPA expression by OPN in human mammary carcinoma (Tuck et al., 2001), with the uPA pathway is also upregulated in breast cancer patients

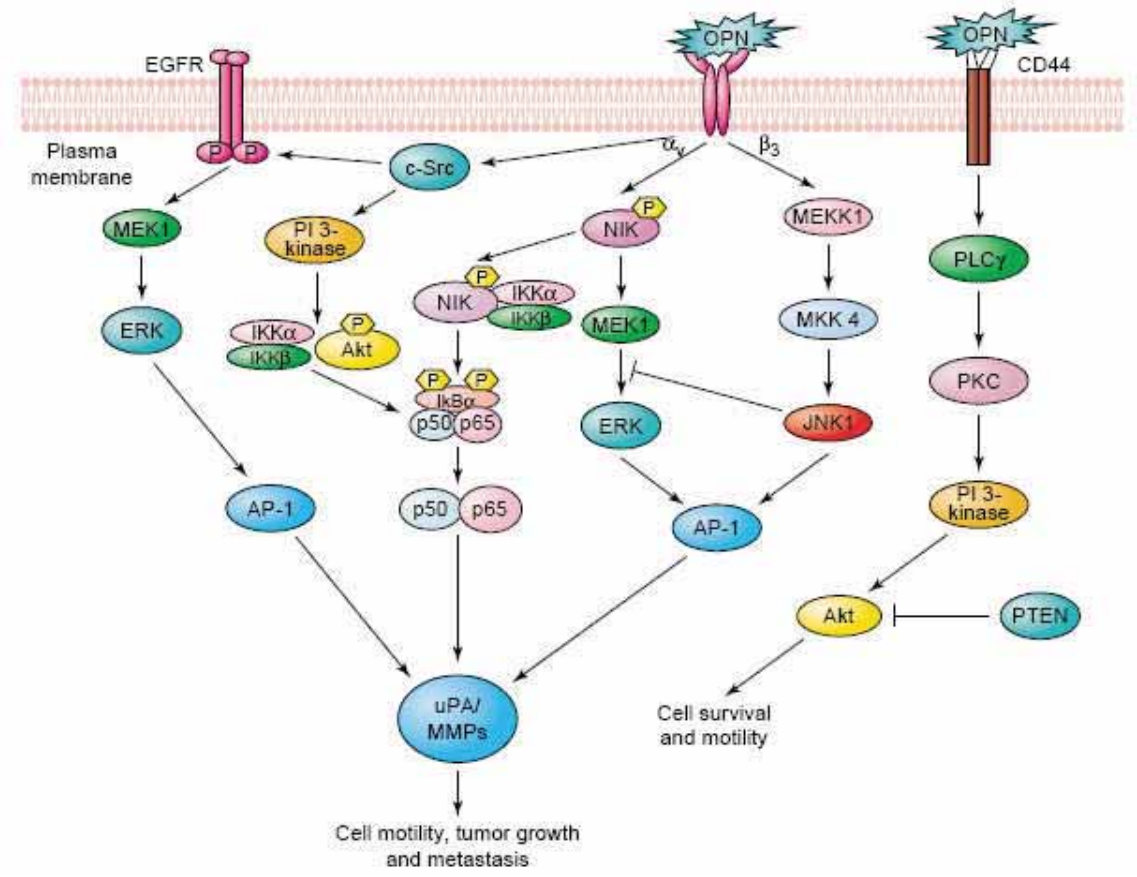


Figure 1.17 Mechanisms of OPN-induced tumor growth and metastasis via CD44 and $\alpha_v\beta_3$ mediated pathways (Rangaswami et al., 2006).

with poor prognosis, malignant cancers, and bone metastases (Andreasen et al., 1997; Fisher, J. L. et al., 2000).

OPN is also thought to be involved in other processes essential for tumor metastasis, including tumor cell migration through the ECM. The EGF and hepatocyte growth factor (HGF) signaling pathways are thought to be important mediators of OPN-regulated cell migration. During OPN induction of breast cancer cell migration, OPN induces EGF receptor (EGFR) mRNA expression and tyrosine kinase activity, HGF receptor (Met) mRNA and protein expression, and Met kinase activity (Tuck et al., 2003). OPN also down-regulates inducible nitric oxide synthase (iNOS) activity, which in turn aids tumor invasiveness and metastasis by altering host immunity (Denhardt and Chambers 1994; Scott et al., 1998; Nagasaki et al., 1999; Guo et al., 2001).

1.3.8 Osteopontin as a biomarker of disease and predictor of poor prognostic outcome in cancer

As previously mentioned, OPN is over-expressed in a variety of human and murine tumors. High levels of OPN have been reported in the sera of tumor-bearing mice, but low levels seen within the tumors themselves (Rittling and Novick 1997), indicating that OPN produced by tumors is readily secreted. Serum OPN concentration has been shown to increase with increasing tumor burden (Kon et al., 2000; Fedarko et al., 2001). Its secreted nature and relative ease of detection in human serum makes OPN a prime candidate as a biomarker of cancer.

Increased OPN expression in cancer is correlated with poor clinical outcome (Tuck et al., 1998). In a study of 333 inoperable stage I and II breast cancer patients, OPN negative specimens correlated with a median survival time from diagnosis of greater than 228 months, compared to 68 months for OPN-positive specimens (Rudland et al., 2002). Another study of 25 lung cancers

indicated that positive immunostaining of OPN in tumor tissue resulted in significantly poorer patient survival (Chambers et al., 1996). Increased OPN expression is also associated with advanced stage and poor outcome in colon, thyroid, gastric, prostate and liver cancer (Fedarko et al., 2004).

Due to the high expression levels in the sera of cancer patients, OPN is an attractive disease biomarker candidate for cancer diagnosis and prediction of prognostic outcome. cDNA microarray technology has identified OPN as a potential biomarker of ovarian cancer (Wong, K. K. et al., 2001). A number of groups have utilised enzyme linked immunosorbent assay (ELISA) techniques to monitor serum levels of OPN in patients with cancer compared to healthy individuals and evaluate its potential as a biomarker (Table 1.1). Plasma OPN levels are significantly higher in preoperative epithelial ovarian cancer patients (Brakora et al., 2004) and metastatic breast cancer patients (Singhal et al., 1997) compared to healthy individuals. In the breast cancer study, patients were divided into one of three groups based on OPN concentration (< 118 μg/L, 118 - 203 μg/L or > 203 μg/L), with Kaplan-Meier adjusted survival estimates concluding that survival time was inversely proportional to plasma OPN concentration (median survival days of 650, 420 and 170 respectively for the three groups). Associations between increasing OPN concentration and decreasing survival or number of metastatic sites were also found by applying a Cox proportional hazards model with OPN as a continuous variable (Singhal et al., 1997).

1.3.9 OPN and HCV-related liver disease

1.3.9.1 OPN expression and function during liver injury

OPN is expressed in bile duct epithelial cells in the healthy liver (Brown et al., 1992; Kawashima et al., 1999; Lee, S. H. et al., 2004; Tajiri et al., 2005). Transcripts have also been identified in

Authors	Publication year	Cancer type	Disease patients (n)	Control patients (n)	Disease OPN (ng/ml)	Control OPN (ng/ml)
Singhal et al.	1997	Breast	70	35	142	47
Le et al.	2003	Head and neck squamous cell carcinoma	54	15	447	318
Koopman et al.	2004	Pancreatic adenocarcinoma	50	22	482	204
Shimada et al.	2004	Oesophageal squamous cell carcinoma	103	n/a	>450	<450
Bramwell et al.	2006	Breast	157	n/a	177	n/a
Reiniger et al.	2006	Uveal melanoma	8	8	170.72	46.78
Kim et al.	2006	HCC	62	60	954	155
Wu et al.	2007	Gastric cancer	132	93	166.5	55.6
Ramankulov et al.	2007	Renal cell carcinoma	80	52	115	28.9
Chang et al.	2007	Non-small cell lung cancer	47	32	158.2	86.2
Ramankulov et al.	2007	Prostate cancer	28	29	1723	625

Table 1.1 Previous studies investigating the potential of plasma OPN concentration as a biomarker of tumor progression and metastasis. In all studies, control patients were healthy non-diseased individuals with the exception of Chang et al. who compared Stage IV non-small cell lung cancer patients (“disease patients”) to Stage I patients (“control patients”). Adapted from Wai and Kuo 2008.

mouse hepatocytes and mouse and rat hepatic stellate cells, however protein expression was not observed in these cells (Lorena et al., 2006). OPN is regularly found at sites of inflammation caused by tissue insult (Denhardt, Giachelli et al., 2001), with expression amplified mainly in activated macrophages (Kawashima et al., 1999; Wang, Y. et al., 2000). OPN mRNA levels are low in normal mouse and rat liver, however induction of fibrosis via bile duct ligation, CCl_4 treatment or heat-killed propionibacterium Acnes results in significant increases in OPN mRNA or protein expression, mainly in Küpffer cells, macrophages and HSCs (Kawashima et al., 1999; Wang, Y. et al., 2000; Lorena et al., 2006). Serum OPN in patients with acute liver dysfunction arising from a number of causes (both viral and non-viral) is significantly increased over patients with no liver disease (631 ng/ml versus 151 ng/ml), with higher OPN levels (≥ 380 ng/ml) resulting in lower probability of survival (approximately 56% survival compared to approximately 77% survival for those with OPN levels < 380 ng/ml, $P = 0.041$) (Arai et al., 2006). Another study found that mean plasma OPN levels in patients with fulminant hepatitis (defined as acute liver failure with hepatic encephalopathy of at least grade II and characterised by massive or sub-massive necrosis and inflammatory cell infiltrate; 3260 ng/ml) were significantly higher than those with chronic hepatitis infection alone (436 ng/ml, $P < 0.05$) and healthy individuals (239 ng/ml, $P < 0.05$), indicating that OPN expression may increase as liver disease progresses (Matsui et al., 2004). These studies all indicate that liver inflammation and fibrosis from a variety of causes can result in increased tumor OPN expression and secretion.

Little is known about the role of OPN during liver injury and fibrosis. Infection with HCV results in inflammatory liver disease with a massive influx of neutrophils and lymphocytes (NKTs and T-cells) followed by hepatocyte necrosis (Fujii et al., 2005; Zhu et al., 2007). OPN is expressed by activated T cells and is thought to act as an essential cytokine during the Th1 immune response, aiding in the infiltration and activation of macrophages during inflammation (Ashkar et

al., 2000; Diao et al., 2004). OPN has been recently described as having a role in neutrophil chemotaxis (Koh et al., 2007). It may also promote adhesion of T cells, amplify a CD3-mediated proliferative response (O'Regan, A. W. et al., 1999), and promote migration, activation and differentiation of dendritic cells towards a Th1-promoting phenotype (Renkl et al., 2005). However, the role of OPN in HCV-related inflammation has not yet been studied.

The role of OPN in NKT-mediated liver disease was recently confirmed in a mouse model of concavalin-A-induced hepatitis (Diao et al., 2004). Diao *et al.* found that significantly less liver damage was seen in OPN deficient mice ($OPN^{-/-}$) compared to wildtype controls. The OPN deficient mice had lower serum transaminase levels and increased survival rates. NKT cells were found to be the major T-cell phenotype in the concavalin-A-induced hepatitis seen in the control group, with concurrent liver damage significantly decreased in the OPN knockout mice. OPN is also thought to be involved in recruiting macrophages to the liver. Subcutaneous injection of purified OPN protein into mice induced massive infiltrate of macrophages (Singh et al., 1990); this infiltrate also included neutrophils, which corresponds with the findings of Diao *et al* (2004). Macrophage accumulation in mice injected with purified OPN can be reduced by 60% by infusion with a neutralising OPN antibody (Giachelli et al., 1998).

1.3.9.2 OPN expression in patients with chronic HBV and HCV infections

OPN is known to be highly expressed in the livers and sera of patients with HCC (Gotoh et al., 2002; Pan et al., 2003; Kim, J. et al., 2006), however expression levels in patients with HBV or HCV infection prior to the development of HCC are less well defined. In a study of 94 patients with HBV infection, significantly increased OPN expression was seen in patients with HBV-related cirrhosis and HCC compared to healthy individuals and those with non-cirrhotic HBV infection (Zhao, L., Li et al., 2008). A significant increase was also seen in HCC patients

compared to those with HBV-related cirrhosis. Increased serum OPN levels have also been seen in patients with HCV infection compared to healthy individuals (Libra et al., 2005), however any correlation to HCV-related HCC was not investigated. Differences in OPN expression between HCV-infected patients with varying levels of liver disease (pre-HCC in particular) represents a gap in our current knowledge and will be addressed in Chapters 3 and 6 of this study.

1.3.9.3 OPN expression in HCC and its potential use as a diagnostic and prognostic biomarker

As OPN is highly expressed in HCC, it is increasingly being touted as not only a diagnostic or prognostic HCC biomarker but also a potential therapeutic target. OPN expression has been visualised in both tumorigenic hepatocytes (Gotoh et al., 2002; Pan et al., 2003; Kim, J. et al., 2006; Xie, H. et al., 2007) and macrophages (Kim, J. et al., 2006) within HCC tumors, however little is known about the effect of this differential expression during HCC development and progression.

As previously discussed in Section 1.2.5.2, there is a need for a more reliable and effective diagnostic marker for HCC. A number of studies have investigated the use of OPN in the diagnosis of HCC, either in conjunction with or in place of currently used markers (including AFP). Plasma OPN has been found to be expressed at significantly higher levels in patients with HCC (from a mixture of etiologies including HCV and HBV infection, alcoholic liver disease and other causes) than those with chronic liver disease ($P < 0.001$) or non-diseased livers ($P < 0.001$) (Kim, J. et al., 2006). This study compared the effectiveness of OPN, AFP and PIVKA II as possible diagnostic markers of HCC. Both OPN (955 ng/ml versus 381 ng/ml, $P < 0.0001$) and AFP (155 ng/ml versus 6 ng/ml, $P < 0.0001$) showed significantly increased expression in HCC patients over those with chronic liver disease, whilst PIVKA II showed no

difference. The sensitivity and specificity of each marker was determined using receiver operating characteristics (ROC) analysis and is summarised in Table 1.2.

NOTE:
This table is included in the print copy of the thesis held in the University of Adelaide Library.

Table 1.2 Diagnostic accuracy of tumor markers for HCC from chronic liver disease (Kim, J. et al., 2006)

OPN had the highest sensitivity but lowest specificity of all tumor markers tested, and the largest area under curve value when diagnostic efficiency was analysed using a ROC curve (Figure 1.18), indicating that OPN is the most diagnostically accurate of the three. Furthermore, OPN possessed a sensitivity of 93.5% and specificity of 95% when discriminating between HCC and non-diseased individuals (at a cut-off value of 534.5 ng/ml), and 93.5% and 84.2% when discriminating between HCC and non-HCC (chronic liver disease plus non-diseased) patients (at a cut-off value of 552.9 ng/ml), indicating that OPN concentration could be used to distinguish between HCC-positive and HCC-negative individuals.

Increased plasma and serum OPN levels have also been shown to be predictive of HCC recurrence following tumor resection. Zhang *et al.* (2006) analysed preoperative plasma levels of OPN in 101 HCC patients undergoing tumor resection and found that HCC patients had significantly higher plasma OPN than healthy controls (176.9 ng/ml versus 63.74 ng/ml, $P < 0.001$), and that patients who suffered a recurrence of HCC following resection had significantly higher levels of OPN than those who did not (213.55 ng/ml versus 153.60 ng/ml,

NOTE:
This figure is included in the print copy of the thesis held in the University of Adelaide Library.

Figure 1.18 ROC analysis of the diagnostic effectiveness of OPN, AFP and PIVKA II. Area under curve (AUC) values are displayed for each marker. A higher AUC value is indicative of diagnostic superiority (Kim et al., 2006).

P = 0.0013) (Zhang et al., 2006). They also determined that a higher level of plasma OPN (≥ 200 ng/ml) was predictive of poorer overall survival (38.46% versus 77.12%) and disease-free survival (16.26% versus 59.02%) over patients with preoperative OPN plasma levels less than 200 ng/ml. Another study by Pan *et al.* (2003) of 240 HCC patients (96% of which had an underlying HBV or HCV infection) found that positive OPN mRNA expression in tumor tissue correlated with a higher early tumor recurrence rate (recurrence within one year of resection) than that for OPN-negative tumor samples (*P* = 0.003). OPN expressing HCC patients (n = 133) also had significantly lower 10-year survival rates compared to OPN-negative HCC patients (n = 107, *P* = 0.00013) (Pan et al., 2003).

Numerous other studies have outlined the prognostic significance of OPN expression in HCC, and have, like Zhang *et al.* (2006), determined that higher levels of OPN are predictive of worse overall survival (OS) and disease-free survival (DFS). Using Kaplan-Meier regression analysis, Xie *et al.* (2007) found that HBV-induced HCC patients with positive OPN immunostaining had a mean DFS of 13.55 months and OS of 18.63 months, compared to 36 months and 42.59 months for HBV-induced HCC patients without OPN immunostaining (DFS: *P* = 0.0002, OS: *P* < 0.00001) (Xie, H. et al., 2007).

1.4 Prelude to this study

The basis for this study arose from previous investigations into differential gene expression in HCV-related liver disease, and HCC that develops in the background of a transgenic HCV mouse model (Lerat et al., 2002), performed by the principal supervisor and colleagues. Through their investigations, OPN was consistently identified as significantly upregulated in both HCV-infected livers with advanced liver damage and HCCs in both humans and mice. Using DNA microarray analysis and RT-PCR confirmation, OPN mRNA levels increased more than

5-fold in Huh-7 cells transfected with HCV core protein (following 72 hour core protein expression using a tetracycline-on/tetracycline-off system) (Li et al., 2002). In a study of 21 patients with chronic HCV infection (also using DNA microarray analysis), relative OPN transcript levels were found to be significantly higher in patients with advanced bridging fibrosis or established cirrhosis (stage 3 - 4, n = 6) than those with early portal or periportal fibrosis (stage 1 - 2, n = 10), no fibrosis (stage 0, n = 5), and uninfected individuals (n = 5) (Lau et al., 2005). OPN transcript was also expressed at significantly higher levels in HCC tumors of mice transgenic for the complete HCV polyprotein compared to surrounding non-neoplastic liver tissue (Beard et al., in preparation). These data suggest that OPN expression is more significant during the latter stages of HCV-related liver disease and may play a role in hepatocarcinogenesis. These preliminary observations and a paucity of information regarding the expression and role of OPN in liver disease and HCC prompted this study.

1.5 Aims of this project

The aims of this research project were four-fold:

- To investigate the expression of OPN and its alternatively spliced variants in cultured hepatocyte-derived carcinoma cells, healthy and HCV-infected liver tissues and HCC tumor tissue, and to quantitate either relatively or absolutely the level of each OPN variant in a given sample
- To investigate the sub-cellular localisation and effect on hepatoma cell proliferation of each OPN variant *in vitro*
- To investigate the effect of each OPN variant on subcutaneous tumor growth in a nude mouse xenograft model
- To investigate the potential of OPN as a biomarker of HCV-related liver disease

CHAPTER 2: Materials and Methods

2.1 General Reagents

2.1.1 Plasmid vectors

pRc/CMV (Invitrogen)

pRc/CMV is a 5.5 kilobase mammalian expression vector containing a multiple cloning cite (MCS), CMV promoter, and resistance cartridges for ampicillin and neomycin (Figure 2.1). This vector was used to facilitate over-expression of the three OPN variants.

pRc/CMV-OPN-A, pRc/CMV-OPN-B, pRc/CMV-OPN-C

The three OPN variants were individually cloned into pRc/CMV. Each variant was amplified from normal human liver cDNA using primers specifically designed to amplify all three variants in their entirety (see Table 2.1). Sequencing was performed to verify that the correct sequences had been obtained. Each variant was ligated into pRc/CMV by digesting the products and destination vector with restriction enzymes *Xba*I and *Hind*III. These vectors were used in transient transfection experiments to investigate the sub-cellular localisation of each variant, and to generate stable transfectant cell lines for growth and tumor studies.

pRc/CMV-e-GFP

A kind gift from Dr. Tom McNaughton (Queensland University of Technology, Brisbane, Australia), this plasmid encodes GFP (green fluorescent protein) from a CMV promoter. This vector was used to determine the transfection efficiency of cultured cell lines.

