CHAPTER 2  REVIEW OF LITERATURE

2.1 Hepatitis:
Hepatitis is a generalized term to characterize the inflammation of liver. Several agents like viruses, parasites, chemicals are involved in causing the liver-inflammation (hepatitis). Autoimmune processes and metabolic dysfunction are also associated with hepatitis. Hepatitis caused by distinct viruses is now known as viral hepatitis and known to present the distinct pathology of liver. Yellow discoloration of the skin and conjunctivae, known as jaundice are the most common features of hepatitis of any origin.

Till date following seven viruses are known causative agents of viral hepatitis. Hepatitis A virus (HAV), hepatitis B virus (HBV), hepatitis C virus (HCV), hepatitis D virus (HDV), hepatitis E virus (HEV), hepatitis G virus (HGV) and TT virus (TTV). Among them HAV, HBV, HCV, HDV and HEV cause the actual clinical problems. All these viruses except for HAV and HEV cause chronic hepatitis and HBV and HCV are the major causative viruses of chronic liver disease (Lino and Hino, 1999).

2.2 Hepatitis B virus infection:
2.2.1 Epidemiology and public health burden:
Approximately one third of the world’s population has serological evidence of past or present infection with HBV and 350 million people are chronically infected. The spectrum of disease and natural history of chronic HBV infection is diverse and variable, ranging from a low viremic inactive carrier state to progressive chronic hepatitis, which may evolve to cirrhosis and hepatocellular carcinoma (HCC). HBV-related end stage liver disease or HCC are responsible for over 1 million deaths per year and currently represent 5–10% of cases of liver transplantation [Ganem and Prince 2004, Hoofnagle et al 2007, Liaw 2005 and Lok and McMahon 2007 ]. Host and viral factors, as well as coinfection with other viruses, in particular hepatitis C virus (HCV), hepatitis D virus (HDV), or human immunodeficiency virus (HIV) together with other co-morbidities including alcohol abuse and overweight, can affect the natural course of HBV infection as well as the efficacy of antiviral strategies.
CHB may present either as hepatitis B e antigen (HBeAg)-positive or HBeAg-negative CHB. HBeAg-positive CHB is due to so-called “wild type” HBV. It typically represents the early phase of chronic HBV infection. HBeAg-negative CHB is due to replication of naturally occurring HBV variants with nucleotide substitutions in the precore and/or basic core promoter regions of the genome and represents a later phase of chronic HBV infection. The prevalence of the HBeAg negative form of the disease has been increasing over the last decade as a result of HBV-infected population aging and represents the majority of cases in many areas, including Europe [Rizzetto 1998, Rizzetto and Ciancio 2008 and Zarski et al, 2006].

Morbidity and mortality in CHB are linked to persistence of viral replication and evolution to cirrhosis or HCC. Longitudinal studies of patients with CHB indicate that, after diagnosis, the 5-year cumulative incidence of developing cirrhosis ranges from 8 to 20%. The 5-year cumulative incidence of hepatic decompensation survival being approximately 80–86% in patients with compensated cirrhosis [Liaw 2005, Chu and Liaw 2006, Fattovich 2003, Fattovich et al, 2008 and 2004]. Patients with decompensated cirrhosis have a poor prognosis with a 14–35% probability of survival at 5 years.

The worldwide incidence of HCC has increased, mostly due to HBV and HCV infections; presently it constitutes the fifth most common cancer, representing around 5% of all cancers. The annual incidence of HBV-related HCC in patients with CHB is high, ranging from 2% to 5% when cirrhosis is established [Fattovich et al, 2004]. However, the incidence of HBV related HCC appears to vary geographically and correlates with the underlying stage of liver disease (Fig 2.1).
2.2.2 Hepatitis B virus (HBV):

Hepatitis B virus is the first member to be discovered of a family of viruses, later designated *Hepadnaviridae*. This family has since been divided into two groups, the ortho-hepadnaviruses and avian hepadnaviruses. These are hepatotropic and partially double-stranded DNA viruses. Their replication strategy is unique for animal DNA viruses and is only shared by cauliflower mosaic virus (also a DNA virus), in that they use an RNA intermediate and reverse transcription step for replication (Seeger and Maragos, 1991). The circular genome is very compact, with four partially overlapping open reading frames (ORFs). There are no non-coding regions in the genome, so that all regulatory signals are also part of protein-encoding sequences; HBV can encode approximately 50% more protein than would be expected from its genome length (Ganem & Varmus, 1987). In terms of HBV evolution, this leads to two opposing tendencies: the use of reverse transcriptase with its lack of proofreading tends to maintain a relatively high mutational rate, whereas the extreme compactness of the genome prevents a large degree of genetic variability from occurring.

2.2.3 Historical perspective:

When Blumberg et al (1965) first reported their findings on the Australia (Au) antigen; later named as the surface antigen (HBsAg) of hepatitis B virus (HBV), they documented some epidemiological results also that continue to be valid today. Its higher prevalence was found in normal populations from Taiwan, the eastern Mediterranean region, some Polynesian islands and Australian aborigines, than in North America. Prince et al, (1968) and Okaochi et al, (1968) reported the presence of ‘Au’ antigen found specifically in the serum of hepatitis B patients.

By establishing a correlation between acute or chronic hepatitis and the presence of Australia antigen in serum, the agent of serum or inoculation hepatitis was found. When serum from Australia antigen-positive
patients was studied by electron microscopy, vast numbers of spheres and filaments of 22 nm in diameter were seen. Additionally, larger particles of 42 nm having a central nucleocapsid and an outer coat were also seen, subsequently known as Dane particles (Dane et al., 1970). Dane particles were shown to constitute the complete virion. Hepatitis B vaccine studies using plasma-derived, highly purified HBsAg particles started in 1975 when it became clear that anti-HBs was the neutralizing antibody for infectious hepatitis B virus. Since then, safety and protective efficacy of these preparations have been established.

2.2.4 Hepatitis B virus structure

HBV remains present in the form of spherical particles approximately 42-47nm in diameter. It is composed of an electron-dense spherical inner core of 22-25 nm diameters and an envelope of approximately 7-nm thickness. Envelope is made up of lipid and surface (S) antigen (HBsAg), against which antibody is directed to neutralize the virus. The spherical 22-nm-diameter inner-core particles consist of the viral core antigen (HBcAg), the viral e antigen (HBeAg), the viral DNA and DNA polymerase (Fig 2.2).