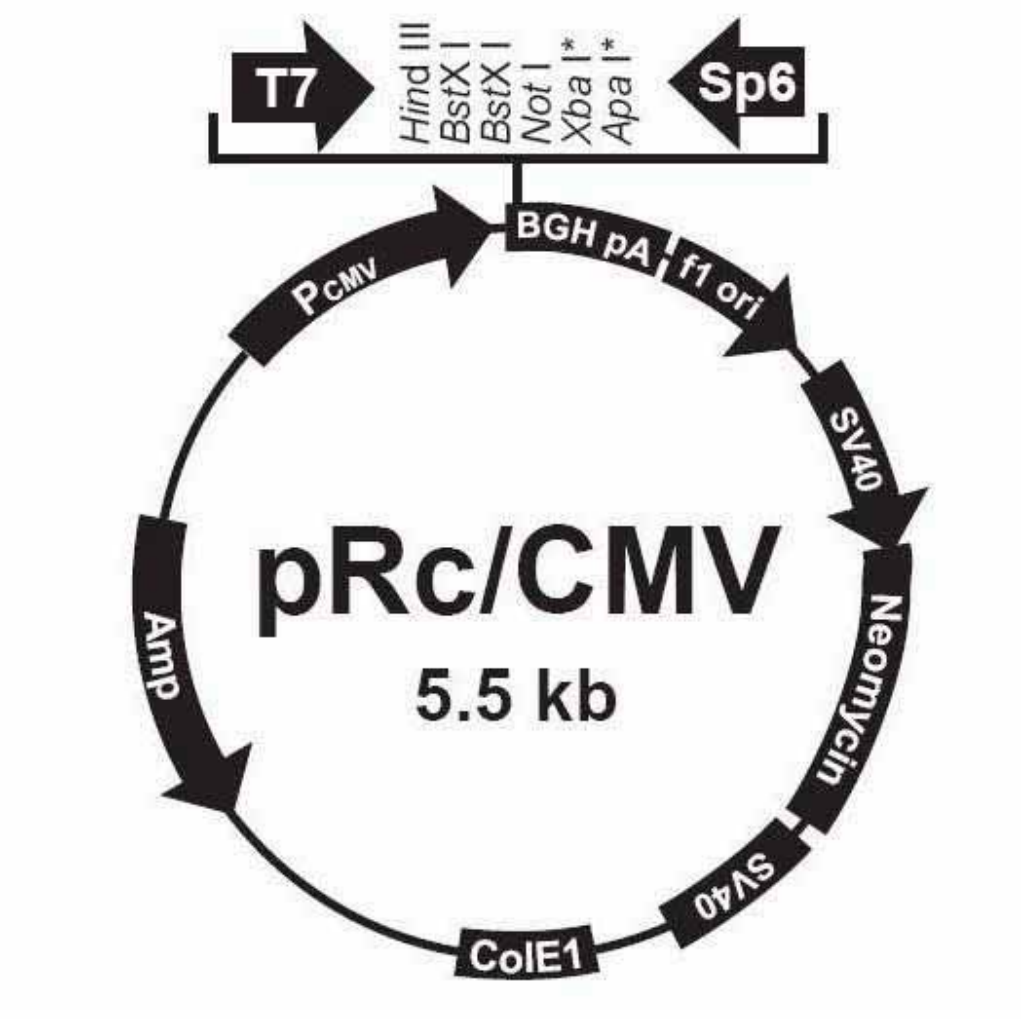


Figure 2.1 Vector map of pRc/CMV (Invitrogen). This vector contains resistance cartridges for ampicillin and neomycin and encodes a CMV promoter. The three osteopontin variants were cloned into the multiple cloning site (MCS) of pRc/CMV by digestion with *Hind*III and *Xba*I.

Primer Name	Sequence 5’ → 3’	Application
OPN-F2	GTT GAA GCT TCT CAC TAC CAT GAG AAT TGC AGT G	Clon., PCR
OPN-R3	TAG TTC TAG ACC TTT TAA TTG ACC TCA GAA GAT G	Clon., PCR
OPN-A/B/C-F	ATG AGA ATT GCA GTG ATT TGC TTT TGC CT	PCR
OPN-A/B/C-R	CAT GGT CAT CAT CAT CTT CAT CAT C	PCR
CD44s-F	GAC ACA TAT TGC TTC AAT GCT TCA GC	PCR
CD44s-R	GAT GCC AAG ATG ATC AGC CAT TCT GGA AT	PCR
hmrGAPDH-F	ACC ACA GTC CAT GCC ATC AC	PCR
hmrGAPDH-R	TCC ACC ACC CTG TTG CTG TA	PCR
albumin-F	CCG ATT GGT GAG ACC AGA GGT TG	PCR
albumin-R	GAA GCC TTC CCT TCA TCC CGA	PCR
OPN-F1	AAT GGT GCA TAC AAG GCC ATC	PCR, qRT-PCR
OPN-R1	TGT CCT TCC CAC GGC TGT	PCR, qRT-PCR
OPN-A/qPCR-F	ATC TCC TAG CCC CAC AGA AT	PCR, qRT-PCR
OPN-A/qPCR-R	CAT CAG ACT GGT GAG AAT CAT C	PCR, qRT-PCR
OPN-B/qPCR-F	CAC AGA CCC TTC CAA GTA AGT C	PCR, qRT-PCR
OPN-B/qPCR-R	CAG ATT CAT CAG AAT GGT GAG	PCR, qRT-PCR
OPN-C/qPCR-F	CTG AGG AAA AGC AGA ATG	PCR, qRT-PCR
OPN-C/qPCR-R	AAT GGA GTC CTG GCT GT	PCR, qRT-PCR
OPN-A-F	TCC TAG CCC CAC AGA ATG	PCR, qRT-PCR
OPN-A-F2	AGC CCC ACA GAA TGC TGT	PCR, qRT-PCR
OPN-A-F3	CAC AGA ATG CTG TGT CCT G	PCR, qRT-PCR
OPN-AC-R	TAC TTG GAA GGG TCT CTT GTT	PCR, qRT-PCR
OPN-B-F	GAG AAT TGC AGT GAT TTG CTT	PCR, qRT-PCR
OPN-B-R	ACT TAC TTG GAA GGG TCT G	PCR, qRT-PCR
OPN-C-F	GTT CTG AGG AAA AGC AGA ATG	PCR, qRT-PCR
36b4-F	AGA TGC AGC AGA TCC GCA T	qRT-PCR
36b4-R	GGA TGG CCT TGC GCA	qRT-PCR
OPN-F-Fam	FAM-CAG GCT GAT TCT GGA AGT TCT GAG G	GS
OPN-R(Fam)	CTT TCG TTG GAC TTA CTT GGA AGG	GS

Table 2.1 All primers used in this thesis. Underlined sequences indicate introduced restriction enzyme sites. Clon. = cloning, PCR = polymerase chain reaction, qRT-PCR = real-time quantitative PCR, GS = GeneScan® Genotyping PCR.

2.1.2 Synthetic oligonucleotides

All oligonucleotides listed in Table 2.1 were obtained from GeneWorks at PCR/sequencing purity. Primers were received in lyphophilised form, diluted in dH_2O to 20 µM and stored at -20 °C until use. Oligonucleotide concentration was determined using the following formula, assuming an average MW of 330 Daltons per nucleotide:

$$\text{Oligonucleotide concentration } (\mu M) = \frac{\text{concentration (mg/ml)} \times 10^6}{\text{nucleotide length} \times \text{nucleotide MW}}$$

Primer design is discussed in Section 2.3.6.

2.2 Tissue Culture Techniques

2.2.1 Tissue culture medium

All cultured cell lines were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L D-Glucose, 25 mM HEPES and 2 mM L-glutamine (Gibco BRL, Invitrogen). Media was supplemented with 10% (w/v) foetal calf serum (FCS; Trace Biosciences), 12 µg/ml penicillin (CSL), 16 µg/ml gentamycin (Pharmacia) and other supplements as required (Table 2.2).

2.2.2 Maintenance of cell lines

Cells were maintained in sterile 0.2 µm vented tissue culture flasks (25 cm^2, 75 cm^2, or 175 cm^2), dishes (3.5 cm^2, 6 cm^2 or 10 cm^2) or trays (6-, 12-, 24-, 48- or 96-well) (Falcon®, Becton Dickinson Labware) at 37 °C in a humidified incubator (5% CO_2). Adherent cell populations were subcultured by aspirating the culture medium, washing with PBS (see Appendix 1), and incubating in a small volume of Trypsin-EDTA (see Appendix 1). Cells were gently dislodged from the culture vessel and resuspended in warm culture medium. Cell viability was determined

Cell Line	Culture Media
Huh-7	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
HepG2	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
Hep3B	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
PLC/PRF/5	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
HeLa	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
MCF-7	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
MDA-MB-231	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
Hepa1-6	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
H-4-II-E	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
Cos-7	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
Vero	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin
Huh-7/mock	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin, 500 μg/ml G418 sulphate
OPN-A, -B, -C stable transfectant Huh-7	DMEM, 10% FCS, 12 μg/ml penicillin, 16 μg/ml gentamycin, 500 μg/ml G418 sulphate

Table 2.2 Cell culture requirements for all cell lines used in this study.

by trypan blue exclusion and cell numbers enumerated using a haemocytometer. Cells were subcultured every 2 - 5 days at a ratio of 1:4 to 1:10 depending on cell type and confluency.

2.2.3 Transient transfection of plasmid DNA

FuGENE 6 Transfection Reagent (Roche) was used to transfect plasmid DNA into various cultured cell lines as per the manufacturer's instructions. Cells were seeded in tissue culture dishes (6 cm^2) or trays (6-, 12-, or 24-well) 24 - 48 hours prior to transfection such that they were 50 - 70% confluent at time of transfection. FuGENE 6 Transfection Reagent and plasmid DNA were diluted in an appropriate volume of serum free DMEM according to the manufacturer's recommendations, generally with a DNA:FuGENE 6 ratio of 1:3. Following a 15 minute incubation period, the transfection mixture was added in a dropwise fashion to cell monolayers which were returned to normal culture conditions (37 °C, 5% CO_2). Assays were performed 24 - 72 hours post transfection.

2.2.4 Stable transfection of plasmid DNA to generate over-expressing cell lines

To generate stably transfected cell lines expressing the three OPN variants, Huh-7 cells were seeded in 10 cm^2 tissue culture dishes at a density of 3×10^5 cells per dish and transfected with vector clones of each variant using FuGENE 6 Transfection Reagent 24 hours later. 24 hours post transfection, 500 μg/ml G418 sulphate (Invitrogen) was added to media to select for cellular integration of plasmid DNA. 500 μg/ml G418 sulphate was chosen as the optimal concentration for selection based on a titration experiment performed on naïve Huh-7 cells (Figure 2.2). Once antibiotic resistance was observed, stably transfected colonies were generated by re-plating cells at a low density in fresh 10 cm^2 tissue culture dishes (diluted 1/10 - 1/30). Individual colonies were isolated using sterile glass cloning rings and expanded until cell number was sufficient for screening of clones for expression of the protein of interest. Successfully transfected lines were

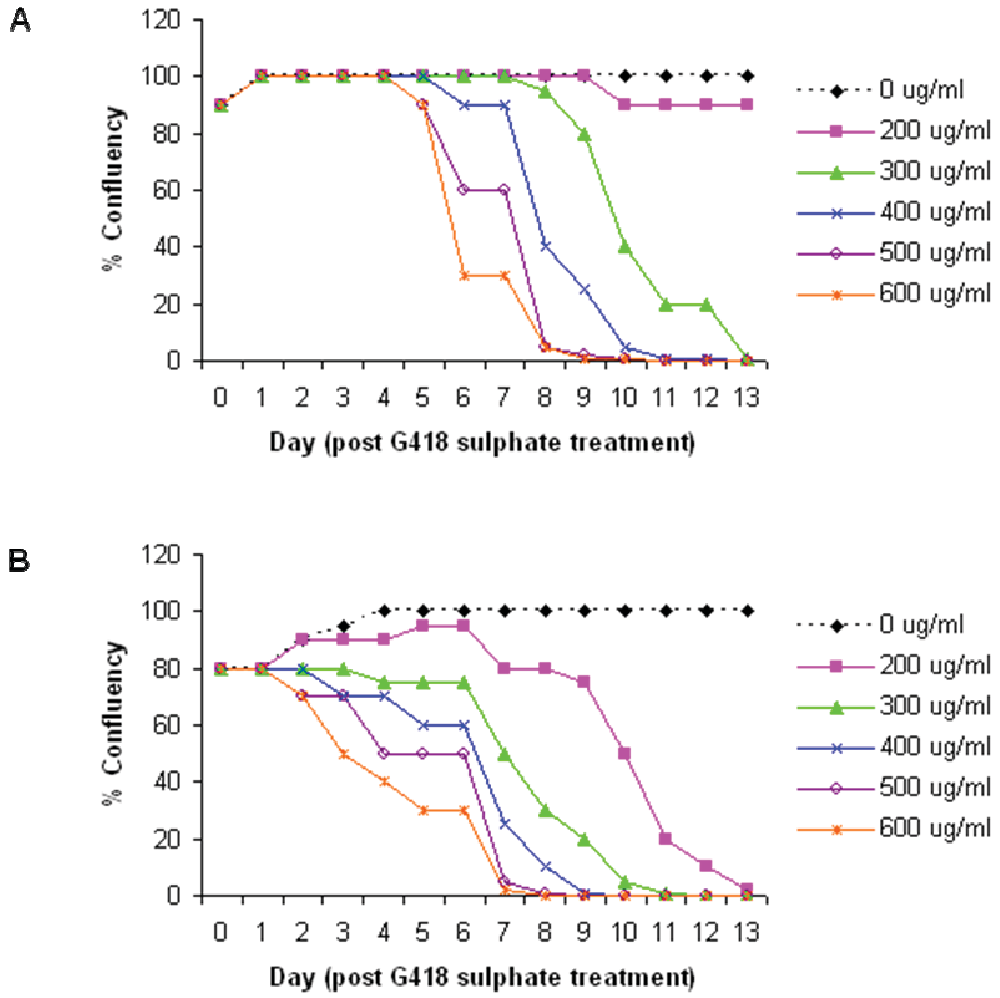


Figure 2.2 Optimal G418 sulfate concentration for the selection of OPN-transfected Huh-7 cells was calculated by seeding naïve Huh-7 cells in 6-well tissue culture plates and allowing cells to grow in the presence of increasing concentrations of G418 sulfate. From a starting cell confluency of 80-90%, repeat experiments **(A)** and **(B)** indicated that 500 µg/ml G418 sulfate was sufficient to reduce cell confluency by 95% after 7-8 days of growth.

maintained under normal culture conditions (37 °C, 5% CO_2) and continued G418 sulphate selection.

2.2.5 Transient transfection of Stealth™ siRNA oligonucleotides

Stealth™ siRNA double stranded RNA oligonucleotides designed to knock down endogenous expression of the OPN receptor CD44 were transfected into Huh-7 and HeLa cells using the reverse transfection method of Lipofectamine™ 2000 (Invitrogen) as per the manufacturer's instructions. Briefly, 100 pmol of siRNA was mixed with 200 µl serum-free DMEM in a 12-well tissue culture tray. 2 µl Lipofectamine™ 2000 was added and complexes incubated for 10 - 20 minutes at room temperature. Cultured cells were diluted in warm media without antibiotics and added to wells such that each well contained 7×10^4 cells in a total volume of 1.2 ml media (including siRNA complex at a final concentration of 40 nM). Cells were assayed for CD44 expression (and growth rate when required) 24 - 96 hours post transfection. Conditions were scaled up or down to accommodate 6- or 24-well tissue culture trays as outlined in the provided Lipofectamine™ 2000 protocol. The siRNA transfection efficiency of each cell line was assessed by performing the above transfection with 20 - 100 pmol BLOCK-iT™ Fluorescent Oligo (Invitrogen) and examination of transfection efficiency by fluorescence microscopy. Three different siRNA oligonucleotide complexes were tested for their ability to minimise CD44 expression in a transient system; two were chosen for further experiments.

2.2.6 Preparation of conditioned media

Conditioned media from Huh-7 cell lines stably transfected with each OPN variant was produced by replacing growth media on 80% confluent cell monolayers with fresh supplemented DMEM (without G418 sulphate) for 48 hours. Media was collected, supplemented with 2 mM L-glutamine (Gibco BRL, Invitrogen) and stored at -20 °C until required.

2.2.7 Cellular proliferation

To monitor proliferation of cultured cells under various conditions, cells were seeded in triplicate at a density of 7×10^4 cells per well of a 12-well tissue culture tray. Cells were seeded in fresh media or conditioned media from stably transfected cell lines, with or without the addition of specific antibodies or Stealth^{TM} siRNA oligonucleotides. Media (including antibody or siRNA) was replenished after two days of growth. Daily, three wells per cell line were trypsinised and counted using trypan blue exclusion to monitor cellular growth over a period of four days.

2.2.8 Trypan blue exclusion

Estimation of cell number for seeding of transient assays and counting during cellular proliferation experiments was performed by trypan blue exclusion. Trypsinised cells were mixed with an equal volume of trypan blue (see Appendix 1) and counted using grids on a haemocytometer. Cell concentration was estimated using the following equation

$$\text{cell concentration (cells/ml)} = \text{cell count per grid} \times \text{dilution factor} \times 10^4$$

2.2.9 Cultured cell lines used in this study

Huh-7

Huh-7 is a human hepatocellular carcinoma cell line of epithelial origin isolated from a 57 year old Japanese male (Nakabayashi et al., 1982). This was the major cell line used for cell culture experiments, and served as the parental cell line for production of stably transfected OPN lines.

HepG2

HepG2 is a differentiated human hepatoblastoma cell line of epithelial morphology isolated from a 15 year old Caucasian male (Aden et al., 1979). These cells were used as an alternative hepatoma line to Huh-7.

Hep3B

Hep3B is a human hepatocellular carcinoma cell line of epithelial origin originally isolated from an 8 year old African American male (Knowles et al., 1980). This cell line contains an integrated hepatitis B virus genome, expresses the HBV surface antigen, and was used as a negative control for CD44 expression. These cells were kindly donated by Prof. Jacob George (Westmead Millennium Institute, Sydney, Australia).

PLC/PRF/5

PLC/PRF/5 is a human hepatoma cell line originally isolated from a 24 year old male from Mozambique that expresses the hepatitis B virus surface antigen (HBsAg) (Alexander et al., 1976; MacNab et al., 1976). These cells were used as an example of a human hepatoma cell line.

HeLa

HeLa is a human cervical adenocarcinoma cell line isolated from Henrietta Lacks, a 31 year old African American woman, in 1951 (Gey et al., 1952). These cells were used in this study as an example of a non-liver carcinoma cell line and as a positive control for high-level expression of CD44, one of the main receptors for OPN.

MCF-7

MCF-7 is a breast adenocarcinoma cell line of epithelial origin isolated from a 69 year old Caucasian female (Soule et al., 1973). This cell line was used as an example of a non-liver carcinoma cell line.

MDA-MB-231

MDA-MB-231 is a breast adenocarcinoma cell line of epithelial origin isolated from a 51 year old Caucasian female (Cailleau et al., 1974). This cell line was kindly provided by Ms Jane

Holland (The University of Adelaide, Adelaide, Australia) and was used as an example of a non-liver carcinoma cell line.

Hepa1-6

Hepa1-6 is a mouse hepatoma cell line of epithelial origin derived from a C57/L mouse (Darlington 1987).

H-4-II-E

H-4-II-E is a rat hepatoma cell line of epithelial origin (Reuber 1961).

Cos-7

Cos-7 is a monkey fibroblast-like kidney cell line, isolated from a normal African green monkey and permissive for SV40 replication (Gluzman 1981).

Vero

Vero is a monkey kidney epithelial cell line, isolated from a normal African green monkey (Yasumura and Kawakita 1963).

2.3 Molecular Biology Techniques

2.3.1 Transformation of competent bacteria

Frozen aliquots of the competent *Escherichia coli* strain DH5α (prepared in-house or purchased from commercial sources) were thawed on ice for 5 minutes prior to the addition of an appropriate concentration of plasmid DNA. Tubes were gently mixed and incubated on ice for 20 minutes. Cells were heat shocked in a 42 °C water bath for 90 seconds followed by incubation on ice for 2 minutes. 400 µl Luria Bertani Broth (LB) was added and cells incubated in a 37 °C water bath for 30 minutes to allow for the induction of antibiotic resistance gene expression.

Cells were plated onto L-Agar containing 100 μg/ml ampicillin and incubated at 37 °C for approximately 16 hours. Colonies were then selected and prepared for plasmid extraction.

2.3.2 Mini-preparation (small scale) of plasmid DNA

5 ml LB cultures containing 100 μg/ml ampicillin were inoculated with a single transformed bacterial colony and incubated overnight with vigorous shaking at 37 °C. Plasmid DNA was isolated from log phase cultures using the Ultraclean™ 6 Minute Mini Plasmid Prep Kit (Mo Bo Laboratories) according to the manufacturer's instructions. Briefly, 2 ml overnight culture was centrifuged at 10,000 × *g* for 1 minute, all traces of supernatant removed and the pellet resuspended in 50 μl of Solution 1 by vortexing. 100 μl Solution 2 was added, tubes inverted, 325 μl Solution 3 added and tubes again inverted. Tubes were centrifuged for 1 minute at 10,000 × *g* and the resultant supernatant transferred to a provided spin filter. Spin filters were centrifuged for 30 seconds at 10,000 × *g*, waste supernatant discarded and the filter centrifuged again after the addition of 400 μl Solution 4. Spin filters were then placed in a fresh collection tube and plasmid DNA eluted with 50 μl Solution 5 by centrifugation at 10,000 × *g* for 30 seconds. Plasmid DNA was stored at -20 °C until further use.

2.3.3 Maxi-preparation (large scale) of plasmid DNA

3 ml LB cultures containing 100 μg/ml ampicillin were inoculated with a single transformed bacterial colony and incubated for 8 hours at 37 °C with shaking. 100 μl of log phase culture was transferred to 100 ml LB containing 100 μg/ml ampicillin and incubated overnight with shaking at 37 °C. Plasmid DNA was isolated using the Plasmid Maxi Kit (Qiagen) as per the manufacturer's instructions. Overnight cultures were centrifuged at 6,000 × *g* for 15 minutes at 4 °C using a JA10 or JA14 rotor in a Beckman J2-21M Induction Drive centrifuge. Cell supernatant was discarded and the cell pellet resuspended in 10 ml Buffer P1. 10 ml Buffer P2

was added and the mixture incubated at room temperature for 5 minutes, after which 10 ml chilled Buffer P3 was added, tubes inverted and incubated for 20 minutes on ice. Tubes were then centrifuged for 30 minutes at 20,000 × *g* at 4 °C in a JA20 rotor in a Beckman J2-21M Induction Drive centrifuge. Meanwhile, supplied QIAGEN-tip 500 columns were equilibrated with 10 ml Buffer QBT. Centrifuged supernatant was passed through the QIAGEN-500 column, which was then washed with two 30 ml washes of Buffer QC. Plasmid DNA was eluted from the column into a fresh collection tube with 15 ml of Buffer QF. DNA was precipitated by adding 10.5 ml isopropanol to the eluted plasmid mixture, gently mixing and centrifuging at 15,000 × *g* at 4 °C for 30 minutes. Supernatant was discarded and the resultant pellets washed with 1 ml 70% ethanol (v/v) by centrifuging at 15,000 × *g* for 10 minutes at room temperature. After removal of the ethanol wash, pellets were air dried for 5 - 10 minutes, re-dissolved in 100 µl sterile dH_2O and stored at -20 °C until further use.