![Figure 2.2: Structural arrangement of Hepatitis B virus with respect to the viral proteins produced.](image)

2.2.5 Viral genome:

The smallest genome is present in hepadnaviruses among all the known viruses. It is small circular molecule and is partially double stranded (Landers et al., 1977). Hepatitis B virus DNA consists of a long negative (-) strand. The sequence of the (-) strand is complementary to that of the viral messenger RNA (mRNA) of 3.2 k. The short positive (+) strand of varying length between 1700 and 2800 bases also remains present. Negative (-) strand is not a closed circle but a nick exists at a unique site approximately 225 bp from the 5’end of the positive strand (Sattler et al., 1979). A 19-nucleotide capped oligoribonucleotide remains covalently attached to the 5’ end of the positive DNA strand of HBV. This RNA functions as a primer for synthesis of this DNA strand (Fig 2.3).

The genome has four overlapping open reading frames in the negative DNA strand. The entire genome gets transcribed and eventually overlapping genes are translated in different reading frames.
Figure 2.3: Structural and genetic arrangement of Hepatitis B virus genome.

**Surface-gene (S-gene):** Includes pre-S1, pre-S2 and S regions having three in-frame initiation codons. Proteins of the virion envelope and that of the incomplete viral forms (surface antigen) are encoded by these surface genes. Glycosylated and non-glycosylated forms of these three proteins are found which are of 24 kd, 33 kd and 39 kd encoded by 'S' alone, by pre-S2 and S, and by pre-S1, pre-S2, and S-regions of the S-ORF, respectively.

**Core-gene (C –gene):** Core-gene encodes for viral core i.e. nucleocapsid polypeptides, which is of 21 kd as well as of 16 kd. The later polypeptide haves the HBeAg specificity. This ORF includes a short pre-C (pre-core) sequence at its N-terminus delineated by an in-frame initiation codon.

**Polymerase-gene (P-gene):** It spans nearly three-fourth of the viral genome and overlaps the C-terminal portion of the C-gene, the entire S-gene, and the amino-terminal portion of the X-gene. P-gene is responsible for the production of polypeptide of nearly 90 kd. This protein contains the virion associated DNA polymerase ('or' reverse transcriptase) activity (Bavand et al, 1989).

**X-gene:**

X-gene encodes for the X-protein of 154 amino acids in length. Transcription of certain viral and cellular genes is activated by the X-protein (Twu et al, 1987). Duck Hepatitis B virus (DHBV) does not have the X-gene. Avian hepadnaviruses contain only three genes (S, C and P).
2.2.6 Viral proteins:

2.2.6.1 Envelope proteins (Surface Antigen):

Both glycosylated and non-glycosylated forms of above described three surface proteins are found. The two smallest polypeptides (24kd & 27kd) encoded by the S-region remain present in greatest amount in HBsAg preparations, which consists of 226 amino acids. Nonglycosylated form is designated as p24 and its glycosylated form as gp27. Two sets of larger polypeptides are minor components of HBsAg preparations. They are (i) 33kd (280-amino acids) polypeptide consisting of 55 amino acids encoded by pre-S2 and 226 amino acids encoded by S region of the S-ORF, which is designated as p33. The glycosylated form of the same polypeptides (36kd) is designated as gp36. (ii) 39kd (389-400 amino acids) polypeptide consisting of 119 amino acids (108-119 in different viruses) encoded by pre-S1 plus the sequences encoded by the pre-S2 and S-regions of the S-ORF, designated as p39 and its glycosylated form as gp42.

2.2.6.2 Nucleocapsid (core) proteins:

Hepatitis B core antigen (HBcAg) contributes to the formation of virion-core. It remains present only as the internal component of virions. No free HBcAg is found in the serum. HBcAg and HBeAg have different specificities of proteins encoded by the viral C-gene. This is the single predominant polypeptide (p21) of 183-185 amino acids in length, which is found in highly purified virion cores. The carboxy terminal of 34 amino acids of p22 is rich in arginine, therefore, highly basic and is involved in binding with viral DNA (Petit et al, 1985). HBcAg specificity for antibody binding is dependent upon protein conformation of intact HBcAg particles which is lost if disruption of the particles takes place.

The polypeptide having HBeAg specificity is secreted in serum. Its approximate weight is 16kd and is a truncated form of p21, where 34 amino acids of the carboxy-terminal (the basic DNA-binding domain) are missing. The pre-core polypeptide has an N-terminal signal sequence that directs this protein to the secretory pathway. The carboxy terminal DNA-binding domain is removed from this polypeptide prior to the secretion from cell.

2.2.6.3 Viral Polymerase:

Polymerase gene of hepatitis B virus has sequence homology to other viral reverse transcriptases also. The protein encoded by P-gene consists of multiple functional domains: the N-terminal portion serves as the primer for reverse transcriptase, the spacer region has not fully been characterized, the reverse transcriptase domain and the ribonuclease H domain (Hussain and Lok 1999). The reverse transcriptase domain has been further divided into 5 sub-domains designated A through E. A three dimensional HBV polymerase structure has also been described as a model of right handed conformation similar to other viral reverse transcriptases resembling with thumb,
palm, and finger (Allen et al, 1998). Domains A, C, and D are involved in dNTP binding and in catalysis, which correspond structurally to the finger domain. Domains B and E interact with RNA template and with primer and have been designated structurally as the thumb and palm domains (Gao et al, 2000). The C domain of HBV reverse transcriptase has a YMDD motif, which is highly conserved in other viral polymerases also. In persons chronically infected with HBV or HIV, substitutions either with valine or with isoleucine for amino acid position 552 (methionine), in the YMDD motif has been associated with drug resistance. Crystallographic studies suggested that YMDD mutation alters the dNTP binding pocket. Other mutations resulting in to drug resistance are also known within HBV reverse transcriptase. They occur outside the dNTP binding site. One of them is L528M substitution in the B domain and emerges in association with the YM552I/VDD mutation. Alone, the L528M mutation has a mild effect on HBV replication compared with the single M552V mutation, however, the replication efficiency has been reported to be improved when combined with M552V mutations (Melegari et al, 1998).