2.3.4 Extraction of total RNA and cDNA synthesis

Total cellular RNA was extracted from adherent cell monolayers and tissue samples using Trizol® Reagent (Invitrogen) according to the manufacturer's recommendations. Cell monolayers were lysed directly in Trizol® Reagent (1 ml per 10 cm^2 culture dish), whilst tissue samples were homogenised in Trizol® Reagent (1 ml per 50 - 100 mg tissue). All samples were incubated at room temperature for 5 minutes, 200 µl chloroform per 1 ml Trizol^R added, tubes shaken vigorously for 15 seconds and incubated again at room temperature for 2 minutes. Samples were then centrifuged at 12,000 × *g* for 15 minutes at 4 °C. Following separation, the upper aqueous layer was transferred to an RNAse-Free 1.5 ml Microfuge tube (Ambion) and RNA was precipitated with isopropanol (0.5 ml per 1 ml Trizol®) for 10 minutes at room temperature. Samples were again centrifuged for 15 minutes at 12,000 × *g* at 4 °C. Supernatants were aspirated and pellets resuspended in 1 ml 70% (v/v) ethanol and centrifuged at 7,500 × *g* for 10 minutes at

4 °C. After removal of the ethanol wash, the resultant RNA pellet was air dried and resuspended in 20 µl RNAse-free dH_2O. All RNA samples were treated with *DNase*I to remove any contaminating DNA. When RNA extraction was not immediately possible, tissue samples were stored in RNAlater® RNA Stabilisation Solution (Ambion) at 4 °C for extraction at a later time.

cDNA was prepared from RNA harvested from cultured cell lines and tissue samples using M-MLV reverse transcriptase (Promega). 1 µg of total RNA was combined with 1 µg of random hexamer primer (GeneWorks) and diluted to a total volume of 14 µl with dH_2O. Tubes were incubated at 70 °C for 5 minutes, then on ice for a further 5 minutes. 1 × M-MLV RT reaction buffer (Promega), 0.5 µmol dNTP mix (0.5 µmol each dATP, dCTP, dGTP, dTTP; Promega), 40 units rRNAsin® RNase Inhibitor (Promega), 200 units M-MLV RT, RNase H(-) Point Mutant (Promega) and 3.25 µl dH_2O were added to each tube, samples mixed gently and incubated at room temperature for 10 minutes, then 42 °C for 50 minutes. Samples were then diluted to a final volume of 100 µl with dH_2O and stored at -20 °C until further use.

2.3.5 RNA quantitation

The RNA concentration of a given sample was determined using a SmartSpec™ 3000 UV spectrophotometer (BioRad Laboratories). Samples were diluted in dH_2O and the absorbance measured at 260 nm (A_{260nm}) using a glass cuvette and a 10 mm light path. The final concentration was calculated using the following equation:

$$\text{RNA concentration (ng/µl)} = A_{260nm} \text{ x dilution factor x } 40$$

RNA purity was determined using the A_{260nm} (nucleic acid):A_{280nm} (protein) ratio, with a ratio of 1.80 being the accepted minimum.

2.3.6 Primer design

Forward and reverse primers were designed for a number of genes of interest and varying applications. Table 2.1 outlines all primers used in this study, and the application for which they were intended. In general, primers were designed to be at least 15 base pairs in length and have a greater than 50% GC content to ensure a high melting temperature, thus reducing the level of non-specific binding. For sequences where it was not possible to have greater than 50% GC content, the primer length was increased. Lengthy G or C sequences were avoided where possible, and primer pairs were compared to prevent binding and subsequent extension between pairs, therefore avoiding primer dimers and potentially increasing the efficiency of the PCR reaction. Introduced restriction enzymes sequences (to facilitate cloning into vectors) are underlined.

2.3.7 Polymerase Chain Reaction (PCR)

All PCR reactions were performed using a MyCycler^{TM} Thermal Cycler (BioRad Laboratories) and the high fidelity polymerase AmpliTaq Gold (Applied Biosystems). Reactions were carried out in 50 µl volumes containing 20 pmol each of forward and reverse primer (Table 2.1), 1 × PCR Reaction Buffer (Applied Biosystems), 0.2 µmol dNTP mix (Promega), 25 mM $MgCl_2$ (Applied Biosystems), 1 unit AmpliTaq Gold polymerase (Applied Biosystems) and the appropriate amount of target plasmid DNA (typically 10 - 100 ng) or cDNA (5 µl of diluted synthesised cDNA). PCR reaction conditions comprised initialisation at 94 °C for 10 minutes, followed by 25 - 40 cycles of denaturation at 94 °C for 30 seconds, annealing at 55 °C for 30 seconds and extension at 72 °C for 30 seconds. A final elongation step was then performed at 72 °C for 7 minutes. PCR reaction outcomes were confirmed by agarose gel electrophoresis. For reverse transcriptase PCR (RT-PCR), all cDNA samples were also used as a template for parallel

reactions using primers for the housekeeping gene GAPDH (Table 2.1) to confirm equivalent cDNA input and integrity.

2.3.8 Agarose gel electrophoresis

Gel electrophoresis was performed using 1 - 2% (w/v) agarose gels, made by dissolving DNA grade agarose (AppliChem) in 1 × TAE (see Appendix 1). Gels were cast in a BioRad Wide Mini-Sub® Cell GT tank. Samples were mixed with 6 × loading dye (see Appendix 1), loaded into wells and subjected to electrophoresis at 50 - 100 volts in 1 × TAE. 0.5 μg of 1 kilobase or 100 base pair DNA molecular weight marker (New England Biolabs) was electrophoresed simultaneously to estimate product size. Following electrophoresis, gels were stained in 3 × GelRed™ Nucleic Acid Gel Stain (Biotium Inc.) for approximately 15 minutes and bands visualised under ultraviolet light on a BioRad Universal Hood II gel documentation system using Quantity One® Version 4.6 Basic software (BioRad Laboratories).

2.3.9 Real-time quantitative PCR

Real-time quantitative PCR was performed to determine relative levels of OPN and its splice variants in a range of samples using the comparative C_T method of SYBR® Green PCR Master Mix (Applied Biosystems). 5 μl cDNA or 0.001 - 1 pg vector DNA was combined with 10 μl SYBR® Green PCR Master Mix and 300 nM each of forward and reverse primers (Table 2.1) in duplicate. A standard curve of 10-fold dilutions of vectors expressing each of the OPN variants was utilised to enable calculation of the relative concentrations of each variant in a given cDNA sample. All cDNA samples were also combined with 10 μl SYBR® Green PCR Master Mix and 300 nM each of forward and reverse primers for the housekeeping gene ribosomal protein large P0 (RPLPO, aka 36b4; Table 2.1) to normalise input cDNA levels. Reaction conditions were controlled by an ABI PRISM 7000 Sequence Detection System (Applied Biosystems) and

comprised denaturation at 95 °C for 10 minutes followed by 40 cycles of 95 °C for 15 seconds and 60 °C for 1 minute. A final dissociation step of 60 °C for 10 minutes followed to facilitate melt curve analysis. Data was analysed using ABI Prism 7000 SDS Software; the threshold was set at 0.2 for all experiments.

2.3.10 GeneScan® genotyping PCR

GeneScan® Genotyping PCR was utilised in an attempt to quantify each of the three OPN variants individually within a given sample. Primers were designed to simultaneously amplify all three variants (Table 2.1), with the forward primer labeled with a 5' 6-FAM tag to allow for quantification of fluorescence for each of the three differently sized sequences. Reactions were set up as per normal PCR conditions, and run through the same cycling run for 15 - 30 cycles. PCR products were diluted 1/3 and the fluorescence of the three variants quantitated on a 3730 DNA Analyser (Applied Biosystems) at the Molecular Pathology Sequencing facility of the IMVS. Data was analysed using ABI Prism GeneScan® 3.1.

2.3.11 Gel purification

To purify electrophoresed DNA from agarose gels for downstream applications, bands were excised using a scalpel blade and DNA purified using the MinElute Gel Extraction Kit (Qiagen) using a microcentrifuge as per the manufacturer's recommendations. Briefly, excised DNA bands were weighed in a clean tube and 300 µl Buffer QG added per 100 mg gel weight. Samples were incubated at 50 °C for approximately 10 minutes to dissolve the gel slice, after which 100 µl isopropanol was added per 100 mg gel and tubes inverted to mix. Samples were applied to a MinElute column (placed in a supplied 2 ml collection tube) and centrifuged at 14,000 rpm for 1 minute. Flow-through was discarded, 500 µl Buffer QG added and tubes centrifuged again at 14,000 rpm for 1 minute. Flow-through was again discarded and the column washed with 750 µl Buffer PE by centrifuging for 1 minute at 14,000 rpm. Final flow-through was discarded, the

MinElute column placed into a clean 1.5 ml collection tube and DNA eluted by adding 10 µl dH_2O to the column filter, incubating for 1 minute at room temperature, and centrifuging at 14,000 rpm for 1 minute. Samples were stored at -20 °C until required.

2.3.12 DNA sequencing

Gel purified PCR products and plasmid DNA were sequenced to ascertain if the desired sequence had been obtained. Sequencing was performed in a 10 µl volume containing 150 ng plasmid DNA or 10 ng amplicon, 0.32 pmol primer, 1 × BigDye® Buffer (Applied Biosystems), 0.5 µl BigDye® Terminator v3.1 (Applied Biosystems) and dH_2O. Sequencing PCR conditions comprised an initial denaturation step of 94 °C for 5 minutes, followed by 25 cycles of 94 °C for 30 seconds, 50 °C for 15 seconds and 60 °C for 4 minutes. A final step of 60 °C for 7 minutes was performed, after which reactions were prepared for sequencing by vortexing with 80 µl of 75% (w/v) isopropanol and incubating for 15 minutes at room temperature. Samples were centrifuged for 20 minutes at 14,000 rpm, supernatant aspirated and pellets resuspended in a further 250 µl of 75% (w/v) isopropanol. Tubes were again vortexed and centrifuged for 5 minutes at 14,000 rpm, supernatant removed and pellets dried at room temperature for 15 minutes. Samples were stored at -20 °C until sequencing could be performed on a 3730 DNA Analyser at the Molecular Pathology Sequencing facility of the IMVS. Data was analysed using ChromasPro version 1.34 and compared to known sequences using NCBI BLAST nucleotide blast search against the human genome (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi).

2.3.13 Restriction endonuclease digestion

Restriction enzyme digests were performed in 10 µl volumes containing 1 µg DNA and 10 units of restriction enzyme(s) (New England Biolabs; NEB) in an appropriate NEB buffer. Digests were incubated at 37 °C for 1 - 3 hours and visualised by agarose gel electrophoresis.

2.3.14 DNA ligation

PCR amplicons of the three OPN variants were digested with *Hind*III and *Xba*I to allow for directional cloning into similarly digested pRc/CMV. Ligation reactions were performed in 10 µl volumes containing 0.4 units T4 DNA Ligase (New England Biolabs), 1 × Ligase Buffer (New England Biolabs), and a 3:1 ratio of amplified OPN variant insert to destination vector. Reactions were allowed to proceed at room temperature for 1 - 2 hours. Ligation mixtures were transformed into competent DH5α cells for colony growth and plasmid purification.

2.3.15 Immunofluorescence staining and microscopy

To visualise chosen antigens in cultured cell lines, cells were seeded onto 13 mm^2 round cover glasses (Menzel-Glaser®) in 24-well tissue culture trays at a density of 5×10^4 cells per well. Following culture or transfection, media was removed, cells washed in PBS and fixed in 1:1 acetone:methanol on ice for 15 minutes. The fixative was removed, cells again washed in PBS and incubated in the appropriate primary antibody diluted in PBS (Table 2.3) for either 1 hour at 37 °C or overnight at 4 °C depending on the antibody requirements. Cells were washed three times in PBS and incubated for a further 1 hour at 37 °C in the appropriate concentration of fluorescently-labelled secondary antibody (Table 2.3). Following another round of three washes with PBS, cells were incubated in 3 mg/ml DAPI nuclear counterstain (Sigma) for 1 minute in the dark and again washed with PBS. Coverslips were then incubated in SlowFade® Anti-Fade Equilibration Buffer (Invitrogen) for 10 minutes and mounted onto glass microscope slides with SlowFade® Anti-Fade Mounting Media (Invitrogen).

Cultures were visualised using a Nikon TE 300 inverted microscope equipped with a CoolsnapFX B&W CCD Cooled Camera or an Olympus IX70 microscope equipped with a Peltier-Cooled Monochrome Interline Transfer CCD Camera, Model 4920 (Canon). Confocal

Name	Species	Application	Concentration/ Dilution	Source
Primary Antibodies				
α-OPN (K-20)	Goat	IF WB IHC Cell growth	1 μg/ml 400 ng/ml 2 μg/ml 200 ng/ml	Santa Cruz Biotechnology
α-OPN	Rabbit	IF WB IHC	1 μg/ml 1 μg/ml 2 μg/ml	NeoMarkers (LabVision)
α-OPN (MAB14331)	Mouse	ELISA	3 μg/ml	R&D Systems
α-HCAM (300)		Cell growth	200 ng/ml	Santa Cruz Biotechnology
α-CD44	Mouse	IF WB	4 μg/ml 200 ng/ml	NeoMarkers (LabVision)
α-OPN	Rat	WB	1/2,000	Chemicon
Unrelated IgG	Rabbit	Cell growth	200 ng/ml	Santa Cruz Biotechnology
Normal rabbit IgG	Rabbit	IHC	2 μg/ml	Cell Signaling
α-β-actin	Mouse	WB	1/10,000	Sigma
Secondary Antibodies				
Biotinylated α-OPN	Goat	ELISA	200 ng/ml	R&D Systems
α-goat HRP	Donkey	WB	33 ng/ml	Rockland
α-mouse HRP	Donkey	WB	50 ng/ml	Rockland
α-rabbit HRP	Donkey	WB	500 ng/ml	Rockland
α-rat HRP	Donkey	WB	100 ng/ml	Rockland
α-goat Alexa-488	Goat	IF	10 mg/ml	Molecular Probes
α-mouse Alexa-488	Goat	IF	10 mg/ml	Molecular Probes
α-rabbit Alexa-488	Donkey	IF	10 mg/ml	Molecular Probes

Table 2.3 Antibodies used in this study. Concentrations given indicate the final concentration at which the antibody was used for its intended purpose. For antibodies which were provided as neat serum or ascite fluid, the final dilution used is provided instead. IF = immunofluorescent staining, WB = western blotting, IHC = immunohistochemical staining, Cell growth = cellular proliferation assay, ELISA = in-house sandwich ELISA.

microscopy was performed using a BioRad Radiance 2100 confocal microscope at the IMVS Detmold Imaging Centre, using 20 × UAPOW and 40 × UPLAPO water immersion optical lenses. Images were automatically merged and processed using Confocal Assistant software.

2.3.16 Extraction of cellular protein

Total protein was extracted from cultured cells and tissue samples using RIPA Buffer (see Appendix 1). Cultured cell monolayers were either trypsinised, centrifuged and resuspended in ice-cold RIPA buffer, or lysed directly in ice-cold RIPA buffer (100 µl per well of a 6-well tissue culture plate, scaled up or down appropriately). Tissue samples were homogenised on ice and resuspended in RIPA buffer (100 µl per 100 mg tissue). 1 µl proteinase inhibitor cocktail (Sigma) was added per 100 µl RIPA buffer and samples incubated on ice for 20 minutes. Samples were then centrifuged at 14,000 rpm for 15 minutes at 4 °C, and protein-containing supernatants transferred to fresh tubes, vortexed, and stored at -20 °C until further use.

2.3.17 Protein quantification

The protein concentration of a given sample was estimated by BioRad Protein Assay (BioRad Laboratories), using bovine plasma γ-globulin as a standard. All samples were diluted 1/10 in dH_2O in duplicate and 10 µl of diluted protein combined with 200 µl provided Dye Reagent Concentrate (diluted 1/5 in dH_2O). A_{600} values were determined using a MR5000 plate reader (Dynatech) and compared to a standard curve of known protein concentrations of bovine plasma γ-globulin to determine sample protein concentration.

2.3.18 SDS-PAGE and protein transfer

SDS-PAGE was performed as previously described (Laemmli 1970). A 12% separating gel (see Appendix 1) was cast in a BioRad Mini-PROTEAN® Tetra Cell gel tank and layered with dH_2O to prevent oxidation and enhance gel polymerisation. Once set, a 5% stacking gel (see Appendix

1) was layered onto the separating gel and a comb inserted to form wells. Protein samples were prepared for electrophoresis by boiling for 5 minutes with 1 × SDS-PAGE sample buffer (see Appendix 1). 50 μg samples were then loaded onto the gel alongside 7.5 μg Precision Plus Protein® Standards - Kaleidoscope (BioRad Laboratories) and run at 100 volts for 1 - 2 hours in SDS-PAGE running buffer (see Appendix 1). Gels were equilibrated in cold western transfer buffer (see Appendix 1) for 5 minutes and proteins transferred to a Hybond ECL membrane (Amersham Pharmacia Biotech) in western transfer buffer for 1 - 2 hours at 100 volts in a BioRad Mini Trans Blot Electrophoretic Transfer Cell.

2.3.19 Western blotting

Following transfer, membranes were blocked with 5% skim milk (Diploma) in 0.1% PBS-T (see Appendix 1) for 1 hour with gentle agitation. Membranes were rinsed twice in 0.1% PBS-T and incubated in the appropriate concentration of primary antibody diluted in 0.1% PBS-T (Table 2.3) overnight at 4 °C. Membranes were rinsed twice in 0.1% PBS-T and incubated once for 15 minutes and three times for 5 minutes each in fresh changes of 0.1% PBS-T to remove any unbound primary antibody. Membranes were then incubated in the appropriate horseradish peroxidase-conjugated secondary antibody (Table 2.3) for 1 hour with gentle agitation and washed as described above. Protein detection was carried out using the ECL Plus Western Blotting Detection Reagent kit (Amersham Pharmacia Biotech) with chemiluminescent detection as per manufacturer's instructions. Protein bands were visualised by exposure to Kodak BioMax film for 30 seconds to 20 minutes depending on signal intensity.

2.3.20 Serum harvest from whole blood

Whole blood samples in 9 ml Vacuettes (Greiner Bio-one) were centrifuged at 2,000 rpm for 15 minutes at 4 °C to separate serum from particulate blood matter. Separated serum was harvested and stored at -20 °C until required.

2.3.21 ELISA

OPN concentration in cell culture media and sera samples was determined using in-house sandwich ELISA. Wells of a 96-well EIA/RIA Flat Bottomed Plate (Corning Costar®) were coated with 100 µl capture antibody (3 µg/ml R&D Systems monoclonal anti-OPN in PBS; see Table 2.3) overnight, washed three times with 0.05% PBS-T (See Appendix 1) and blocked with 200 µl 5% skim milk (Diploma) in PBS for 2 hours. After washing again with 0.05% PBS-T, cell culture supernatant (diluted 1/2 in PBS + 1% (w/v) BSA) or sera samples (diluted 1/3 and 1/5 in PBS + 1% (w/v) BSA) were incubated in duplicate for 1 hour. Serial dilutions of purified recombinant human OPN protein (R&D Systems) were used to create a standard curve. Plates were washed and probed with detection antibody (200 ng/ml R&D Systems polyclonal anti-OPN in 0.05% PBS-T; see Table 2.3) for 1.5 hours. Following thorough washing, wells were incubated in 0.1 ng/ml peroxidase-conjugated streptavidin (Rockland) for 1 hour, washed, then incubated in SigmaFast™ O-Phenylenediamine dihydrochloride tablet set (Sigma; made to manufacturer's directions in dH_2O) for 15 minutes to provide substrate for colour detection. Once a yellow colour had developed, the reaction was stopped by adding 50 µl 2.5 M HCl and A_{490} readings were taken on a MR5000 plate reader (Dynatech).

2.3.22 *In vivo* tumor growth in nude Balb/c mice

Five-week-old female Balb/c nude mice were injected subcutaneously into both dorsal flanks with 5×10^6 cultured cells (either Huh-7 or stably transfected OPN cell lines) suspended in 100 µl PBS, from which subcutaneous tumors formed. Cells were injected between the skin and peritoneal cavity using a sterile 1 ml syringe and 26 G BD PrecisionGlide™ Needle (BD Biosciences). Tumor diameter was measured every day using a caliper ruler. Tumor volume was determined by the following equation:

$$([\text{longest diameter}]^2 \times [\text{shortest diameter}]^2) / 2$$

At the conclusion of the experiment, blood was harvested from animals by heart venupuncture and all mice euthanised and tumors excised. One tumor portion was fixed in 10% buffered formalin for subsequent histological analysis, another stored in RNAlater® at 4 °C for RNA extraction and the last portion snap frozen in liquid nitrogen and stored at -80 °C for protein extraction. Formalin fixed tumor portions were processed for histological analysis by Dr. Andrew Ruszkiewicz of SA Pathology, Adelaide, Australia.

2.3.23 Murine blood collection via eye bleed

Eye bleeds were performed on nude Balb/c mice prior to initiation of subcutaneous tumor growth to obtain a baseline serum sample for OPN ELISA. Mice were anaesthetised by isoflurane inhalation for 2 - 5 minutes until non-responsive to pressure on the footpad. Blood was harvested by gently rupturing a blood vessel directly behind the eye with a sterile glass capillary tube (Livingstone). Samples were pooled into a 1.5 ml Eppendorf tube and serum isolated by centrifugation at 2,000 rpm for 5 minutes and harvest of the upper layer. Serum was stored at -20 °C until required.

2.3.24 Murine blood collection via heart venupuncture

Heart venupunctures were performed on nude Balb/c mice at the conclusion of subcutaneous tumor growth studies to obtain serum for OPN ELISA. Mice were anaesthetised by isoflurane inhalation as described above. A 26 G BD PrecisionGlide™ Needle (BD Biosciences) was inserted directly into the heart and blood collected into a sterile 2 ml syringe (BD Biosciences). Blood was transferred to a sterile 1.5 ml Eppendorf tube, from which serum was isolated as described previously. Serum was stored at -20 °C until required. Following heart venupuncture, all mice were euthanised.

2.3.25 Mouse urine collection

Urine from tumor-bearing nude Balb/c mice was collected at the conclusion of subcutaneous tumor growth studies to allow for analysis of excreted human OPN. Immediately following euthanasia of mice, a 1.5 ml Eppendorf tube was used to capture urine expelled from the animal due to involuntary muscle spasms during animal death. Urine was stored at -20 °C until required.

2.3.26 Immunohistochemical staining

Paraffin embedded sections of human liver and Huh-7 derived tumor tissue from mice were used for immunohistochemical staining. 4 micron sections were de-waxed by incubations in two treatments of 100% xylene (5 minutes), followed by rehydration in two washes of 100% ethanol, one of 90% (v/v) ethanol and one of 70% (v/v) ethanol, each for 10 minutes. After two 5 minute washes in PBS, antigen retrieval was performed by boiling slides in 10 mM sodium citrate (pH 6) in a microwave for 3 minutes on high power and 10 minutes on low power. Slides were cooled to room temperature and endogenous peroxide inactivated by incubating slides in 0.5% (w/v) hydrogen peroxide (in PBS) for 10 minutes on a stirring block. Slides were again washed twice for 5 minutes in PBS then incubated in the appropriate concentration of primary antibody (diluted in Antibody Diluent [Dako]; Table 2.3) for 1 hour. Slides were again washed in PBS twice for 5 minutes each and incubated in ADVANCE^{TM} HRP-Link (Dako) for 30 minutes, washed in PBS, then incubated in ADVANCE^{TM} HRP Enzyme (Dako) for 30 minutes. Slides were again washed and incubated in diluted DAB+ Chromagen (Dako; 20 µl DAB+ Chromagen in 1 ml DAB+ Substrate Buffer) for 10 minutes. Following a 5 minute wash in dH_2O, slides were counterstained with Gill's Haematoxylin (see Appendix 1) for 5 minutes then tap water for a further 5 minutes. Slides were then dehydrated in two washes of 50% ethanol for 5 minutes each, then two of 100% (v/v) ethanol and two of 100% (v/v) xylene each for 10 minutes. After air drying slides were mounted onto glass cover slips using DePeX mounting media (BDH). Slides

were evaluated in a blind manner by an expert liver pathologist (Dr. Andrew Ruszkiewicz, SA Pathology, Adelaide, Australia).

2.4 Data Analysis

All data was initially analysed and stored in Microsoft Excel 2003. Statistical analysis by way of unpaired student t-tests was performed utilising SPSS 9.0 (SPSS Inc.) and all graphing and statistical analysis by way of ANOVA performed in GraphPad Prism 5.

CHAPTER 3: Expression and quantification of osteopontin and its alternatively spliced variants in cultured cell lines and liver tissue

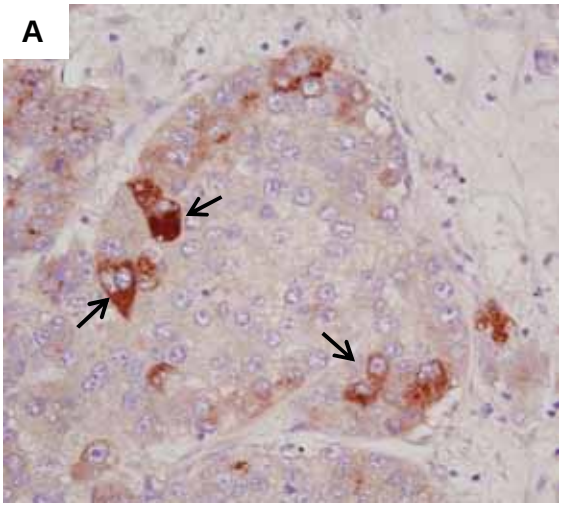
3.1 Introduction

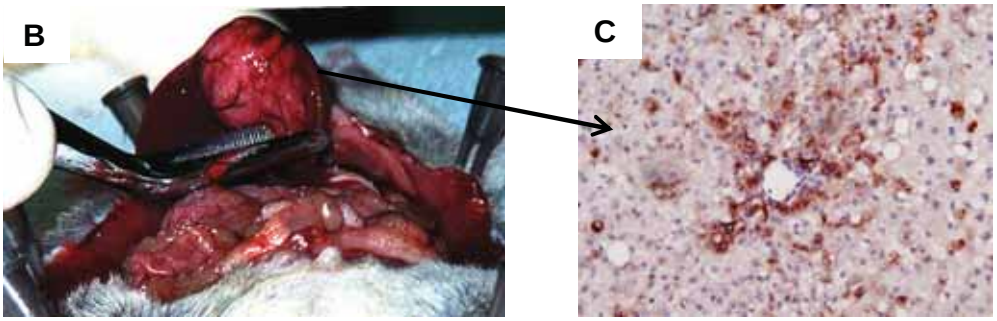
Expression of OPN at the mRNA level has been well documented in a wide range of tumorigenic cell lines including those derived from liver tumors (Gao et al., 2003; Sahai et al., 2004; Said et al., 2005; Forootan et al., 2006; Shevde et al., 2006; Emani et al., 2008; Zhao, J., Lu et al., 2008). However, analysis of cell lines and clinical HCC specimens has largely focused on the full length OPN-A mRNA transcript. Two splice variants for OPN also exist (see Section 1.3.5) termed OPN-B and –C and their expression and role in HCC biology (and other tumors) has been relatively unexplored. Alternative splicing is extremely important in expanding protein function of the approximately 30 - 40,000 genes present in humans (Graveley 2001). It is therefore highly likely that these variants of OPN may have specific functions in different cellular processes.

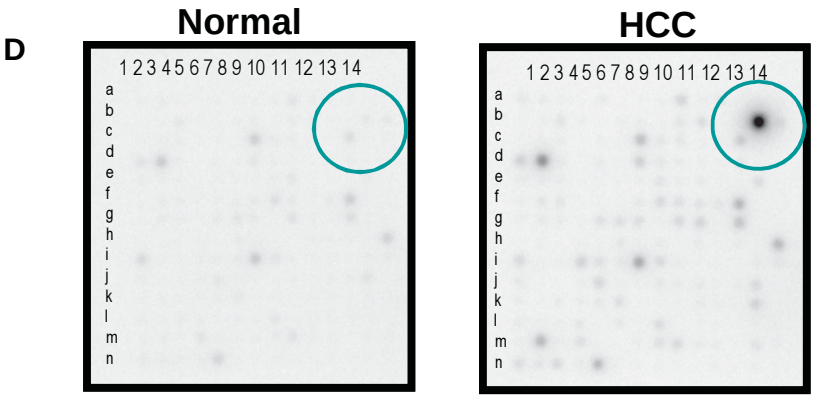
Through the course of microarray studies conducted in the Beard laboratory to investigate transcriptional profiles in HCC and non-diseased liver we noted a significant increase in OPN mRNA that not only consisted of OPN-A but also significant levels of the two splice variants OPN-B and –C. The identification of OPN variants in HCC has not been explored and thus **the first aim of this chapter was to investigate expression of OPN variants in HCC derived cell lines and clinical HCC samples.**

Studies of OPN expression in human liver disease have largely been restricted to HCC, with tumorigenic hepatocytes expressing OPN at the leading edge of tumor nodules as shown by the principal supervisor (Figure 3.1a) and others (Gotoh et al., 2002; Xie, H. et al., 2007). Previous

Figure 3.1 OPN protein is highly expressed in both human and murine HCC. **(A)** Strong OPN staining is observed in the cytoplasm of tumorigenic hepatocytes **(→)** in human HCC at the leading edge of a regenerating nodule. **(B)** A HCC tumor in the liver of a HCV transgenic mouse (Lerat *et al.*, 2002). **(C)** OPN IHC of the tumor shown in **(B)** shows strong OPN staining in the cytoplasm of tumorigenic hepatocytes. **(D)** Microarray analysis (Atlas Arrays, Clontech) of the murine HCC shown in **(B)** indicates that OPN expression is increased in HCC (right) compared to cognate non-diseased liver (left). All data was generated by M. Beard at The University of Texas Medical Branch, Galveston, Texas, USA.


A


B
C


D
Normal
HCC
1 2 3 4 5 6 7 8 9 10 11 12 13 14
a b c d e f g h i j k l m n

work by Dr. Beard and colleagues has also visualised strong OPN expression in tumorigenic hepatocytes of murine HCC (Figure 3.1c) in a HCV-transgenic mouse system (Figure 3.1b), and confirmed that OPN expression is significantly increased in murine HCC compared to surrounding non-neoplastic liver through the use of DNA microarray analysis (Figure 3.1d). In non-diseased liver, OPN is specifically expressed in bile duct epithelial cells and, to a lesser degree, Küpffer cells and HSCs (Kawashima et al., 1999; Lee, S. H. et al., 2004; Tajiri et al., 2005). No expression has been documented in normal hepatocytes (Gao et al., 2003). However, a recent study documented increased OPN expression in HBV-infected cirrhotic livers (Zhao, L., Li et al., 2008) while microarray analysis of HCV-related liver disease suggested OPN expression was also increased (Lau et al., 2005). To date there has been very little work on OPN expression in HCV-related liver disease, therefore **the second aim of this chapter was to investigate mRNA expression of OPN in HCV-related liver disease and correlate expression with histological stage of liver disease.**

While OPN expression has been well documented, characterisation of the splice variants has been relatively ignored. The quantification of individual splice variants by absolute or relative methods can be a powerful tool in the diagnosis and prognosis of disease. **The final aim of this chapter was to develop a method whereby we could individually quantify the three variants of OPN in a given cell culture or tissue sample.** This was attempted using two techniques, a real-time quantitative RT-PCR method and a GeneScan® fluorescent genotyping PCR method, with stringent validation methods.

3.2 Results

3.2.1 mRNA expression of OPN and its variants in cultured cell lines

Expression of OPN mRNA representing the full length OPN gene was examined in a panel of cultured human, murine, rat and monkey cell lines of both hepatocyte and non-hepatocyte origin, including cell lines from a number of human cancers. Total cellular RNA was harvested from cultured cells, cellular DNA degraded by *DNAse*I treatment, and cDNA prepared. Full length OPN mRNA expression was detected by 30 cycle RT-PCR using primers OPN-F2 and OPN-R3 (see Table 2.1). OPN was expressed in almost every cell line tested, including the human hepatoma cell lines Huh-7, HepG2, Hep3B, and PLC/PRF/5, the non-hepatocyte derived carcinoma cell lines HeLa (cervical) and MDA-MB-231 (breast), the murine hepatoma cell line Hepa1-6, the rat hepatoma cell line H-4-II-E, and the African green monkey kidney cell lines Cos-7 and Vero (Figure 3.2). OPN was not expressed in the breast cancer cell line MCF-7 (consistent with previous reports by He *et al.,* 2006), nor the no template control as expected. Interestingly there was significant variation in the amount of OPN mRNA present, ranging from very little in Hep3B and H-4-II-E to significant levels in HepG2 and Cos-7.

Next we investigated expression patterns of the three OPN variants in the same cell lines by RT-PCR as described above. Primers OPN-A/B/C-F and OPN-A/B/C-R (Table 2.1) were designed to sit within exons 2 and 6 of the OPN gene, allowing for concurrent amplification of all three variants from a 35 cycle PCR (Figure 3.3). Most human cell lines showed expression of all three variants at varying levels (Figure 3.4 top and middle). The majority of human cell lines (including Huh-7, HepG2, Hep3B, and HeLa) showed highest expression of OPN-A, lower expression of OPN-B (at up to half that seen for OPN-A), and relatively low expression of OPN-C. However, PLC/PRF/5 showed equal expression of OPN-B and OPN-C but no OPN-A,

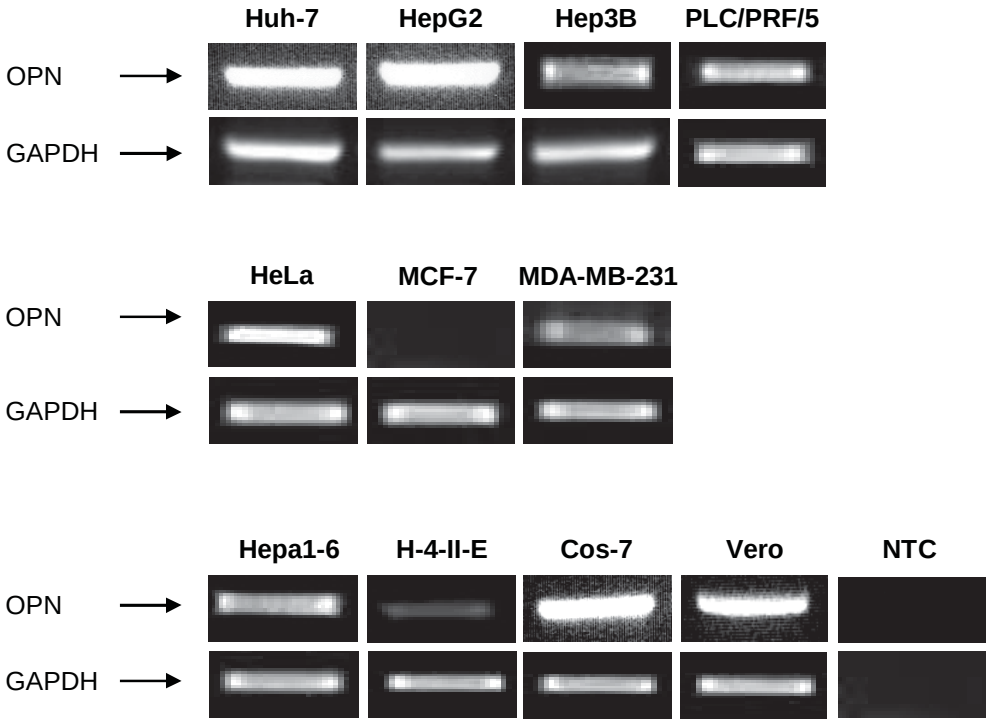


Figure 3.2 Full-length OPN mRNA expression (30 cycle PCR) in a range of cultured cell lines, including human hepatoma (top), human non-liver carcinoma (middle), and rodent hepatoma and African green monkey kidney (bottom) cell lines. OPN was expressed in all cell lines tested except MCF-7. The negative control sample contains water in place of cDNA (no template control, NTC). The housekeeping gene GAPDH served as an internal loading and integrity control. The primers for GAPDH are able to amplify product from human, mouse and rat templates.

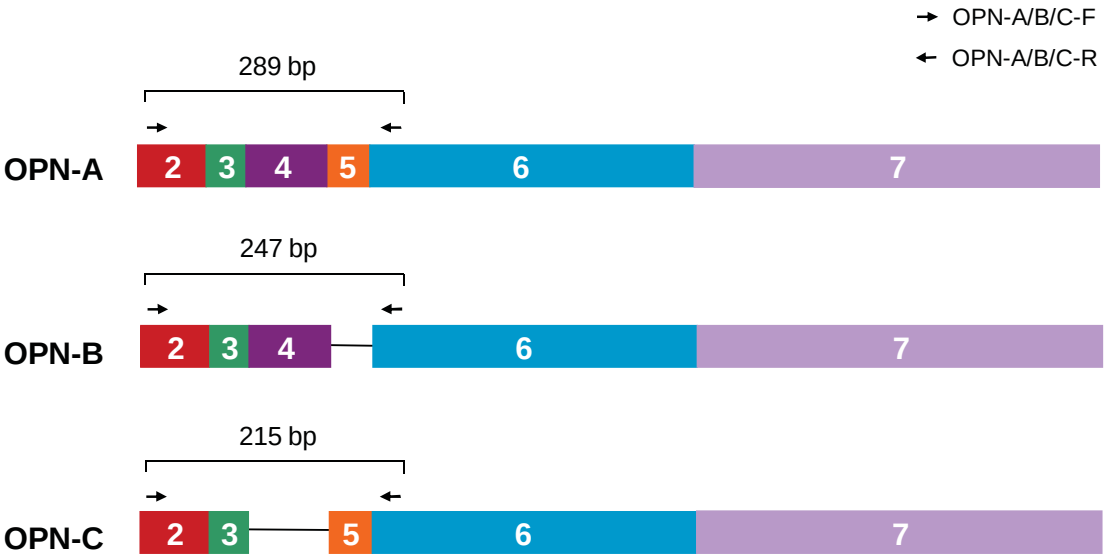


Figure 3.3 Location of primers for concurrent amplification of the three OPN variants. Primers are located within exons 2 and 6 to flank the missing exons 5 and 4 of OPN-B and OPN-C respectively. Expected amplicon size for each variant is noted above.

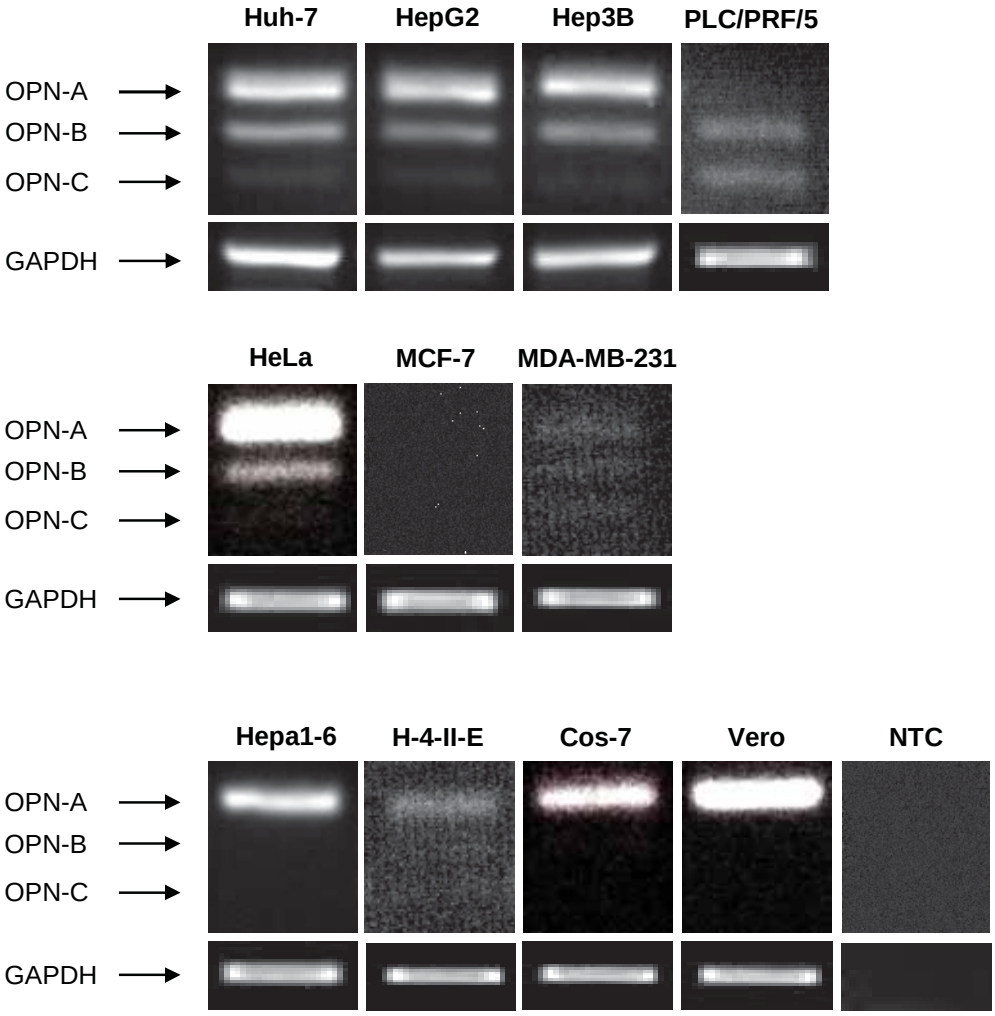


Figure 3.4 Expression patterns of all three OPN variants (35 cycle PCR) in a range of cultured cell lines, including human hepatoma (top), human non-liver carcinoma (middle), and rodent hepatoma and African green monkey kidney (bottom) cell lines. All cell lines showed expression of all variants except PLC/PRF/5 (which lacked OPN-A) and MCF-7 which showed no OPN expression as expected. OPN-A was the dominant species for the majority of cell lines.