2.2.6.4 The X-gene product:

The X-ORF is located downstream to enhancer I and is partly overlapped by P ORF at its N-terminus and by the C-ORF at its C-terminus. It encodes for a 17 kd protein designated as pX (X-protein). The X-gene is conserved among mammalian hepadnaviruses. The X-protein, pX, is essential for viral propagation at least in the woodchuck. Since last decade, efforts have been made to study the oncogenic role of pX in hepatitis B virus. Now it is known that pX is a multi functional regulatory protein which regulates many host functions. pX modulates the transcription machinery and/or protein kinase signaling cascades, transactivates many host genes involved in cell proliferation, cytokine networks, and acute immune response. It also modulates DNA repair process by interacting with p53.

2.3 Hepatitis B virus replication:

Process involved in entry of the HBV to the hepatocytes is still not well understood. The critical cellular factors, in particular the viral receptor, that control entry have not yet been identified.

Formation of covalently closed circular (ccc) DNA: Viral replication cycle starts with the formation of covalently circular (ccc) DNA from the incoming viral DNA to the hepatocytes. Viral DNA is a relaxed circular (rc) molecule. It has a complete minus strand and an incomplete plus strand. Events involved in the conversion of rc to ccc DNA have not fully been characterized. However, structural differences between the two DNA species signify certain steps. First of all, viral reverse transcriptase (RT) is removed. Then one copy of the terminally redundant sequences is deleted from the 5' end of the minus strand. Thereafter, modification on plus strand DNA involves its completion. An RNA oligomer from the 5' end of the plus strand is removed. Furthermore, the 5' and 3' ends of both the strands ligate to form a covalently closed circular DNA. Cellular DNA repair enzymes are believed to be involved in the formation of ccc DNA.
However, on the other hand viral DNA polymerase is also speculated to be involved in plus strand elongation up to the primer-binding site at DR2.

**Transcription of viral RNA:** Once formed, ccc DNA acts as the template for the transcription of viral RNA from several distinct promoters (Fig. 2.3). Recently, Newbold and colleagues (1995) demonstrated the association of ccc DNA with proteins. Thus, ccc DNA exists as a mini-chromosome. The chromatin like structures of ccc DNA provides stability to it within the nucleus of the infected cell. Stability and the half-life of ccc DNA is becoming an important issue for antiviral therapy that is aimed to cure for hepatitis B virus infection. Experiments conducted in the presence of inhibitors of hepadnaviral DNA synthesis have shown that ccc DNA has longer than 4 days half-life *in vivo*.

Core subunit, reverse transcriptase (RT) and pregenomic RNA (pgRNA) are the major three viral components for the replication of the hepadnaviruses genome. PgRNA is roughly 10% larger than ccc DNA because it is terminally redundant. Core and RT are translated from pgRNA, which subsequently serves as the template for reverse transcription for minus strand (Huang *et al*, 1991).

**Assembly of core proteins:** Core proteins assemble into icosohedral capsids. PgRNA and RT are encapsidated in these structures. The packaging reaction depends on the presence of the packaging signal on pgRNA, termed epsilon, which has been proposed to fold into an RNA hairpin with a loop and a bulge (Junker-Niepmann *et al*, 1990). RNA packaging and the priming of reverse transcription both depend on the interaction between the polymerase and epsilon RNA. The binding between the RT and the epsilon leads to stable ribonucleoprotein complex (RPN), which, apart from its role in packaging, also serves as the primer for the reverse transcription.

**Synthesis of minus strand DNA:** Unlike all other known reverse transcriptases that use RNA as a primer for minus strand DNA synthesis, the hepadnavirus polymerase employs the hydroxyl residue of one of its own tyrosine residues to initiate DNA synthesis. As a result of this strategy, the 5’ end of minus strand DNA present in virion are covalently attached to protein.

**Translocation of the nascent minus strand:** The nascent minus strand so formed is translocated to the 3’ end of the pgRNA template, where it anneals to complementary sequences and continues reverse transcription.

**RNaseH activity:** Then the 5’ end of the RNA template, which is simultaneously degraded, as it is reverse transcribed, by an RNAaseH activity on the polymerase polypeptide. At the end of this step, the synthesis of the minus strand of the genomic DNA is complete, and a short portion of the template RNA remains undegraded.

**Translocation of the RNA primer:** The undegraded RNA is then translocated to the DR II close to the 5’ end of the minus strand of newly synthesized DNA, and anneals to it via a short 11-12 nucleotide long
homology between DR I and DR II, where it primes the synthesis of the plus strand DNA, which proceeds from the DR II to the 5’ end of the minus strand template.

Circularization of the minus strand: For continuation of synthesis there must be a switch to the other end of the template, which is apparently achieved by circularization of the template. This step is facilitated by the terminal redundancies in the minus strand DNA.

The cause and significance of the premature termination of plus strand synthesis in mammalian viruses is unclear. It is believed that structural factors of the capsid and the RT prevent the completion of this reaction during virion morphogenesis.

Transcription: Four promoter and two enhancer elements regulate the expression of viral genes. These transcriptional regulatory elements direct the synthesis of multiple viral transcripts that are approximately 3.5, 2.4, 2.1 and 0.9 kilobases in length, which contain the C- and P-ORFs, large S-ORF, middle and small S-ORFs, and X-ORF, respectively. These transcriptional regulatory elements contain binding sites for liver-enriched transcription factors, which provide tissue specificity to the hepadnaviruses (Di et al, 1997). After the completion of transcription, transcripts are modified by the addition of 5’-cap structure, and then 3’ polyadenylation takes place. All of these four transcripts co-terminate at an identical polyadenylation site (Schaller and Fischer, 1991) (Fig. 2.3). Once in the cytoplasm, each of the four transcripts serves as the mRNA. In general, the same interaction with the translation machinery operates for viral protein synthesis as it works in case of cellular mRNA (Fig 2.4).