MDA-MB-231 which showed equally low expression of all three variants, and MCF-7 did not express any of the variants as expected from the lack of OPN expression found previously (see Figure 3.2). No splicing was evident in the mouse, rat or African green monkey cell lines (Figure 3.4 bottom). No expression was noted in the no template control as expected.

3.2.2 mRNA expression patterns of OPN and its variants in non-diseased, HCV-infected and HCC-tumor bearing liver tissue

Having established OPN and OPN variant expression in liver derived cell lines, we next investigated OPN mRNA expression in liver tissue from non-diseased individuals (normal human liver; NHL), patients chronically infected with HCV (CHC), and HCV-HCC together with cognate cirrhotic tissue. Semi-quantitative expression was monitored by RT-PCR using primers OPN-F2 and OPN-R3 as described in Section 3.2.1. OPN expression was visualised in all samples tested (Figure 3.5). While only semi-quantitative, all CHC samples showed higher OPN expression than NHL; of these samples, the patient with probable cirrhosis (CHC 13) showed visually higher expression of OPN than patients with earlier stages of liver disease (CHC 6: no fibrosis and CHC 10: mid level fibrosis as determined by METAVIR scoring of liver sections). OPN expression varied between sample pairs for HCC patients (HCC tumor tissue and cognate cirrhotic, non-neoplastic liver tissue); only one patient showed higher OPN expression in HCC compared to surrounding non-neoplastic tissue (Patient 2). No OPN was seen in the no template control as expected.

Expression patterns of the three OPN variants were visualised in the same tissue samples by RT-PCR as described previously. Expression of all three variants was observed in NHL, with OPN-A being the dominant species (Figure 3.6 top) as noted in cell lines (see Figure 3.4). This therefore suggests that OPN splicing occurs in the non-diseased liver and contradicts the claim of He *et al.,* (2006) that OPN splicing is a phenomenon of oncogenic cells. Similar patterns were

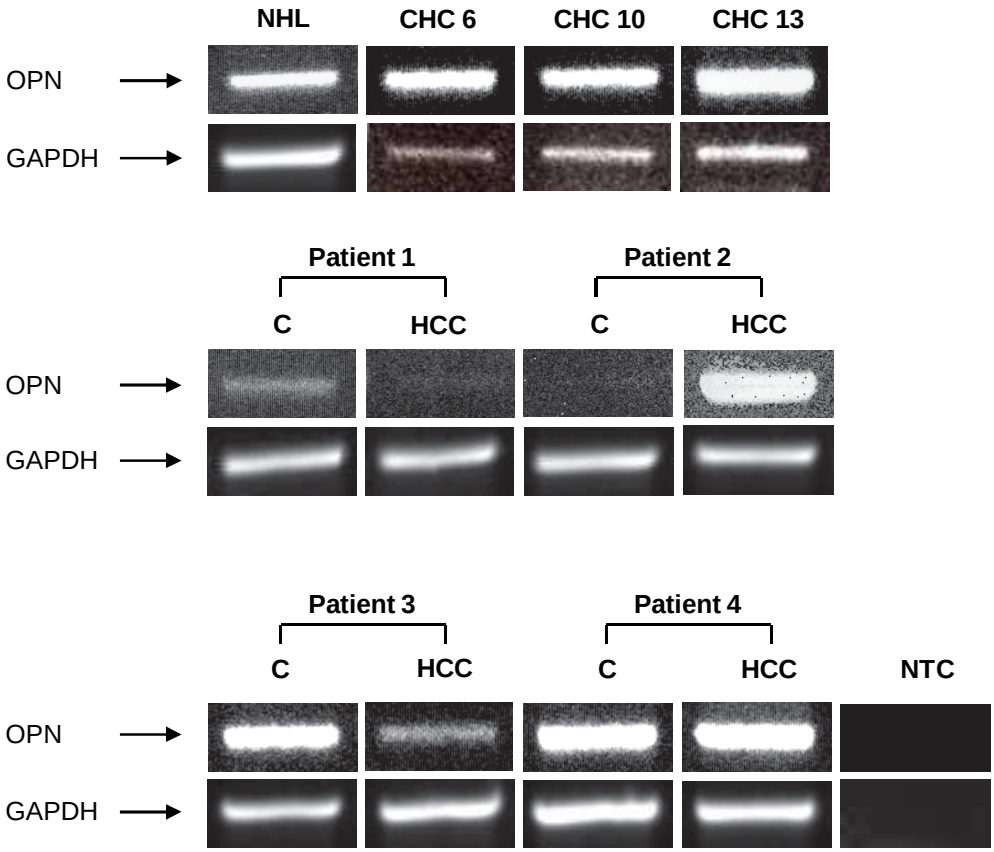


Figure 3.5 mRNA expression of full-length OPN in a range of human liver and tumor tissues, including normal human liver (NHL), chronic HCV-infected liver (CHC), HCC tumor tissue (HCC), and cirrhotic tissue surrounding HCC tumors from the same individual (C). OPN was visualised in all samples tested, albeit at varying levels.

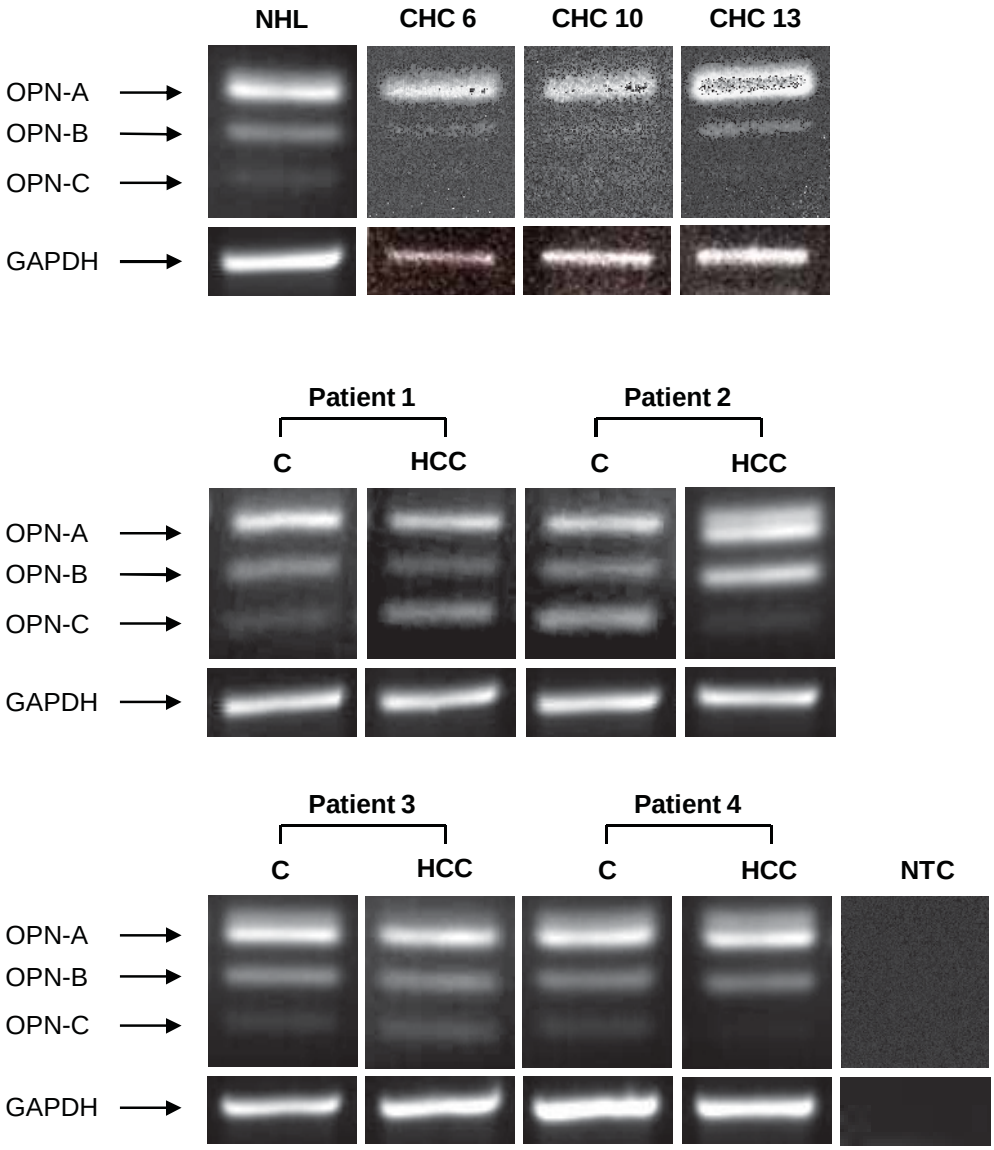


Figure 3.6 Expression patterns of OPN variants in a range of human liver and tumor tissues. All three OPN variants are expressed in normal human liver (NHL), chronic HCV-infected liver (CHC), HCC tumor tissue (HCC), and cirrhotic tissue surrounding the HCC of the same individual (C). All variants were expressed in all samples, with OPN-A the dominant species.

seen in liver tissue from individuals chronically infected with HCV (Figure 3.6 top). OPN-A was the dominant species for all paired HCC samples (Figure 3.6 middle and bottom), followed by OPN-B. While OPN-C was present in all samples (albeit weakly), its expression varied significantly. These results suggest for the first time that all OPN variants are expressed in cirrhotic and tumor tissue in the liver.

To compare OPN expression between human and murine HCC tissue, expression of OPN variants was also monitored in healthy mouse liver tissue (NML), and HCC derived from HCV transgenic mice (Lerat et al., 2002). No alternative splicing was observed in any sample, suggesting that alternative splicing of OPN is unique to humans (Figure 3.7). A noticeable difference in OPN-A expression was observed between paired tumor and non-neoplastic samples, whereby both murine HCC samples showed higher expression of OPN-A compared to their cognate non-neoplastic tissue. This is particularly evident for Mouse 1, and is consistent with results from DNA microarray experiments conducted using this tissue (Beard, M., unpublished data; see Figure 3.1d). No expression was noted in the no template control as expected.

3.2.3 Quantitation of OPN in different stages of HCV-related liver disease

The above results suggest that in addition to a role in HCC, OPN may also play a role in HCV-related liver disease prior to development of HCC. We therefore investigated quantitative mRNA expression of OPN in CHC of varying fibrosis grades using a real-time PCR approach. Liver samples from 20 patients were collected and a portion fixed in RNAlater® RNA Stabilisation Solution (Ambion) while the remainder was sent for histological analysis which was performed by Dr. Andrew Ruszkiewicz (Pathologist, SA Pathology). Samples were grouped by their level of liver fibrosis according to the METAVIR scoring system (Table 3.1).

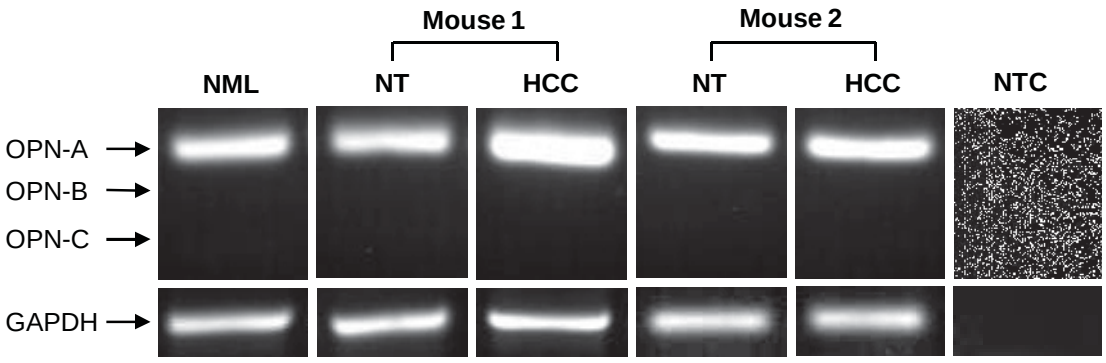


Figure 3.7 Expression patterns of OPN variants in a range of murine and rat normal liver tissue, HCC tumor tissue and cognate cirrhotic non-tumor tissue. No alternative splicing was noted in any sample.

Level of fibrosis	METAVIR fibrosis score	HCV Patient #
None visible	F = 0	6, 8, 12, 15, 22
Low	F = 1	11, 23, 29, 39, 50
Medium	F = 2/3	1, 4, 10, 24, 26
High; probable cirrhosis	F = 4	13, 14, 41, 51, 62

Table 3.1 Grouping of HCV liver samples based on METAVIR fibrosis score as assigned by Dr. Andrew Ruszkiewicz (Pathologist, SA Pathology)

Total cellular RNA was harvested from each liver biopsy and cDNA samples prepared as previously described. OPN expression for each sample was detected in duplicate by real-time RT-PCR using an ABI PRISM 7000 PCR machine and primers OPN-F1 and OPN-R1 (see Table 2.1) designed to amplify 76 bp of exon 7. Expression levels of the housekeeping gene RPLPO (aka 36b4) were also monitored using primers 36b4-F and 36b4-R (Table 2.1) to normalise levels of input cDNA between samples. A no template control containing water in place of cDNA was included in every experiment to investigate possible contamination.

Good amplification and melt curves were obtained for both OPN and RPLPO (see Appendix 2). OPN expression levels were monitored for each sample and expressed relative to normal human liver pooled from 4 healthy individuals (NHL; Figure 3.8a). Overall, relative OPN levels were low in patients with no (F = 0) or low (F = 1) fibrosis, but higher in those with medium (F = 2/3) or high (F = 4) fibrosis. Mean OPN levels were calculated for each group of patients (relative to NHL), with a trend indicating that OPN expression increases with increasing fibrosis grade (Figure 3.8b). This result however was not significant. The seemingly higher average OPN expression for the F = 0 group can be attributed to the high level recorded for patient 12, which was much higher than all other patient samples within that group. Clearly this analysis requires more samples to be tested before significant correlations can be drawn, however this is a

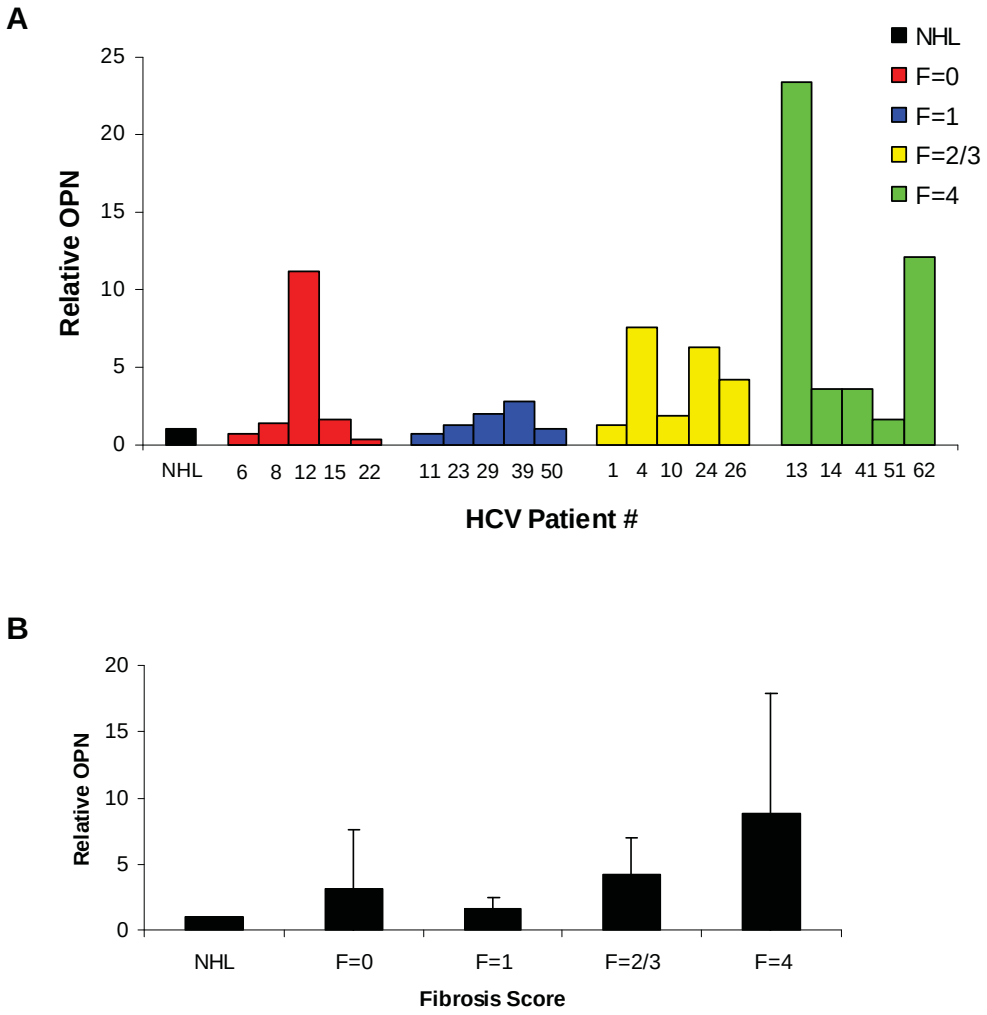


Figure 3.8 OPN expression levels in CHC livers with varying degrees of fibrosis. **(A)** OPN expression levels for individual HCV-infected liver samples. All OPN levels represent the fold change compared to pooled NHL. **(B)** Average OPN expression levels relative to NHL for each group of samples based on fibrosis score. A trend is evident whereby OPN expression increases with the progression of HCV-related liver disease.

significant trend towards increased OPN as liver disease progresses. This is further confirmed by the significant expression of OPN in the cirrhotic liver surrounding HCC as noted in Figure 3.4.

3.2.4 Quantitation of OPN in HCC tumor tissue compared to cognate non-neoplastic liver

OPN is known to be over-expressed in individuals with HCC (Gotoh et al., 2002), however relative expression in the surrounding non-neoplastic tissue compared to the tumor itself remains unknown. To investigate this, OPN expression was monitored by real-time RT-PCR and compared between HCC tumor tissues and their surrounding non-neoplastic tissues from 4 individuals with chronic HCV-infection. Positive OPN expression was previously visualised in these samples by RT-PCR (see Section 3.2.2). cDNA samples were prepared and run in duplicate on the ABI PRISM 7000 as previously outlined. Two of the four individuals showed increased OPN expression in HCC tumor tissue compared to cognate non-neoplastic tissue (Figure 3.9, Patient 2: 208-fold and Patient 4: 7-fold), one individual showed decreased OPN expression in HCC tissue (Patient 3: 25-fold) and one individual showed no change in OPN expression between the paired tissues (Patient 1). Within this small sample size, 50% of HCC-bearing individuals showed marked increases in OPN compared to the surrounding tissue. Lack of HCC tissue samples precluded a more detailed analysis, although our data are consistent with the literature.

3.2.5 Quantitative assessment of OPN variant expression: real-time RT-PCR

The ability to accurately quantitate OPN variant expression has not been reported to date. Providing such an assay may be useful to predict tumor progression, treatment response and metastatic potential. The most utilised method for quantifying splice variants is real-time quantitative PCR, using specific primers for each variant. In splicing events that result in inclusion of an intron or alternative exons, unique nucleotide sequences will be present that allow for specific primers to be designed for each variant. However, in the case of exon deletion (which

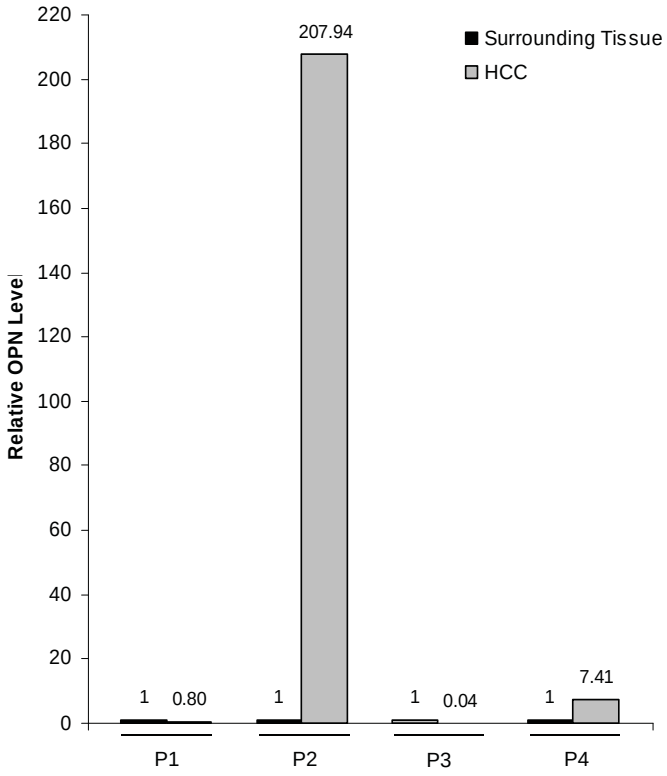


Figure 3.9 Relative OPN expression in HCC tissue compared to cognate non-neoplastic tissue. OPN expression of each non-neoplastic tissue sample is set at 1 and expression in each HCC tissue sample is shown relative to its paired non-neoplastic tissue. OPN expression was varied, however 50% of patients showed greatly increased OPN in HCC compared to cognate tissue. Px = patient x.

occurs during OPN splicing), these unique sequences are not present. Therefore, novel exon-exon junction sequences must be utilised in the hope of designing primers that will amplify each variant individually. Expression can be calculated as an absolute concentration, or relative amount compared to the other variants. A recent study was able to individually amplify and quantitate the three OPN splice variants in human breast cancer tissue (Mirza et al., 2008), however no mention is made of the specificity of each primer set to its corresponding variant or any validation methods used to confirm the validity of the protocol. Another technique, GeneScan® Genotyping PCR, utilises GeneMapper ID v3.2 (Applied Biosystems) to analyse and quantitate fragments of DNA via fluorescently labeled primer amplification. It is used extensively in applications where precise quantification and sizing of fragments are required, including microsatellite analysis, single nucleotide polymorphism (SNP) validation and screening, and amplified fragment length polymorphism (AFLP) analysis.

3.2.5.1 Primer design and validation

Individual primer sets were utilised to amplify each of the three OPN transcript variants. The unique exon-exon junctions created by the lack of exons 5 and 4 in OPN-B and OPN-C respectively were exploited to create primers that would specifically amplify each of the three variants with no cross-reactivity. Therefore, one primer of each pair was designed to lie over the unique exon-exon junctions found in each variant; namely, the exon 4-5 junction for OPN-A, exon 4-6 junction for OPN-B, and exon 3-5 junction for OPN-C. Primer sequences for OPN-A (OPN-A/qPCR-F and OPN-A/qPCR-R) and OPN-C (OPN-C/qPCR-F and OPN-C/qPCR-R) were sourced from a previous publication (He et al., 2006), whilst a unique primer set was designed for OPN-B (OPN-B/qPCR-F and OPN-B/qPCR-R).

To evaluate the specificity of each primer pair, a standard PCR was performed using 10 ng of each OPN variant vector alone and in combination as a template. A no template (water only)

control was also included. All three primer pairs were also tested on Huh-7 cDNA to ensure that correct sized amplicons could be amplified from human samples. All primer pairs amplified bands of expected size (207, 177 and 149 bp respectively for OPN-A, OPN-B and OPN-C), however cross-reactivity was seen for all primer pairs (Figure 3.10) and it was therefore determined that, in practical terms, these primers sets were not variant-specific and were of no use in real-time quantitative RT-PCR.

In collaboration with Dr. Eric Smith (Department of Surgery, University of Adelaide), a second set of OPN variant real-time PCR primers were designed. These primers also exploited the unique exon-exon junctions but differed slightly in sequence to the first primer set (Figure 3.11). A standard PCR was used to check that each primer set would amplify a band of correct size from validated OPN templates and Huh-7 cDNA samples as performed above. All three primer sets successfully amplified bands of correct size (OPN-A: 69 bp, OPN-B: 188 bp, OPN-C: 72 bp) from each of the templates with no cross-priming (Figure 3.12).

3.2.5.2 Evaluation of primer efficiency

If we were to employ real-time RT-PCR as a tool to determine quantitative assessment of OPN variants then it is imperative that all primer pairs are working at the same efficiency. Primer efficiency is a measure of the proportion of template successfully amplified during each PCR cycle. For example, an efficiency value of 1 indicates that every single template sequence is being duplicated during each PCR cycle, resulting in an exact doubling of template amount (which is the ideal situation). Primer pair efficiency can be calculated by plotting a standard curve of C_T value versus known input template amount, determining the slope of the curve, and solving the following equation: Efficiency = $10^{-1/slope}$ -1. This calculation of primer pair efficiency is known as the Pfaffl method (Pfaffl 2001). A standard curve of average C_T value versus known input vector amount was generated to evaluate the efficiency of each primer set (Figure 3.13).

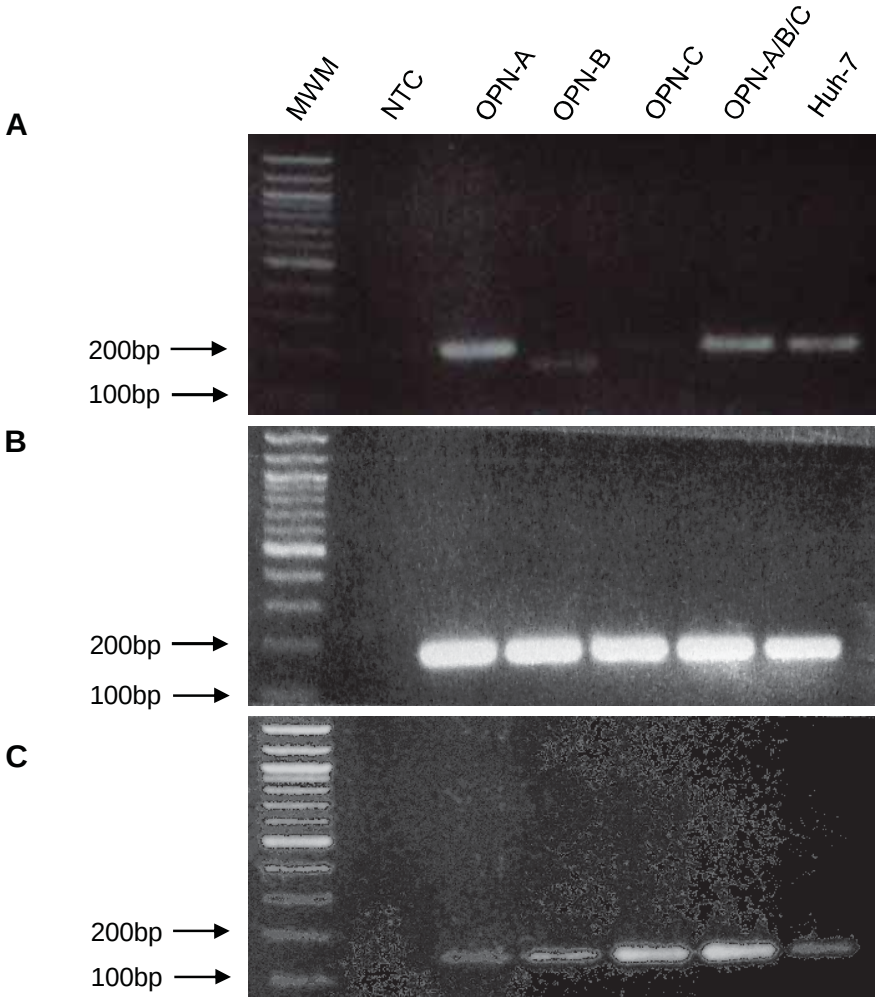


Figure 3.10 Amplification of the three OPN variants from samples containing variant vectors individually (OPN-A, OPN-B, OPN-C) or in combination (OPN-A/B/C), and from Huh-7 cDNA with individual primer sets designed for use in real-time quantitative RT-PCR. **(A)** Amplification of all vector samples with the OPN-A primer set. **(B)** Amplification of all vector samples with the OPN-B primer set. **(C)** Amplification of all vector samples with the OPN-C primer set. Cross-reactivity was seen for all primer pairs, therefore they were deemed unusable for real-time RT-PCR quantitation of the OPN variants. MWM = molecular weight marker.

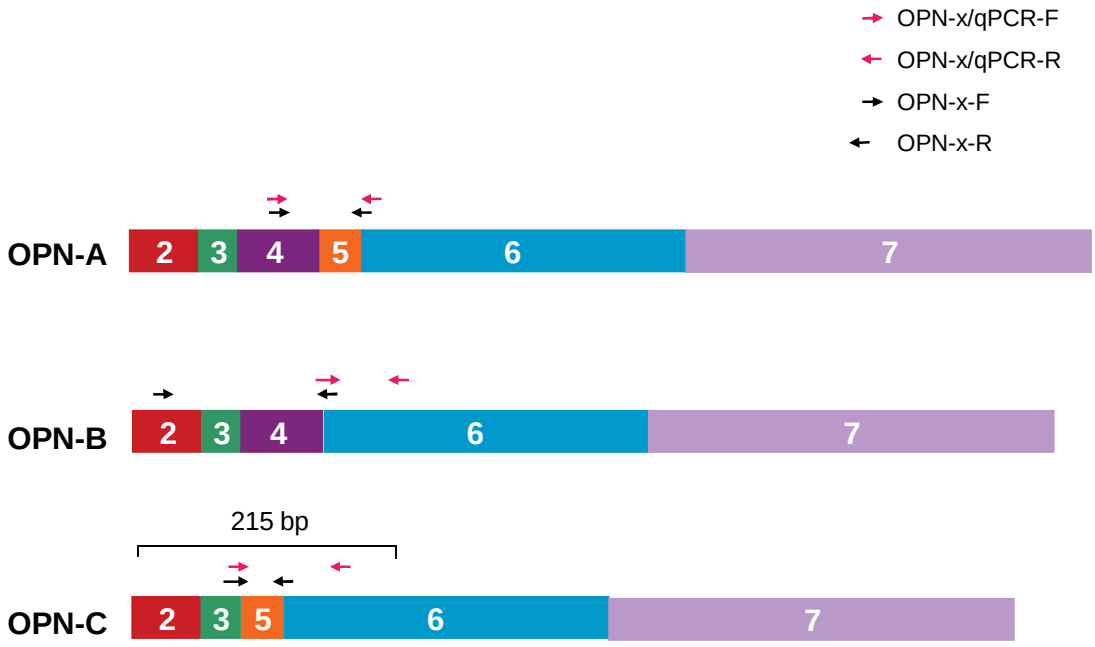


Figure 3.11 Location of primers for real-time RT-PCR amplification of the three OPN variants. The first set of primers designed are indicated in pink (→); these were deemed to be of no use in quantifying OPN variants using real-time RT-PCR. The second set of primers designed in conjunction with Dr. Eric Smith are indicated in black (→). Note that for the second set of primers, the reverse primers for OPN-A and OPN-C are identical.

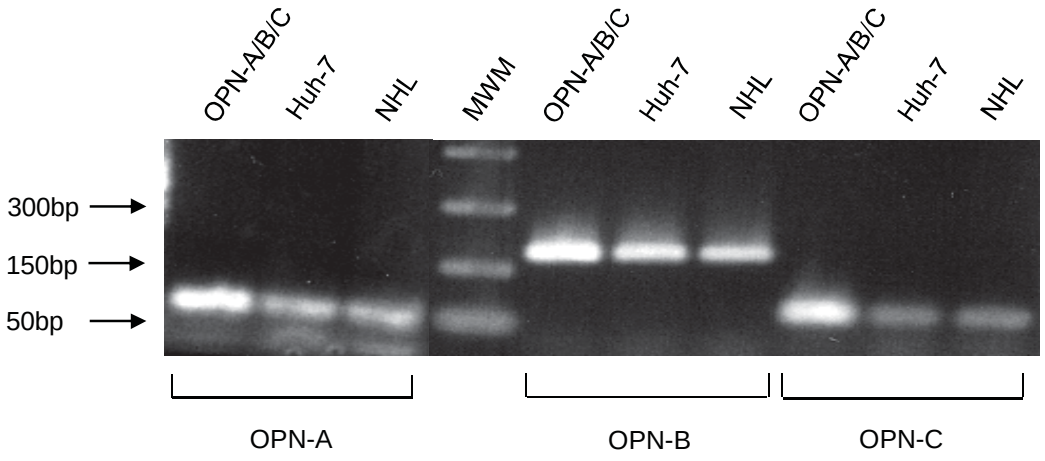
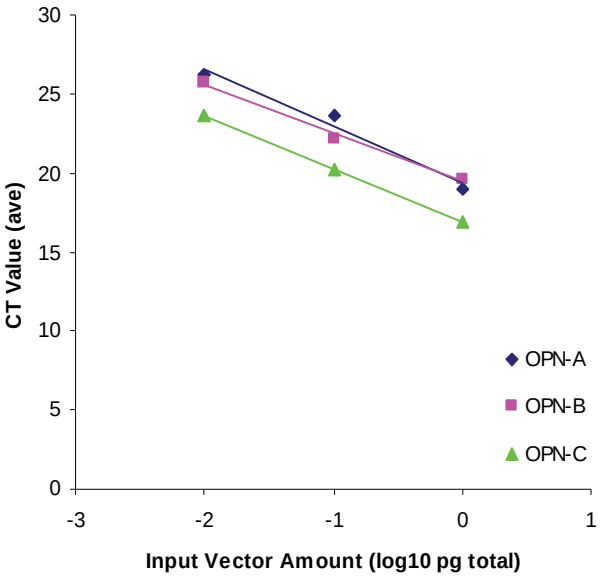


Figure 3.12 Amplification of the three OPN variants with a second set of primers for real-time quantitative RT-PCR. Each of the three primer sets was tested against a mixture containing 10 ng each of the three variants (OPN-A/B/C), Huh-7 cDNA and NHL cDNA . All primer pairs amplified bands of expected size from all templates with no cross priming.



Primer Pair	R^2	Slope	Efficiency
OPN-A	0.9748	-3.62	0.89
OPN-B	0.9929	-3.08	1.11
OPN-C	0.9995	-3.34	0.99

Figure 3.13 Evaluation of efficiency of OPN variant primer pairs. Standard curves of CT values were obtained for the three OPN variant primer pairs using known input amounts of corresponding OPN template (on a $\log_{10}$ scale). Trendlines for each standard curve were utilised to calculate the slope of each curve and the efficiency of each primer set. Data correlation was excellent and efficiency ideal for OPN-B and OPN-C primer pairs but not for OPN-A.

Data correlation was excellent for OPN-B and OPN-C ($R^2 > 0.99$) but not for OPN-A ($R^2 = 0.97$). Solving for the efficiency equation provided above, both OPN-B (1.11) and OPN-C (0.99) primers showed at or above ideal efficiency, however efficiency of OPN-A primers was less than ideal (0.89).

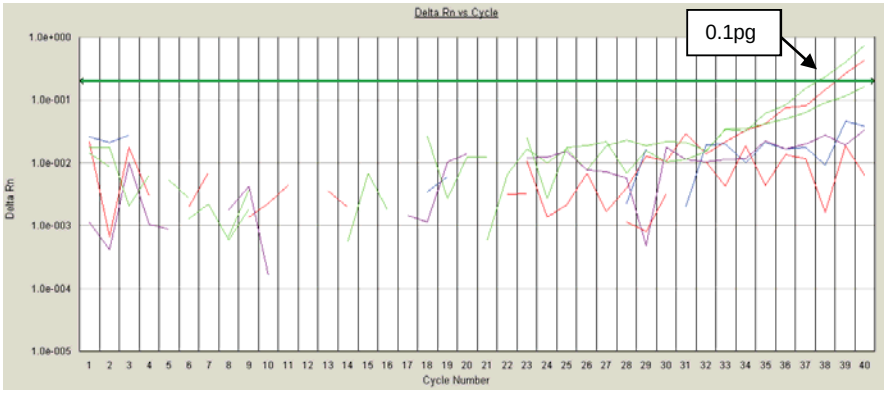
An attempt was made to increase the efficiency of OPN-A amplification through the design of a new forward primer. Two new forward primers were designed (OPN-A-F2 and OPN-A-F3, see Table 2.1), with both sequences shifted in the 3' direction in respect to the original forward primer (OPN-A-F) but still situated across the exon 4-5 junction. Each of the newly designed primers was paired with OPN-AC-R and tested against the original OPN-A primer pair (OPN-A-F + OPN-AC-R) to compare their specificity and efficiency during amplification of OPN-A. The OPN-A-F3 primer failed to amplify OPN-A at concentrations less than 0.1 pg (Figure 3.14a) and was therefore deemed unsuitable for OPN-A amplification. C_T values obtained for the new OPN-A-F2 primer (Figure 3.14b) were comparable to those obtained from the original OPN-A primer pair (Figure 3.14c).

A standard curve of average C_T value versus known input vector amount was generated to evaluate the efficiency of the primer pair containing OPN-A-F2 compared to the original OPN-A primer pair (Figure 3.15). The new OPN-A primer set showed much higher efficiency (1.02) than the original primer pair (0.57). The newly designed primer set of OPN-A-F2 + OPN-AC-R was deemed to be more efficient at specific amplification of OPN-A than the original primer pair.

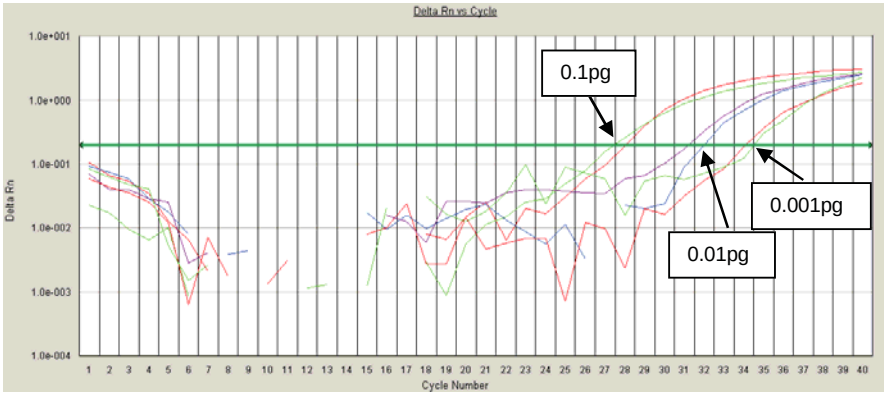
3.2.5.3 Ability of primers to accurately quantitate relative OPN levels from known input dilution series

Each OPN variant primer pair was tested under real-time PCR conditions with 10-fold dilution series of corresponding input OPN template. It was hoped that this would develop a standard

A



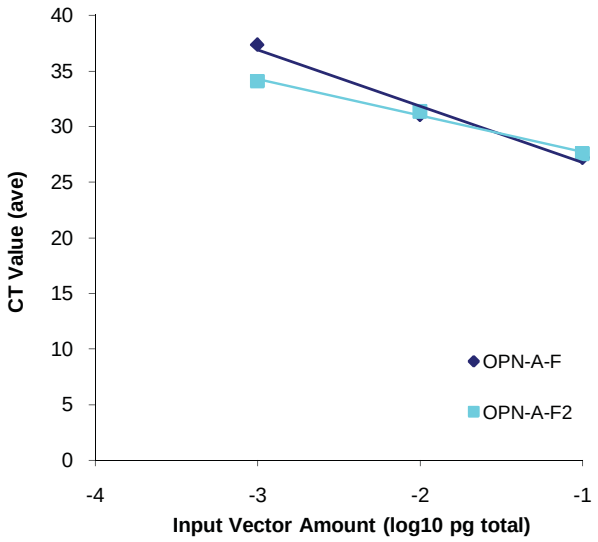
B



C

	Input Template Amount		
Forward Primer	**0.1 pg**	**0.01 pg**	**0.001 pg**
OPN-A-F	27.21	31.04	37.34
OPN-A-F2	27.54	31.33	34.29
OPN-A-F3	37.41	n.d.	n.d.

Figure 3.14 Amplification of OPN-A sequence using alternative OPN-A forward primers OPN-A-F3 **(A)** and OPN-A F2 **(B)**. C_T values generated by the OPN-A-F2 primer were comparable to those from the original primer pair (using OPN-A-F) **(C)**. n.d. = not detected.



Primer Set	R^2	Slope	Efficiency
OPN-A-F + OPN-AC-R	0.9806	-5.07	0.57
OPN-A-F2 + OPN-AC-R	0.9910	-3.27	1.02

Figure 3.15 Standard curve of CT value obtained for OPN primer pairs at varying concentrations with known input amounts of corresponding OPN variant template ($\log_{10}$ scale). Trendlines for each standard curve were utilised to calculate the slope of each curve and the efficiency of each primer set. The primer pair containing OPN-A-F2 showed increased efficiency compared to that containing OPN-A-F.

curve method by which absolute quantitation of OPN variant transcript levels could be performed on patient samples. Each primer pair was tested against 1 pg, 0.1 pg and 0.01 pg of their corresponding OPN template to derive a standard curve of C_T value against input vector amount. Due to the 10-fold difference in input template, an approximate 3.3-fold difference in C_T value was expected; a decrease of 1 C_T value indicates a doubling in the amount of PCR product ($2^1 = 1$), therefore a decrease in C_T value of 3.3 will result in an approximately 10-fold increase in PCR product ($2^{3.3} = 9.85$).

The OPN-C primer pair showed almost ideal C_T differences between 10-fold dilutions of input vector amount (3.21 for 0.1 pg:1 pg and 3.46 for 0.01 pg:0.1 pg; Figure 3.16c and Table 3.2), which corresponded very closely to 10-fold decreases in relative OPN expression (9.25 decrease from 0.1 pg:1 pg and 11-fold decrease from 0.01 pg:0.1 pg). OPN-B primers also gave C_T value decreases close to those expected (2.63 for 0.1 pg:1 pg and 3.53 for 0.01 pg:0.1 pg) with corresponding fold decreases of 6.17 and 11.51 respectively (Figure 3.16b and Table 3.2). OPN-A however did not follow the expected 3.3 C_T and 10-fold decrease pattern, with C_T decreases of 4.17 and 2.63 and OPN expression fold decreases of 18.00 and 6.19 for 0.1 pg:1 pg and 0.01 pg:0.1 pg (Figure 3.16a and Table 3.2). Melt curve analysis showed a single peak for each primer pair, indicating that a single product was obtained (data not shown). Control samples indicated no amplification for any primer pair from the backbone pRc/CMV vector as expected.