The final step is virion morphogenesis (Nassal et al, 1996), which involves budding of the newly formed cytoplasmic nucleocapsids into a pre-Golgi compartment of the cell's secretory pathway (Roingeard and Sureau, 1998), i.e. the endoplasmic reticulum (ER) or the intermediate compartment. The synthesized envelope-proteins initially get inserted into ER membrane. Most of the S proteins are released from the cell as particles, which bud into the lumen of these compartments and are exported by vesicular transport machinery. Most of the L protein is retained. Hence a certain concentration of L, M and S proteins is available at these membranes for interaction with the nucleocapsids. Mutational studies have indicated that a specific sequence within the preS1 region is required for capsid envelopment (Bruss et al, 1997).
2.4 Natural history of hepatitis B virus infection:

Hepatitis B virus infection may be associated with a very large spectrum of clinical presentations that may vary with age and the immunological status.

2.4.1 Acute hepatitis B:

It is usually a benign disease that spontaneously resolves in more than 90% of cases. It may also evolve into chronic hepatitis in 8-10%. Rarely, 0.1-1% of the patients may develop a fulminant course (Decker, 1998). When clinical symptoms develop, the serum is positive for the ‘surface’ and ‘e’ antigens (HBsAg and HBeAg, respectively), there are high levels of IgM antibodies to viral core antigen (IgM anti-HBc), and the HBV-DNA is usually detectable by direct hybridization techniques (Decker, 1998). When the virological/clinical course is self-limited, viremia quickly becomes undetectable; HBeAg disappears within a few weeks and is replaced by the corresponding antibody (anti-HBe). HBsAg positivity is highly variable and usually becomes negative in 2-4 months (Decker, 1998). The appearance of the antibodies to HBsAg (anti-HBs) is the best serological indicator of recovery from the infection, and may take several months after HBsAg seroclearance. Persistence of high HBV-DNA levels and of HBeAg positivity predicts evolution of the infection (Whalley et al, 2001), conventionally diagnosed when HBsAg persists for more than 6 months from its first detection. A remarkable characteristic of fulminant acute hepatitis could be HBsAg and HBV DNA negativity due to massive hepatocellular necrosis (Liang et al, 1991; Pollicino et al, 1997); in such cases anti-HBc IgM is the most reliable assay for the correct diagnosis.
2.4.2 Chronic HBV infection:

Chronic HBV infection may show different patterns of HBV serology and DNA replication. Considering the viremia levels as an indicator of HBV replication, the following three categories can be distinguished:

(A) HBsAg positive patients with persistently active viral replication: The vast majority of HBeAg positive patients and a minority of anti-HBe positive subjects show constantly active HBV replication with levels of circulating viruses > $10^5$ copies/mL that can easily be detected by direct hybridization techniques. In these cases, Southern blot of HBV DNA from the liver reveals a high amount of HBV replicative intermediates (Raimondo et al, 2003).

(B) HBsAg positive individuals with permanent or temporary suppression of viral replication: The category of chronic HBsAg positive patients with long-lasting or intermittent inhibition of HBV replication is highly heterogeneous and may comprise subjects co-infected with hepatitis Delta (HDV) or hepatitis C virus (HCV), the so called inactive HBsAg carriers, and individuals showing episodes of virological and clinical reactivation occurring in the course of an apparently quiescent infection. Inactive HBsAg carriers or healthy HBsAg carriers constitute a large proportion of the HBV infected individuals worldwide. They have normal liver biochemistry and ultrasonography, and if a liver biopsy is performed the histology is normal or reveals minimal changes. These subjects share anti-HBe positive and persistently anti-HBc IgM negative status. Levels of viremia in these cases may vary between $10^3$ and $10^5$ copies/mL in most cases (Cacciola et al, 2000; Martinot-Peignoux et al, 2002). Several HBsAg carriers show a peculiar virological and clinical picture characterized by prolonged phase of inhibited HBV replication and apparent quiescence of the liver disease alternating with phases of reactivation of the viral replication leading to the exacerbation of the liver damage (Perrillo, 2001). Such types of reactivation in these cases may take a fulminant course also (Tassopoulos et al, 1987) and carries a high risk of cirrhosis and HCC development (Perrillo, 2001). Some of these individuals are HBeAg positive, but in some geographical areas like the Mediterranean basin this subgroup is typically anti-HBe positive (Tassopoulos et al, 1987; Raimondo et al, 1990; Colloredo et al, 1994).

(C) HBsAg negative subjects with occult HBV infection: HBV infection may persist in the liver even in the absence of HBsAg (Raimondo et al, 2000; Raimondo et al, 2001; Brechot et al, 2001; Conjeevaram and Lok, 2001; Torbenson and Thomas, 2002). This occult infection is mostly found in anti-HBc and/or anti-HBs positive individuals. It also occurs in individuals negative for all serum markers of HBV infection (Caccicola et al, 1999). These patients with occult HBV infection may carry both integrated and free HBV DNA (Marusawa et al, 2000). In some cases, lack of HBsAg detection is due to changes or variations in the HBV genome that interfere with gene expression or lead to the production of an antigenically modified S-protein that is not recognized by commercially available kits (Yamamoto et al, 1994; Carman et al, 1997; Hou et al, 2001).
2.5 Natural history of chronic hepatitis B:

Acquisition of HBV in adulthood generally results in the development of a symptomatic transient infection, which resolves in the majority of cases and is accompanied with the development of neutralizing antibodies during late convalescence. However, infection of infants and children frequently leads to persistent HBV carriage with a significant risk for the development of chronic active hepatitis, cirrhosis and hepatocellular carcinoma (Bortolotti, 1994). It is mainly host factors such as age, gender and immune-competence, which determine the outcome of HBV infection.

Lok and colleagues (1995) have developed a model of the natural history of chronic hepatitis B acquired early in life. The clinical, immunological, pathological and virological features have been divided into three phases.

**Phase I:** It is the immunotolerant phase characterized by high viral load, modest immune response and minor changes in liver biopsy. The carrier in this phase generally remains asymptomatic. The hepatitis B e antigen (HBeAg), (actively secreted in wild type HBV infection), is regarded as a tolerogen and this acts to mask the major T cell determinants of hepatitis B core antigen (HBcAg), later is located on the viral nucleocapsid protein.

**Phase II:** The immuno-elimination, or second phase, has been proposed to be associated with flares of hepatitis in this model. This phase is associated with a conversion (seroconversion illness) from HBeAg to anti-HBe positivity in serum. It may be asymptomatic or mildly symptomatic in a large number of cases and may go unnoticed. In this phase, the host generally overcomes the tolerogenic effects of chronic infection and the immune response remains actively engaged in eliminating HBV infected hepatocytes from the liver.

**Phase III:** Finally, it comes the third phase where negligible viral loads are found. While many patients still have minimal changes on liver histology, a significant proportion progresses to cirrhosis. During this phase most of the viral DNA is in the form of integrated HBV DNA in the host cell genome (i.e. latent state). It is in this third stage where the risk for the development of the complications of chronic liver disease including end stage liver failure and hepatocellular carcinoma is most pronounced.

Variations on this model are seen and one common variation is the patient with HBeAg negative chronic hepatitis B, which is due to persistent infection with a pre core mutant virus. These patients have abnormal liver function tests, are HBeAg negative, and/or anti-HBe positive, with low but detectable viral DNA. This group of patients most likely represents the effects of the immune response applying a selective pressure to the virus, but being unable to eradicate all the HBV quasispecies. In this group, the seroconversion to anti-HBe does not reflect the end of the replicative activity of the virus, and does not suggest imminent viral clearance (Hadziyannis et al, 1995).
Correct staging of chronic hepatitis B according to the particular phases in the natural history of chronic HBV is important in order to best plan and interpret the results of clinical trials, and to individualize patient management. For example, patients found in phase I are typically non-responders to IFN-alpha and might be better managed by first reducing the high viral load with agents like nucleoside analogues. If such therapy is effective in blocking viral replication completely, including the elimination of circulating serum HBeAg, recognition of viral antigens may occur on reappearance of viremia and antigenaemia that occurs with cessation of nucleoside analogue therapy. The breaking of immunotolerance may then result in viral clearance. Finally, patients with HBeAg negative hepatitis (infection with the pre core mutant of HBV) with grade 4 fibrosis or cirrhosis may be better managed on long term chemosuppression with a nucleoside analogue such as famciclovir/lamivudine, where the viral load and associated necroinflammatory activity is reduced, allowing the liver to regenerate.

2.6 Genetic variability in hepatitis B viruses:

2.6.1 HBV subtypes:
The first report of variability in HBV came from Le Bouvier (1971) who described two mutually exclusive subtype determinants, \(d\) and \(y\). These reside in the surface protein together with the main antigenic determinant ‘\(a\)’ (Levene & Blumberg, 1969). Two additional determinants, \(w\) and \(r\), were described by Bancroft et al, (1972), who found that each HBV strain could be characterized as belonging to either subtype \(adw\), \(adr\), \(ayw\) and \(ayr\). In a larger study additional subtypes were characterized by Courouce-Pauty et al, (1983). The nine subtypes described were \(ayw1\) to \(ayw4\), \(ayr\), \(adw2\), \(adw4\), \(adrq\) and \(adrq^+\). A geographical pattern for the distribution of subtypes was confirmed. Their geographical distribution was found to be stable over two decades. The subtype ayw was described in the Mediterranean countries (\(ayw2\), \(ayw3\)), in West- and Central-Africa (\(ayw4\), \(ayw2\)) and in Vietnam (\(ayw1\)). The rare subtype ayr is found in Vietnam. The subtype \(adw\) is predominant in North-West-Europe, in America, in East and South Africa, in India (\(adw2\)) and in South-East Asia (\(adw4\)). The subtype \(adr\) can be found in Polynesia (\(adrq\)) and in South-East-Asia (\(adrq^+\)) (Kidd-Ljunggren, 1996; Courouce-Pauty et al, 1983).

During the 1980s, it became increasingly clear that the subtype determinants are specified by one single amino acid, at positions 122 (d or y) and 160 (r or w) in the S protein, respectively (Okamoto et al., 1987a,b; Ashton-Rickardt & Murray, 1989a,b; Norder et al., 1991). Subtype determinants d and w have a lysine at both positions, whereas an arginine at both positions indicates subtype determinants y and r. Additional subtype determinant reactivities have been mapped to amino acid positions 127, 144, 145, 158, 159, 177 and 178 (Okamoto et al, 1989; Norder et al, 1992a). Subtyping of HBV strains was used for epidemiological purpose, and, to find correlations between disease and a particular subtype. Over the last decade, however, subtype determination has gradually been replaced by genotyping.
2.6.2 HBV genotypes:

In 1988, Okamoto et al, (1988) first suggested that the traditional subtypes could be complemented or replaced by a classification of different HBV strains into genetic subgroups. Comparing the full nucleotide sequences of 18 HBV strains they found that these clustered into four groups, A to D, with more than 8% divergence between the groups. This degree of divergence has since become the definition for HBV genotype. The correspondence between subtypes and genotypes is shown in the table.

Using the polymerase gene instead of the whole genome Orito et al, (1989) also found four separate subgroups, differing slightly from the groups described by Okamoto et al, (1988). Comparisons of S gene sequences were done by Norder et al, (1992b). In addition to results similar to those by Okamoto et al, (1988), two more groups, E and F were described. In a larger study, they compared the S gene sequences from 122 strains and confirmed the existence of the two new groups i.e. E and F (Norder et al, 1993b). Recently, two more genotypes G and H have been described (Stuyver et al, 2000; Arauz-Ruiz et al, 2002).

2.7 Geographical distribution of HBV genotypes:

HBV Genotype A is mainly found in Northeastern areas of Europe and North America (Norder et al, 1993b). Some genotype A strains have also been found in the Philippines (Norder et al, 1993b; Kidd-Ljunggren et al, 1995). Genotype B and C belong in the indigenous population of Southeast Asia (Theamboonlers et al, 1999). Genotype D is the most widely distributed genotype and has been found universally, with its highest prevalence in a belt stretching from Southern Europe and North Africa (Norder et al, 1993b; Borchani-Chabchoub et al, 2000) to India. Genotype E is found in West and South Africa. The most divergent genotype F is found in South and Central America (Norder et al., 1993a; Arauz-Ruiz et al., 1997; Nakano et al, 2001). Genotype G has been found in France and USA (Stuyver et al, 2000). Table 2.1 shows the geographical distribution of HBV genotypes.