3.2.5.4 Ability of primers to accurately quantitate relative OPN levels from mixed input samples

Primer sets were tested against samples containing a mix of all three OPN variant templates to determine if the primers could successfully amplify the correct variant from a mixture of templates. Each primer pair was tested against a sample mixture containing 1 pg each of OPN-A, OPN-B and OPN-C templates, melt curves analysed and C_T values compared to those obtained

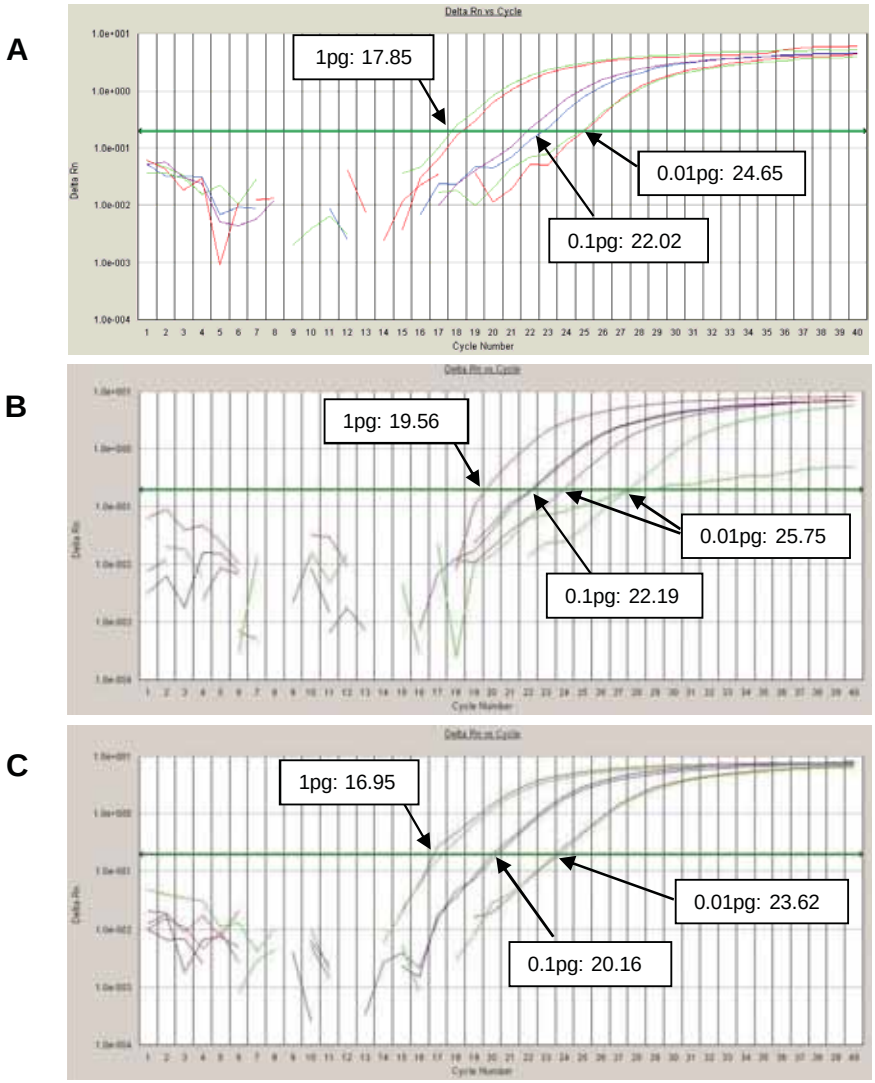


Figure 3.16 Validation of individual primer pairs for OPN-A **(A)**, OPN-B **(B)**, and OPN-C **(C)**. 10-fold dilutions of each variant template were prepared and subjected to quantitative real-time RT-PCR using their matching primer pairs. A difference in C_T value of 3.3 was expected down each dilution series to correlate with the known 10-fold difference in template amount. OPN-C gave almost ideal C_T differences down the dilution series, OPN-B also gave C_T differences close to ideal, however OPN-A did not. Average C_T values for each duplicate template are indicated.

Primer Pair	Input (OPN)	Ave C_T	C_T ↑	Fold ↓
OPN-A-F2 + OPN-AC-R	1 pg	17.85	-	-
OPN-A-F2 + OPN-AC-R	0.1pg	22.02	4.17	18.00
OPN-A-F2 + OPN-AC-R	0.01 pg	24.65	2.63	6.19
OPN-B-F + OPN-B-R	1 pg	19.56	-	
OPN-B-F + OPN-B-R	0.1pg	22.19	2.63	6.17
OPN-B-F + OPN-B-R	0.01 pg	25.71	3.53	11.51
OPN-C-F + OPN-AC-R	1 pg	16.95	-	-
OPN-C-F + OPN-AC-R	0.1pg	20.16	3.21	9.25
OPN-C-F + OPN-AC-R	0.01 pg	23.62	3.46	11.00

Table 3.2 Differences in C_T value and fold expression of OPN variants generated by a 10-fold dilution series of input OPN variant template. OPN expression from 10-fold dilutions of each of the three OPN variant templates was tested by their corresponding primer pairs. An increase in C_T value of 3.3 would correspond to a 10-fold decrease in expression.

from Section 3.2.5.3 where primer pairs were tested on their corresponding variant alone (Table 3.3).

Primer Pair	1 pg Individual Template	1 pg Template Mix	C_T Difference
OPN-A	17.85	16.04	1.81
OPN-B	19.56	19.16	0.40
OPN-C	16.95	17.28	0.34

Table 3.3 CT values of amplification of each of the three OPN splice variant by their primer pairs from samples containing the specific template individually or a mixture of all three templates

OPN-B and OPN-C primer pairs showed very similar C_T values compared to amplification from a single template, however this was not the case for OPN-A amplification. Dissociation analysis showed that the melting temperature of each variant amplified from the vector mixture was consistent with that seen from amplification of each variant individually (data not shown). Melt curve analysis was indicative of single product amplification. This data indicates that the primer sets can specifically amplify their corresponding OPN variant from a sample mix containing all OPN variants. However, discrepancies were noted with the OPN-A primer pair compared to single amplification of its specific template alone and thus these OPN-A primers may not be suitable for quantitative amplification.

3.2.5.5 Analysis of primer cross-reactivity

Chosen primer pairs for the three OPN variants were tested against individual 0.1 pg samples of both of their non-preferred OPN variant templates to evaluate cross-reactivity of the primer pairs. OPN-A primers were able to amplify products with similar melting temperatures from all three variant templates (Figure 3.17), as were primers for OPN-B (Figure 3.18) and OPN-C

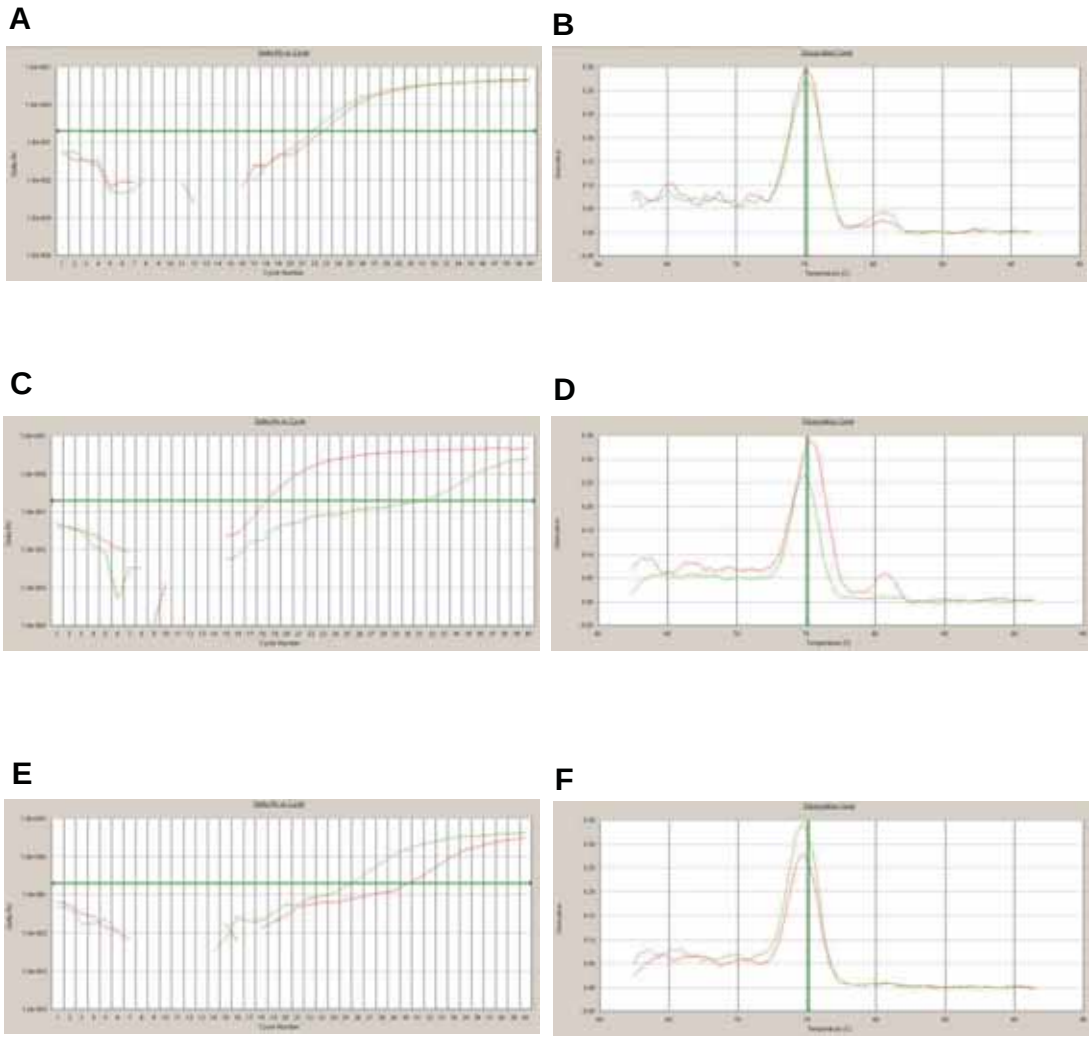


Figure 3.17 Cross-reactivity test of OPN-A primer pair. The primers OPN-A-F2 and OPN-AC-R were tested against 0.1 pg of OPN-A **(A-B)**, OPN-B **(C-D)** and OPN-C **(E-F)** to determine the specificity of the primers. This primer pair amplified OPN product from all OPN variant templates, indicating that it was not specific for OPN-A and could therefore not be used to individually quantify OPN-A.

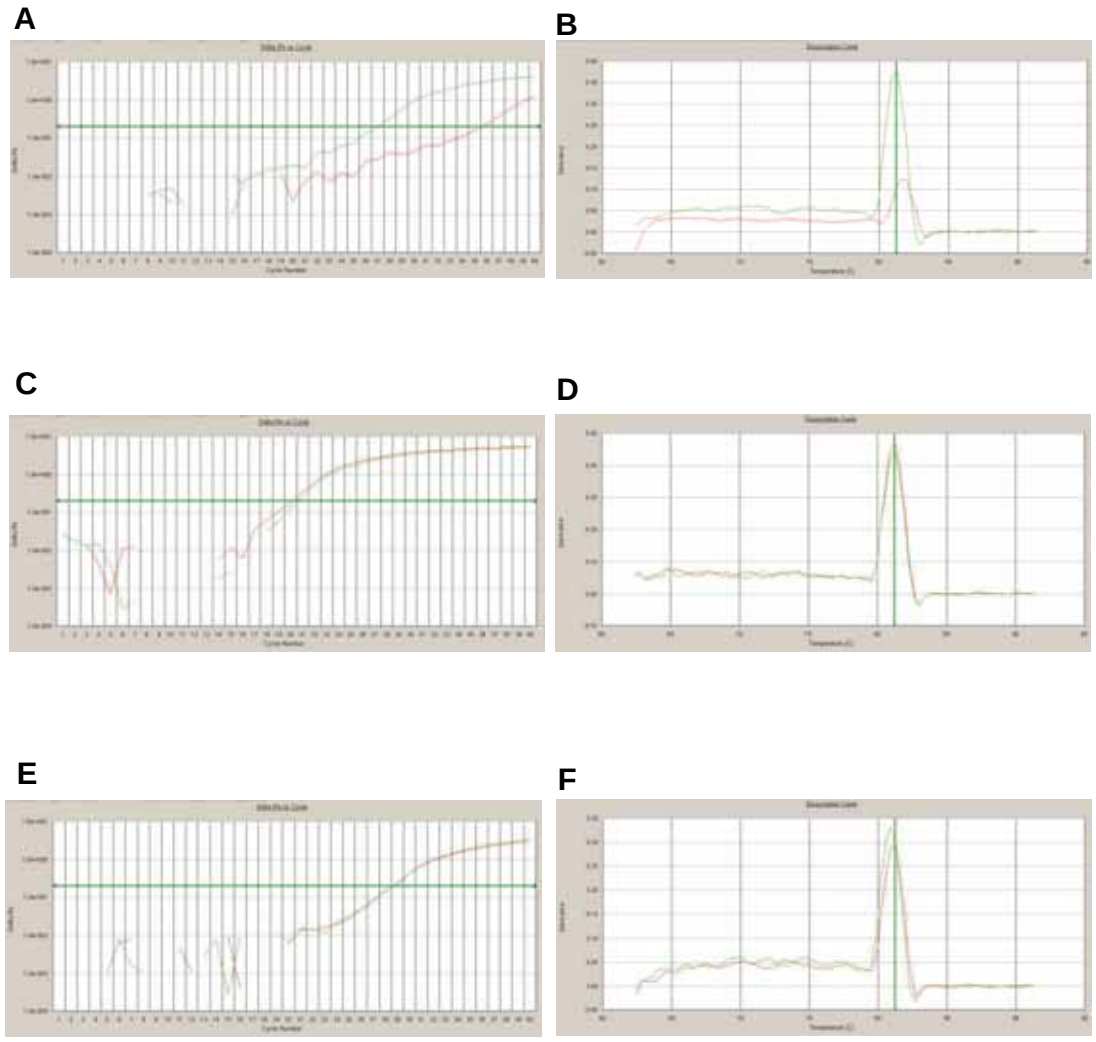


Figure 3.18 Cross-reactivity test of OPN-B primer pair. The primers OPN-B-F and OPN-B-R were tested against 0.1 pg of OPN-A **(A, B)**, OPN-B **(C, D)** and OPN-C **(E, F)** to determine the specificity of the primers. This primer pair amplified OPN product from all OPN variant templates, indicating that it was not specific for OPN-B and could therefore not be used to individually quantify OPN-B.

(Figure 3.19), indicating that none of the primer sets could specifically amplify their corresponding variant without cross-reactivity to the other variants of OPN.

Following these stringent validation techniques, it was determined that the three OPN variants could not be amplified individually from a single test sample without some degree of cross-amplification, and therefore this method could not accurately quantitate the OPN variants.

3.2.6 Quantitative assessment of OPN variant expression: GeneScan® Genotyping PCR

Given the unreliability of the quantitative real-time RT-PCR described above, we employed the GeneScan® Genotyping PCR method to quantitate OPN variants. This method was optimised and validated using known input amounts of OPN cDNA in an attempt to use this assay to quantitate OPN variants in cell lines and patient samples by determining the ratio of their corresponding immunofluorescence as measured by fluorescence peaks that indicate the amount of PCR product present. If successful, this method would provide an accurate, high throughput and reproducible method for measuring relative amounts of OPN variants in a given sample.

Primers were designed to flank exons 4 and 5 of the OPN gene to allow for simultaneous amplification of all three variants (Figure 3.20). PCR amplification using these primers generates amplicons of 182 bp, 140 bp, and 101 bp respectively for OPN-A, OPN-B, and OPN-C. The forward primer (OPN-F-Fam) was labeled with a fluorescent 5' 6-FAM tag to allow for quantitation of PCR product by level of fluorescence. FAM fluoresces in the blue range of the ultraviolet spectrum.

Initial experiments were performed to determine primer specificity and optimal cycle number using known input amounts of OPN template cDNA. A standard OPN PCR program was performed (see Section 2.3.7) with varying amounts of input vector (10 ng or 50 ng) and cycle

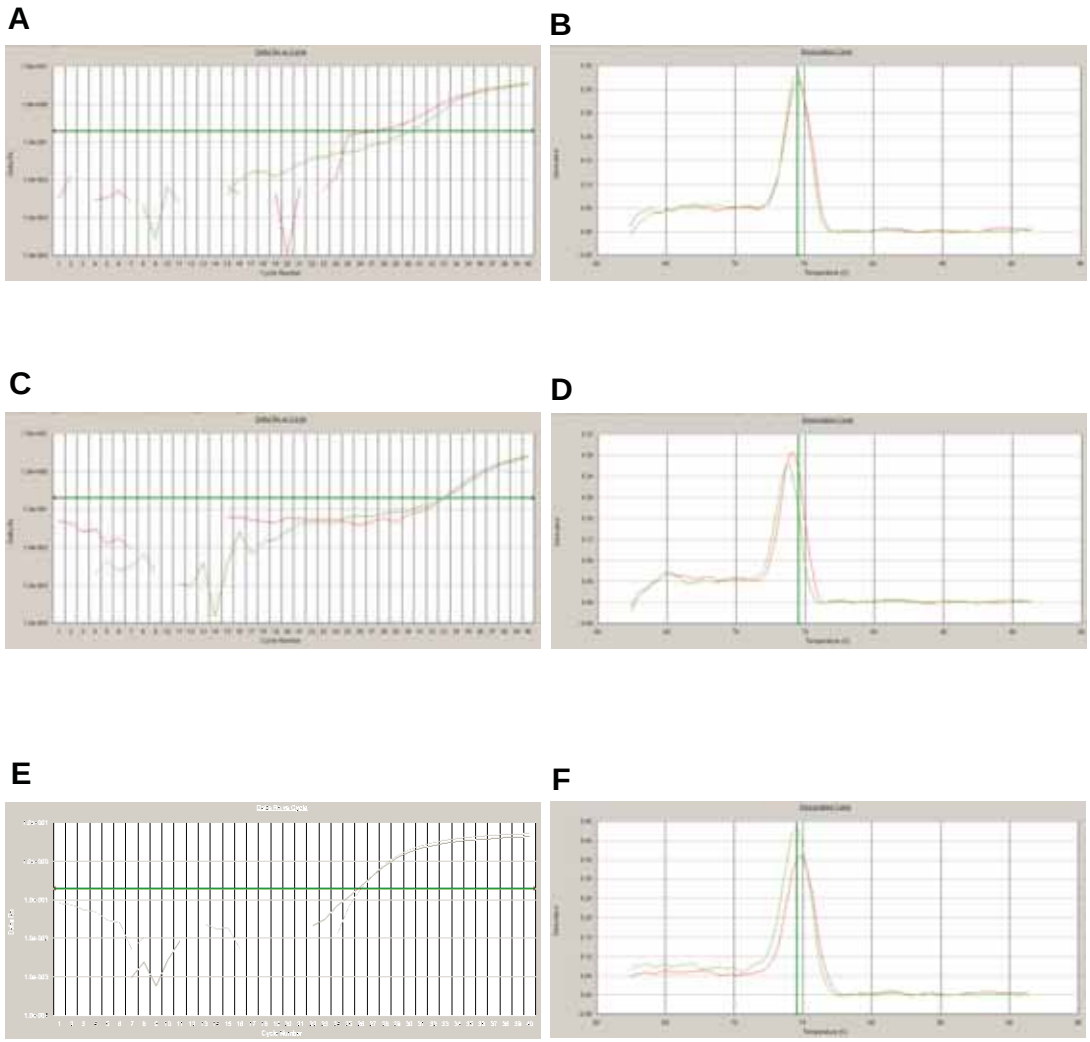


Figure 3.19 Cross-reactivity test of OPN-C primer pair. The primers OPN-C-F and OPN-AC-R were tested against 0.1 pg of OPN-A **(A, B)**, OPN-B **(C, D)** and OPN-C **(E, F)** to determine the specificity of the primers. This primer pair amplified OPN product from all OPN variant templates, indicating that it was not specific for OPN-C and could therefore not be used to individually quantify OPN-C.

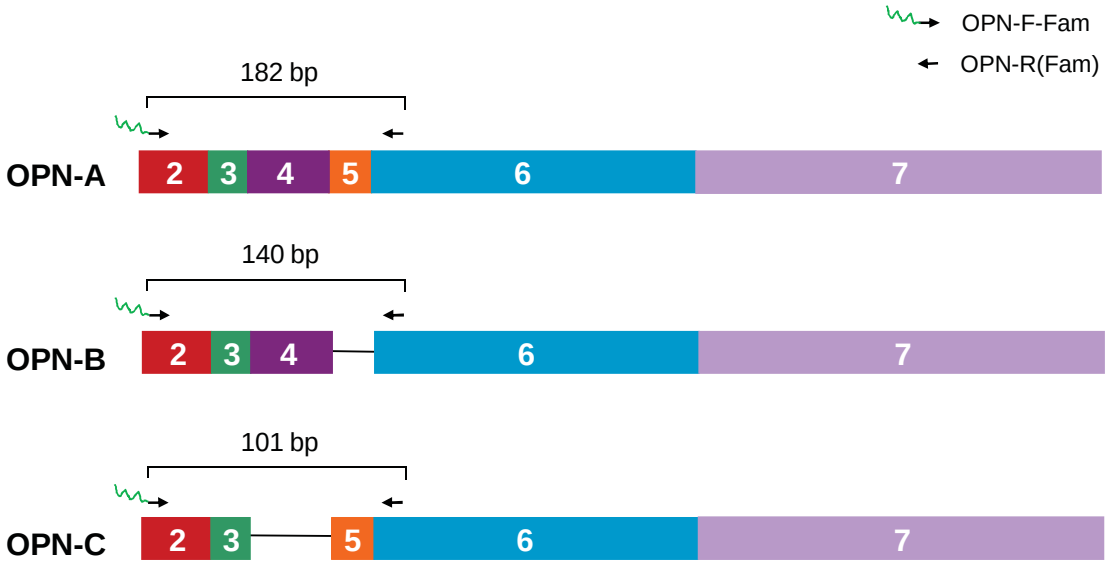


Figure 3.20 Location of primers for concurrent amplification of the three OPN splice variants by GeneScan® Genotyping PCR. Primers are located within exons 2 and 6 to flank the missing exons 5 and 4 of OPN-B and OPN-C respectively. Expected amplicon size for each variant is noted above. The forward primer has a 5' 6-FAM tag attached to allow for fluorescent identification of resultant DNA fragments.

number (15, 20, 25, and 30 cycles). Products were electrophoresed on 1% agarose gels and visualised with GelRed™ Nucleic Acid Gel Stain (Figure 3.21). Each OPN variant was successfully and specifically amplified with the optimal input amount and cycle number combination being 10 ng template and 15 cycles respectively. This resulted in non-saturating levels of amplicons.