2.8 Structural differences between HBV genotypes:

Amongst HBV strains belonging to HBV genotype D, there is 33 nt. deletion in the preS1 region (Heermann & Gerlich, 1991). No such deletions have been seen in other genotypes, even in members of more closely related genotype E (Norder et al, 1994; Bowyer et al, 1997). Similarly, genotype A has 6 extra nucleotides at the carboxy terminal, in its core region that remains absent in all the HBV genotype sequences (Stuyver et al, 2000).
2.9 Recombinant genotype:

The major issue related to the clinical significance of HBV genotypes is its precise determination in a given isolate. Unless whole genomic sequences are used, differentiation of HBV strains into separate genotypes could depend on which part of the genome is used in the phylogenetic analysis. The rooted phylogenetic tree of different HBV genes from several strains varied considerably and illustrated the importance of selecting the appropriate regions for analysis (Norder et al., 1994). One possibility for such large variability in the phylogenetic tree is the recombination (Robertson et al., 1995a; Robertson et al., 1995b).

Recombination, although not genotype related, was reported from HCC-associated HBV (Georgi-Geisberger et al., 1992). When analyzing individual ORFs and complete genomes, respectively, from a number of HBV strains, Bollyky et al., (1996) found evidence for recombination in two strains. In both cases, the mosaic strains originated from geographical areas where several genotypes are known to exist. Several additional studies support the theory of recombination between genotypes that co-circulate in some geographical regions, such as genotype B/C switching and genotype A/D switching (Bowyer et al., 2000; Morozov et al., 2000; Owiredu et al., 2001). Two studies analyzing Vietnamese strains found recombination between genotypes C and A (Hannoun et al., 2000) and B and C (Yuasa et al., 2000), respectively.

Recently, Sugauchi et al., (2002) classified hepatitis B virus of genotype B with or without recombination with genotype C over the pre core region plus the core region and provisionally named as Ba (‘a’ standing for Asia) and Bj (‘j’ for Japan). They have also shown the varying epidemiological and virologic characteristics of genotypes Ba and Bj, subsequently (Sugauchi et al., 2003).

2.10 Influence of HBV genotype on the natural course of chronic HBV infection:

HBV subtypes and genotypes can influence the natural course of chronic HBV infection and the outcome to therapy (Kao and Chen, 2000; Chen, 1993; Chu, 2000). Genotype A is associated with chronic liver disease more frequently than genotype D, whereas most patients with genotype D present with acute hepatitis in Europe (Mayerat et al., 1999). A higher disease-inducing capacity of genotype C than genotype B is observed in Asia (Kao et al., 2000; Lindh et al., 1999; Orito et al., 2001b). In Taiwan, however, genotype B (HBV/B) is reported to enhance the development of HCC in hosts of younger ages (Kao et al., 2000), in striking contrast to genotype B in Japan (Orito et al., 2001a) (Table 2.1).

There is no clear explanation for these different clinical outcomes in patients infected with apparently similar genotype. There is a possibility that such clinical differences in carriers of genotype B could be due to distinct HBV strains infecting the hosts of genotype B. We have recently shown that the genotype D is associated with more severe liver disease and may predict the occurrence of HCC in young Indian patients (Thakur et al., 2002). On the other hand, the lack of clinical significance of hepatitis B virus genotypes and serotypes has been reported in patients from western India (Gandhe et al., 2003).
2.11 Molecular structures and functions of the three viral antigens

Molecular cloning and sequencing of the HBV genome led to the redefinition of the three HBV antigens as viral gene products endowed with specific functions in viral life cycle. The HBCAg and HBeAg are alternative translation products of the core gene, with HBeAg translation requiring an upstream precore region ATG codon (Fig. 2). The HBeAg (called “core protein” nowadays) assembles into viral nucleocapsid (core particle), which packages the pregenome (an RNA copy of viral DNA) and polymerase. Inside the core particle, the viral polymerase directs the synthesis of minus strand DNA from the RNA template. It then degrades the RNA pregenome and generates the plus strand DNA via the minus strand template. The HBsAg is the envelope protein of the virus, and actually comprises three co-terminal proteins (large,
middle, and small) due to the presence of multiple transcripts and alternative translation initiation sites in the gene. They contain preS1/preS2/S domains, preS2/S domains, and S domain, respectively.

The small envelope protein, composed of S domain alone, is the most abundantly expressed. The envelope proteins interact with the nucleocapsid to initiate its envelopment, and the resultant virus particle (virion) is released into the bloodstream. Thus, HbcAg is not detectable in patient blood unless the envelope is removed. In addition to their incorporation into virus particles, the envelope proteins can be secreted alone as non-infectious subviral particles, which constitute the bulk of HBsAg as detectable in patient blood.

The N-terminal 29 residues of the HBeAg precursor are specified by the precore region, the first 19 of which serve as the signal peptide to target the protein to the endoplasmic reticulum, where it is cleaved off. Further down the secretory pathway the arginine rich C-terminus of the molecule is removed, thus releasing mature HBeAg into blood stream. Therefore, HBeAg differs from core protein by a longer N-terminus and shorter C-terminal tail. However, thanks to an intramolecular disulfide bond HBeAg has a secondary structure quite different from that of core protein [Nassal et al, 1993 and Wasenauer et al, 1993]. Only one of the two major B cell epitopes of HBeAg is shared with the core protein. HBeAg is not part of the virus particle, and its true function remains not fully understood. Expression of HBeAg is not required for virus replication in vitro [Tong et al, 1991]. Ablation of e antigen expression had no effect on the in vivo infectivity of the duck hepatitis B virus, but curtailed infection for the woodchuck hepatitis virus (which is more closely related to the human virus) [Chang et al, 1987, Chen et al, 1992 and Schlicht et al, 1987]. It was proposed that expression of HBeAg during perinatal infection, the major mode of HBV transmission in Asia, induces immune tolerance. Another potential role of HBeAg in promoting persistent infection is to mimic core protein so as to buffer immune attack of the infected hepatocytes by the anti-HBc antibodies.

2.12 Role of anti-HBe immunity in clearing HBV infection

Of the three antibodies against HBV, anti-HBc develops first, whereas anti-HBs antibody is detected last. The reason for this sequence remains unknown. The HBsAg is the most abundantly expressed protein of HBV, whereas core protein has probably the lowest abundance due to its location inside virus particles. Whether large excess of subviral particles, a unique feature of hepatitis B virus family, delays the development of anti-HBs antibody, has not been experimentally tested. Anti-HBc antibody rises soon after infection but is not associated with change in viremia titer. This could be related to the presence of HBeAg, the variant core protein, as a decoy. The anti-HBe antibody is not expected to directly neutralize viral infectivity, because virus particle does not contain HBeAg. The declined viremia following anti-HBe
development could be attributed to loss of HBeAg, which unleashes the anti-viral effect of the anti-HBc immunity. Alternatively, anti-HBe antibodies could destroy infected hepatocytes by recognizing HBeAg on the cell surface, although this aspect remains more or less speculative. The anti-HBs antibodies are known to bind envelope proteins on viral surface to prevent infection. This is the basis for using HBsAg as preventive vaccine against HBV infection.

2.13 Types of HBeAg variants

The anti-viral effect of anti-HBe immunity may explain the frequent emergence of HBeAg variants in patients with anti-HBe. Since HBeAg expression is not essential for virus replication, the simplest way for the virus to evade the anti-HBe immunity is to switch off HBeAg expression altogether. The so-called “precore mutants” are the first discovered major immune escape mutants of HBV. These mutants are characterized by a G1896A nonsense mutation in the precore region that truncates the precore/core protein into a 28-aa peptide [Brunetto et al, 1989, Carman et al, 1989 and Tong et al, 1990]. Other nonsense and frameshift mutations inside the precore region have also been found, although less frequently. Point mutations of the precore ATG codon have also been observed, which prevent initiation of translation. We recently found that triple mutation at the –5, -3, and –2 positions of the precore ATG codon, as occasionally found in some South African strains of HBV, greatly reduced translation efficiency [Ahn et al, 2003]. The selective disruption of HBeAg expression through mutations affecting the precore region rather than the core gene can be easily understood in terms of the indispensable role of core protein for viral replication.

The second common HBeAg variants are the core promoter mutants. They are characterized by point mutations in the promoter for both HBeAg mRNA and core protein mRNA (also called pregenomic RNA) [Okamoto et al, 1994]. These mutations were found by transfection experiments to down regulate HBeAg mRNA production, resulting in reduced protein levels [Buckwold et al, 1996 and Scaglioni et al, 1997]. Core promoter mutants are the dominant viral species at not only the anti-HBe stage, but also the late HBeAg stage of infection (Fig. 1). It should be pointed out that the common core promoter mutations, A1762T/G1764A, reduced HBeAg expression by a mere 20% in a genotype A clone that we examined [Parekh et al, 2003]. It is not clear why a moderate reduction in HBeAg expression offers survival advantage, and why the selection is in place well before the rise of anti-HBe.

Considering the many steps required for the secretion of HBeAg, we have recently systemically tested other possible avenues whereby HBeAg expression can be regulated. A V17F missense mutation at the –3 position of the signal peptide cleavage site has been proposed to impair HBeAg production. However, our
transfection experiments failed to find a major impact of this mutation on HBeAg secretion (Guarnieri et al., 2006). On the other hand, naturally occurring mutations at the C-terminal cleavage site were found to reduce HBeAg secretion (Kim et al., 2009). HBe antigenicity can also be abolished by mutation of one of the two cysteines implicated in the disulfide bond. While mutation of the core gene-derived cysteine into serine or phenylalanine did not interfere with viral replication or virion secretion, substitution of the precore-derived cysteine greatly compromised viral genome replication (Bang et al., Virology, in press). The relevant cysteine codon, when present at the 5’ end of pregenomic RNA, constitutes the loop of pregenome encapsidation signal. Finally, we observed an E77Q mutation in the core gene of many naturally occurring core promoter mutants of genotype A (but not from clones with wild-type core promoter sequence), which abolished recognition of both core protein and HBeAg by a rabbit polyclonal antibody (Kim et al., 2009). This finding is consistent with the localization of the immunodominant epitope of the core protein within residues 77-84 [Salfeld et al., 1989 and Belnap et al., 2003] Whether the selection of this missense mutation is driven by the need to escape the anti-HBc or anti-HBe immunity is an open question.

2.14 Infection with core promoter mutants is associated with more severe forms of liver diseases

Among the HBeAg variants, the core promoter mutants deserve special attention. Many cases of fulminant hepatitis have been traced to infection with core promoter mutants. However, since fulminant hepatitis is a rare form of acute infection, most reports are descriptive. A case-controlled study is needed to rigorously test whether core promoter mutants are more likely to produce fulminant hepatitis than the wild-type isolates from the same region. Since HBV is considered a noncytolytic virus, the immune response plays a crucial role in generation of liver injury. Therefore, contribution of the genetic makeup of the host in fulminant hepatitis should not be overlooked. It is also necessary to point out that a core promoter mutant often contains many other genetic alterations within the viral genome. Thus, even when a core promoter mutant elicits fulminant hepatitis, mutations elsewhere in the viral genome could be responsible. In one well-documented case, transmission of a core promoter mutant resulted in outbreak of fulminant hepatitis [Liang et al., 1991]. In another study, a core promoter mutant associated with fulminant hepatitis was found to induce more severe liver damage when experimentally inoculated into chimpanzees [Ogata et al., 1993]. These observations provide compelling evidence for the intrinsic virulence of some core promoter mutants.

During chronic infection, core promoter mutants have been linked to more severe forms of liver diseases including liver cancer. A study from South Africa revealed prevalence of core promoter mutations in 66% of HCC patients but only 11% of asymptomatic carriers matched in age and HBeAg / anti-HBe status [Baptista et al., 1999]. Similarly, core promoter mutations were present in only 3% of Taiwanese inactive carriers but up to 64% of HCC patients [Kao et al., 2003]. Certainly, prospective epidemiological studies will be needed
to demonstrate that rise of core promoter mutations precede cancer development. Another piece of evidence for the enhanced pathogenicity of core promoter mutants came from comparative studies of HBV genotypes. East Asian patients are primarily infected with genotype C or B of HBV, with a North to South transition. Interestingly, genotype C patients often suffer from more severe liver diseases, delayed HBeAg to anti-HBe seroconversion, and accelerated HCC development as compared with genotype B patients [Chu and Lok 2002]. Further analysis revealed that genotype C isolates are more likely to develop core promoter mutations than genotype B [Kao et al, 2003, Orito et al,2001a, Sumi et al,2003].