While the primer set could amplify each variant individually, to be of use they must be able to amplify all OPN variants simultaneously when all are present as they are in a clinical sample. To mimic this situation, validation was performed by using the conditions outlined above with varying ratios of OPN cDNA template. This information would confirm if the primers had equal or competing affinities for the three variants. cDNA mixtures were prepared containing either equal (10 ng each of OPN-A, OPN-B and OPN-C; 10:10:10) or differing (10 ng OPN-A and 5 ng each of OPN-B and OPN-C; 10:5:5 or 10 ng OPN-A, 5 ng OPN-B and 1 ng OPN-C; 10:5:1) amounts of the OPN variants. Empty pRc/CMV vector was added to a number of samples and also used on its own to confirm that amplification was occurring from the OPN sequences inserted into pRc/CMV rather than the vector itself. The primer set was able to successfully amplify the three variants at visually correct band intensity differences from an individual sample when the amount of each variant differed (10:5:5 or 10:5:1), but not when similar input amounts were used (10:10:10; Figure 3.22). The presence or absence of empty pRc/CMV did not affect amplification, and no amplification was visualised in the pRc/CMV control as expected.

PCR products from the sample OPN mix 10:5:5 were diluted 3-fold and sent to the IMVS Molecular Pathology laboratory for GeneScan® analysis. Using capillary based fluorescent recognition, fragments representing the three OPN variants were detected in all samples and fluorescent peaks obtained (Figure 3.23). Fragments corresponding to amplified OPN-A

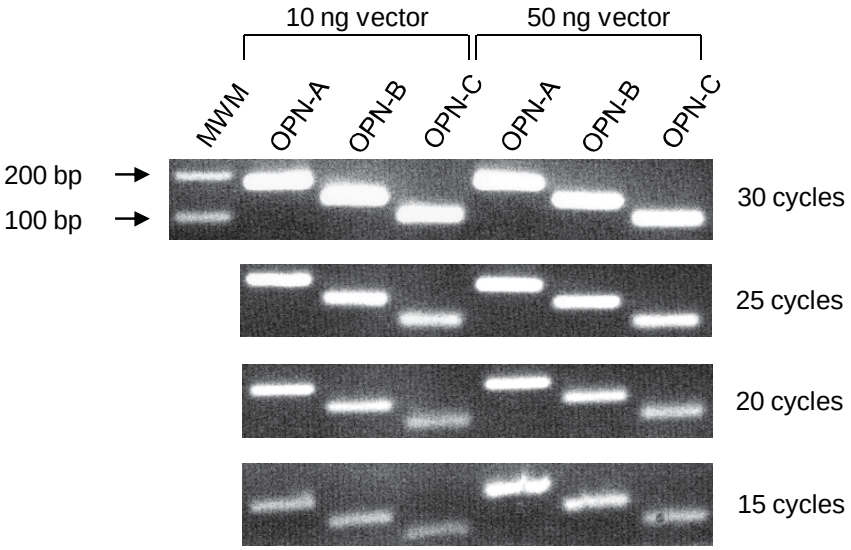


Figure 3.21 Optimisation of GeneScan® Genotyping PCR method for fluorescent amplification of the three OPN variants. OPN variant templates were used to assess specificity of amplification from the designed primers. 10 ng and 50 ng samples were prepared individually for each template and subjected to 15, 20, 25 and 30 cycles of standard OPN PCR to determine the optimal input amount and cycle number and to test primer specificity. All gels are representative of multiple experiments. The optimal input amount and cycle number was determined to be 10 ng template and 15 cycles.

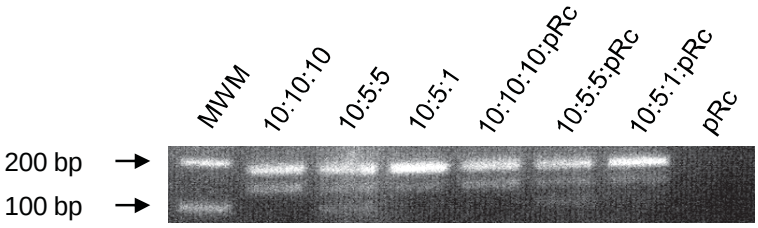


Figure 3.22 Validation of GeneScan PCR method for fluorescent amplification of the three OPN variants. pRc/CMV vectors containing sequences for each variant were used as templates to assess affinity of amplification from the designed primers. Mixtures were prepared containing the three variant templates in the following combinations: 10 ng each of OPN-A, OPN-B and OPN-C (10:10:10), 10 ng of OPN-A plus 5 ng each of OPN-B and OPN-C (10:5:5) and 10 ng of OPN-A, 5 ng of OPN-B and 1 ng of OPN-C (10:5:1). Empty pRc/CMV vector was added to some samples to control for non-specific amplification (10:10:10:pRc, 10:5:5:pRc, 10:5:1:pRc). Empty pRc/CMV was also used on its own as a negative control (pRc). Gel is representative of multiple experiments.

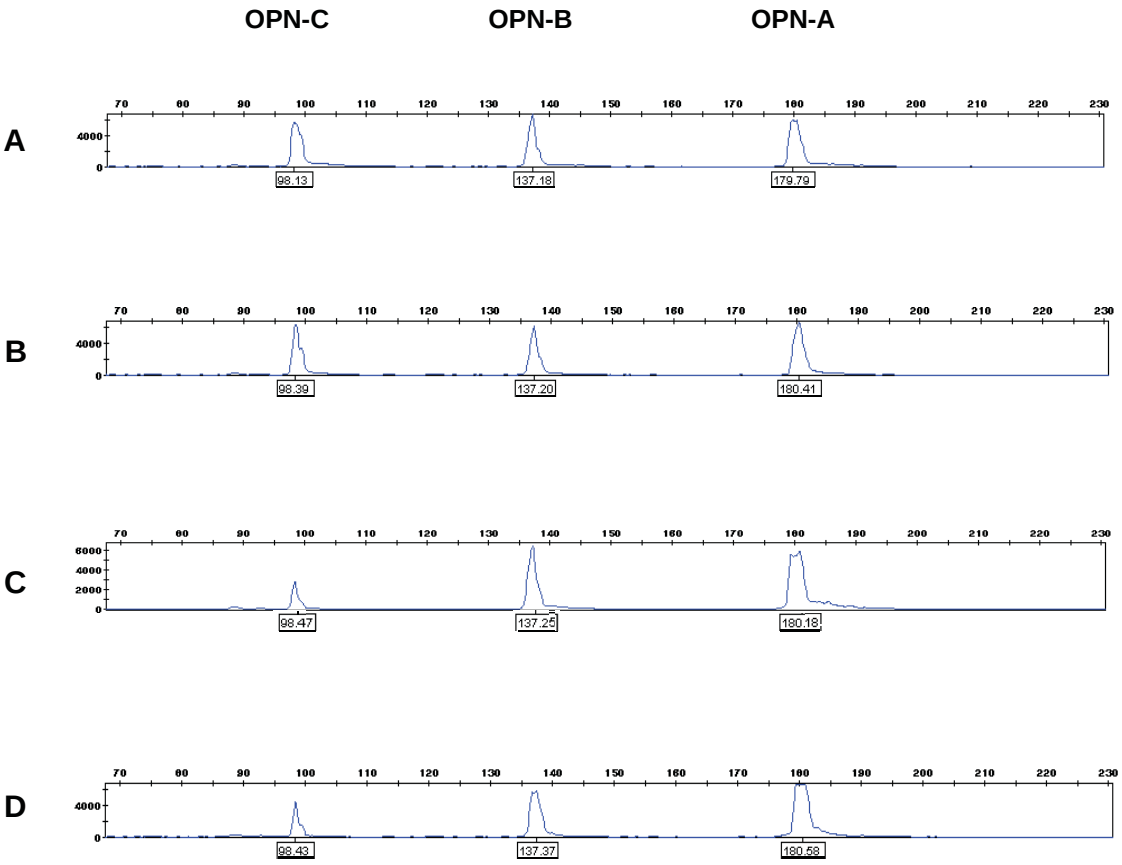


Figure 3.23 ABI Prism GeneScan PCR output readings of fluorescently-labelled PCR products derived from OPN variant templates. In all graphs, the 180 bp peak represents OPN-A, the 137 bp peak represent OPN-B and the 98 bp peak represents OPN-C. **(A)** GeneScan peaks obtained from OPN-A:OPN-B:OPN-C input vector ratio of 10:5:5 ng. **(B)** GeneScan peaks obtained from identical sample to **(A)** with the addition of 100 ng pRc/CMV (10:5:5:pRc). **(C)** GeneScan peaks obtained from OPN-A:OPN-B:OPN-C ratio of 10:5:1 ng input vector. **(D)** GeneScan peaks obtained from sample identical to **(C)** with the addition of 100 ng pRc/CMV (10:5:1:pRc).

(180 bp), OPN-B (137 bp) and OPN-C (98 bp) were obtained in all samples as expected. Peak heights and areas under the peak were calculated by the GeneMapper ID v3.2 software, and ratios between transcript variants calculated to determine if the ratios of peak area or height corresponded with known ratios of input vector (Table 3.4). Neither peak height nor area ratios correlated with ratios of known input vector amount, thus this method was deemed unsuitable for individual quantification of the three OPN splice variants within a single sample.

3.3 Discussion

3.3.1 Expression of OPN and its variants in HCC-derived cell lines and HCV-infected and HCC tissues

Increased OPN expression in patients with HCC is well described in the literature (Gotoh et al., 2002; Pan et al., 2003; Kim, J. et al., 2006). In this chapter we investigated OPN mRNA expression in a range of cultured cell lines derived from liver tumors and other tumors, which confirmed previous reports showing positive OPN mRNA expression in liver, breast and prostate cancer cell lines (Gao et al., 2003; Sahai et al., 2004; Said et al., 2005; Forootan et al., 2006; Shevde et al., 2006; Zhao, J., Lu et al., 2008). In contrast to hepatocellular carcinoma, little is known regarding OPN expression in liver disease prior to HCC development. Therefore, we also investigated OPN mRNA expression in a range of clinical samples including normal liver, HCV-infected liver at varying stages of disease, and HCC tumor tissue and its surrounding cirrhotic liver. Consistent with previous reports (Brown et al., 1992) OPN was expressed in the normal liver, most likely in bile duct epithelial cells (see Figure 6.4). In HCC we also noted significant expression of OPN mRNA over cognate cirrhotic liver in 50% of patients and, while our sample set is small, this is consistent with the literature (Gotoh et al., 2002; Pan et al., 2003). However, our most interesting finding was the observation that OPN mRNA expression increased as liver disease progressed in the HCV-infected liver. Our small sample set precluded

Variant	Input ng	Input Ratio	Fragment Size	Peak Height	Height Ratio	Peak Area	Area Ratio
OPN-A	10	1	179.79	6067	1	112622	1
OPN-B	5	0.5	137.18	6825	1.12	82575	0.73
OPN-C	5	0.5	98.13	5694	0.94	83846	0.74
OPN-A	10	1	180.18	5808	1	123679	1
OPN-B	5	0.5	137.25	6457	1.11	81183	0.66
OPN-C	1	0.1	98.47	2869	0.49	25037	0.20
OPN-A	10	1	180.41	6743	1	106466	1
OPN-B	5	0.5	137.20	6235	0.92	74819	0.70
OPN-C	5	0.5	98.39	6430	0.95	73735	0.69
pRc/CMV	100	10	n.d.	n.d.	n/a	n.d.	n/a
OPN-A	10	1	180.58	6744	1	140633	1
OPN-B	5	0.5	137.37	5899	0.87	60942	0.43
OPN-C	1	0.1	98.43	4549	0.67	36245	0.26
pRc/CMV	100	10	n.d.	n.d.	n/a	n.d.	n/a

Table 3.4 Peak height and area values and ratios obtained for samples containing differing input amounts of each of the three OPN variant vectors. Known ratios of OPN-B and OPN-C vector amounts compared to OPN-A were calculated (**Input Ratio**) and compared with the same ratios obtained for peak height (**Height Ratio**) and peak area (**Area Ratio**). The empty control vector pRc/CMV was not detected (n.d.) during GeneScan® Genotyping PCR as expected.

us from reaching a significant correlation, however we have suggested for the first time that OPN mRNA expression in the HCV-infected liver may be correlated to disease stage. Others have shown correlation between serum OPN expression and liver disease progression in HBV and fulminant hepatitis (Matsui et al., 2004; Zhao, L., Li et al., 2008), so it is not inconceivable that a similar correlation between OPN mRNA and liver disease could be found for HCV.

The majority of research into OPN expression and function has solely focused on the full length protein (OPN-A) and for the most part neglected the alternatively spliced variants OPN-B and OPN-C. Work performed in this chapter and previous work performed by the principal supervisor's laboratory has confirmed these findings for OPN-A and we have now extended investigations to characterise the alternatively spliced variants of OPN. Studying the expression of these alternative variants may lead to the discovery of differential functions for each, which may in turn help solve the mystery of OPN's role in liver disease and cancer development. Particularly important may be the effect of the missing exons in these variants (exon 5 for OPN-B and exon 4 for OPN-C), although later in this study we show that the lack of these exons does not affect the ability of OPN to bind to its receptor CD44 (see Chapter 4). Using semi-quantitative PCR, we visualised expression of all variants in non-diseased liver, HCV-infected liver and HCC tumor tissue, with all tissues expressing predominantly OPN-A regardless of disease stage. Expression of the three variants in HCV-infected liver prior to the development of HCC suggests a role for each in virally-induced liver disease pathogenesis. At the time of this study, the expression of OPN-B and –C in healthy liver had not been described and contradicted a previous suggestion that OPN splicing occurs only in cancerous tissue (He et al., 2006). However, a subsequent study by the same group showed OPN-B mRNA expression in normal human breast tissue (Mirza et al., 2008) and a recent study has shown normal liver to express predominantly OPN-C (Chae et al., 2009) which contradicts our result of predominantly OPN-A with very little

OPN-C expression. This same study showed that similar to our results HCC tissue expressed predominantly OPN-A and OPN-B, and suggested that OPN-A and –B are involved in tumorigenicity and cell migration and that OPN-C may suppress migration. A similar study in lung cancer suggested that OPN-A expression increased tubule length and VEGF secretion, whereas OPN-C abrogated both of these functions (Blasberg et al., 2009). OPN-C has also been associated with metastatic potential in pancreatic cancer and HCC (Takafuji et al., 2007; Sullivan et al., 2009). These studies strongly indicate divergent functions for OPN-A and OPN-C, and we further investigate the functions of the three variants in Chapter 4 of this thesis.

At the time of this study, expression of OPN-B and –C had only been noted the invasive breast carcinoma lines MDA-MB-231 and MDA-MB-435 (He et al., 2006), however very recent studies have also shown expression in the pancreatic ductal adenocarcinoma line BxPC-3 (OPN-C only) (Sullivan et al., 2009) and a range of HCC-derived cell lines (Chae et al., 2009). Our results showed expression of OPN variants in cell lines derived from hepatocellular (Huh-7, HepG2, Hep3B, PLC/PRF/5), breast (MDA-MB-231) and cervical (HeLa) cancers. In our hands, the majority of human cell lines showed a similar expression pattern of predominantly OPN-A, followed by a lesser amount of OPN-B and a very small amount of OPN-C. The majority of these cell lines are know to possess low metastatic potential (Ye et al., 2003), and given the association of OPN-C with metastasis described above it is not surprising that these cell lines showed low OPN-C expression. Our results differ significantly to those obtained by Chae *et al.,* (2009); for example, they showed predominantly OPN-C expression in Hep3B whereas we showed predominantly OPN-A with very little OPN-C. This demonstrates the inter-laboratory variation that commonly occurs from use of the same cell line that has been handled differently and outlines why it is important to validate even the most well described methods for use under your own conditions. Nevertheless, our results demonstrate expression of all three variants in HCC.

We also noted a complete lack of expression of OPN-B and –C in cell lines of any species other than human, suggesting for the first time that OPN splicing may be specific only to humans. The reasons for this remain unknown and are beyond the scope of this study, but suggest that OPN-B and –C may be involved in complex processes found only in highly sophisticated species.

3.3.2 Quantitation of OPN variants

Quantification of absolute or relative expression of individual splice variants may aid in the diagnosis and prognosis of certain disease states. For example, higher expression of the alternative variant of the SMN gene (SMN2, lacking one exon compared to the full length gene SMN1) has prognostic significance for spinal muscular atrophy (Feldkotter et al., 2002). Given that divergent functions have been suggested for OPN-A and OPN-C (as discussed above), the quantification of individual OPN variants in a clinical sample, for example a liver biopsy from a HCV-infected patient with early-stage HCV-related liver disease, may be indicative of the aggressiveness of disease progression that the patient may endure in the future and may also help predict risk of developing HCC. Alternatively, quantification of OPN variants in HCC may help to predict survival, metastatic potential and treatment response.

Attempts have been made to quantify alternatively spliced variants using various techniques, including northern blotting, ribonuclease protection assay, semi-quantitative RT-PCR, competitive RT-PCR and microarray hybridisation, however each method possesses its own limitations. Recent efforts have focused on real-time quantitative PCR, with success observed in quantifying individual splice variants of nuclear factor 1, mouse glial cell-line derived neurotrophic factor family receptor alpha 2, and T cell immune regulator gene 1 (Vandenbroucke et al., 2001; Wong, Y. W. et al., 2002; Shulzhenko et al., 2003). A recent study was able to individually amplify and quantitate the three OPN splice variants in human breast cancer (Mirza et al., 2008). This study compared OPN variant expression in breast cancer tissue, cognate

non-diseased breast tissue and normal breast and showed a significant increase in OPN-C in breast cancer compared to normal breast. However, no mention is made of the specificity of each primer set to its corresponding variant or any validation methods used to confirm the validity of the protocol. Given our results (see below) and without stringent validation it is difficult to be confident in the quantification of OPN variants presented by Mirza and colleagues.

This chapter attempted to individually quantify the three OPN variants using a real-time RT-PCR method and individual primer sets for each variant. However, after two attempts we were unable to design primer pairs that could amplify their specific OPN variant without some degree of cross-reactivity to the other two variants. Cross-priming presented a constant issue for the OPN variants, as there are no unique sequences present in any variant as would be found if OPN splicing resulted in exon insertion. OPN splicing in fact results in exon deletion, so we attempted to use the unique exon-exon junctions as a source of unique sequence around which to design real-time PCR primers. However, each primer pair still contained sequences of up to 12 nucleotides that could be found in all three variants and as such presented common areas for primer annealing and subsequent sequence elongation. As such we determined that it would be extremely difficult to individually amplify the three OPN variants by real-time RT-PCR due to the cross-reactivity problems encountered in designing unique primer sets.

We also attempted to utilise GeneScan® Genotyping PCR as a tool for the individual amplification of the OPN variants, and whilst this method successfully amplified the three variants individually, the resultant fluorescence ratios generated by the GeneScan® analysis did not correlate with known input ratios. Therefore, within the time constraints of this thesis we were unable to individually quantitate the three OPN variants.

In conclusion, the results of this chapter suggest that (i) all OPN variants are expressed in the human liver from normal liver to varying stages of HCV-related liver disease, with the full-length variant OPN-A dominating, (ii) that OPN expression increases with degree of fibrosis in the HCV-infected liver, and (iii) due to the nature of the OPN variants, individual quantification of each variant is technically difficult. Our results also show expression of all three OPN variants in normal human liver and contradict previous reports that OPN splicing occurs only in cancerous tissues.

3.3.3 Future directions

This chapter suggests a trend for correlation between increased OPN mRNA expression and increased degree of liver disease in the HCV-infected liver. Increasing our sample set of HCV-infected patient biopsies may allow us to find a significant correlation. It would also be interesting to re-investigate OPN mRNA expression in a larger cohort of paired HCC tumor tissues as our results were highly variable. To achieve this, close collaboration with clinical hepatologists is required to obtain a sufficient number of biopsy samples.

The obvious next step for this study would be to establish a validated method for the individual amplification of the three OPN variants, as attempted in this chapter. We could attempt to alter the conditions of our real-time PCR method using the same primer pairs by adjusting the temperatures and cycle times. This may increase the specificity of the primers such that they do not cross-amplify the other variants. Failing this, it would be interesting to utilise the primers described in Mirza *et al.*, (2008) in our system to see if we could successfully amplify the three OPN variants individually according to our stringent validation methods. These primers have been optimised for use with RNA from human breast cancer tissue samples but should amplify similar sequences in human liver samples and HCC-derived cell lines.