It has been recently suggested that core promoter mutations, rather than genotype C per se, are the primary risk factor for liver cancer [Yuen et al,2004]. Like core promoter mutations, the G1896A HBeAg-negative precore mutation develops late in the course of HBV infection. However, the prevalence of the precore mutation was not elevated in cancer patients relative to matched controls [Yuen et al,2004]. Thus, the association between core promoter mutations and liver cancer is genuine.

2.15 Core promoter mutations cumulatively enhance viral genome replication in vitro

While a single nucleotide change or insertion/deletion is often present in precore mutants and sufficient to abolish HBeAg expression, the number and position of mutations in the core promoter vary. The mutations are clustered around nucleotides 1750—1770, with the A1762T and G1764A being the most common. Therefore, the A1762T/G1764A double mutation has been chosen for further characterization through site-directed mutagenesis and transfection experiments. Many independent studies have been performed and the double mutation appears to reduce HBeAg expression (by only 20% in our hand) and double the genome replication capacity [Buckwold et al,1996, Scaglioni et al,1997, Tang et al,2001,Yu and Mertz 2003].

From a different perspective, the HBV isolate implicated in fulminant hepatitis outbreak was found to replicate at least 10 times higher levels than a wild-type clone [Hasegawa et al,1994]. This genome is both a core promoter mutant and precore mutant. It contained A1762T, G1764A, C1766T, and T1768A mutations in the core promoter. Mapping experiments and site-directed mutagenesis revealed the 1766/1768 double mutation, rather than common mutations at 1762 and 1762, as responsible for the enhancement of viral replication [Baumert et al., 1996].

Clones derived from highly viremic patients with wild-type virus shows low replication capacity, while some clones derived from low viremia patients with mutated viruses has been shown to have much higher replication levels [Orito et al, 2006]. Sequencing analysis revealed core promoter mutations in the high replicating but not low replicating clones. The highest replicating clones contained T1753C/A1762T/
G1764A/C1766T quadruple mutation or 1762/1764/1766 triple mutation, and the next highest replicating clone contained 1753/1762/1764 mutations. Site-directed mutagenesis of a wild-type clone revealed 2-, 4-, 8-, and 8-fold enhancement of viral replication by the 1762/1764, 1753/1762/1764, 1762/1764/1766, and 1753/1762/1764/1766 mutations, respectively [Parekh et al, 2003]. These mutations reduced HBeAg expression by 20%, 30%, 75%, and 80%, respectively. These results provide compelling evidence that core promoter mutations enhance viral genome replication and reduce HBeAg expression in a cumulative manner. In this regard, the 1762/1764 mutations emerge first, followed by the less common mutations in the core promoter. Our findings suggest the gradual loss of HBeAg expression and enhancement of viral replication capacity over the course of chronic HBV infection.

Although the core promoter mutants were initially identified as HBeAg variants, they have been associated with fulminant hepatitis and liver cancer in vivo and found to display enhanced replication and sometimes impaired virion secretion in vitro. It is tempting to suggest that the enhanced replication capacity and reduced virion secretion may increase viral load in the liver, thus triggering liver damage either directly or indirectly through the immune response. When massive liver damage occurs during acute infection, fulminant hepatitis may ensue. When such damage occurs during chronic infection, it increases hepatocyte turnover, induces fibrosis, and enhances the chance of hepatocellular transformation and malignancy.

2.16 Viral load

Several recent studies suggested that HBV viral load is associated with HCC (Yu et al, 2005; Chen et al, 2006, 2009; Chan et al, 2008; Wu et al, 2008) (Fig 2.5). The REVEAL-HBV study found that after adjusting for age, sex, smoking, alcohol use, HBeAg status, ALT level and cirrhosis, the risk of HCC was approximately six times higher for persons with viral load above $10^5$ copies per ml compared with persons with undetectable viral load (Chen et al, 2006). HBV viral load is also closely related to the development of cirrhosis (Iloeje et al, 2006).

HBV DNA detection and HBV DNA level measurement is essential for the diagnosis, decision to treat and subsequent monitoring of patients. There are multiple methods used for screening of viral loads regularly, with different ranges of quantitation and detection; and based on different methodology (Table 2.2). The detection titers are based on varies from one system to another (Fig 2.6).
Figure 2.5: Increased viral load is associated with increased risk for developing HCC. (Chen CJ, et al, JAMA. 2006; 295(1):65-73.)

Table 2.2: Different procedures for HBV viral load detection and their detection range

<table>
<thead>
<tr>
<th>Test</th>
<th>Method</th>
<th>range of quantitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digene</td>
<td>RNA-DNA hybrid</td>
<td>$1.4 \times 10^5 - 5.6 \times 10^9$ genomes/mL</td>
</tr>
<tr>
<td>Abbott</td>
<td>Liquid hybridization</td>
<td>$&lt;10^9$ - $\infty$ genomes/mL</td>
</tr>
<tr>
<td>Bayer</td>
<td>Branched DNA</td>
<td>$7 \times 10^5 - 5 \times 10^9$ DNA equiv/mL</td>
</tr>
<tr>
<td>Amplicor monitor</td>
<td>Quantitative PCR</td>
<td>$4 \times 10^2 - 4 \times 10^7$ genomes/mL</td>
</tr>
</tbody>
</table>
Figure 2.6: HBV quantification range via different strategies. (adapted from Mutimer D. *EASL*. 2001.)

**Lacunae and importance of the current study:**

Limited data is available on Hepatitis B infection, family based Seroprevalence, and the genotype distribution in entire Northeast India. Since the distribution of genotype determines the predisposition to different grade of HBV related liver disease severity, a large cohort based planned study is indispensable in Northeast India which has an ethnically distinct population compared to other parts of India. Moreover, correlation of the difference of HBV genotypes needs to be correlated with HBeAg status and viral load and assessed as risk factors in different stages of HBV related chronic liver disease cases, which will give us a comprehensive picture of severity and extent of HBV related liver disease in NE India and will in turn help to manage the HBV infected patients and the HBV related disease burden better in the future.