CHAPTER 2: REVIEW OF LITERATURE
2.1 Viruses that cause Hepatitis:

Hepatitis means 'inflammation of liver'. Essentially, any abnormality that results in compromised liver function is termed as hepatitis. It can be caused by a number of agents such as viruses, bacteria, toxic agents, drugs and alcohol. A wide group of viruses are known to cause hepatitis besides hepatotropic viruses. These include Cytomegalovirus, Epstein-Barr virus, Herpes Simplex virus, yellow fever virus, Rubella, Parvovirus, and viruses such as Lassa, Ebola, and Marburg. Infection with these viruses occasionally induces hepatitis, especially in certain age groups or in certain parts of the world. They do not fall under the category of hepatitis viruses as they are not primarily hepatotropic.

Hepatitis induced by hepatotropic viruses represents a disease entity caused by at least six unrelated agents whose primary tissue tropism is the liver. The liver diseases resulting from hepatitis viruses share the common characteristic of causing inflammation of the liver. Hepatitis viruses affect a significant population in the world and are a serious public health concern requiring considerable effort to ensure that the blood, water, and food supply remains free of these viruses. Hepatitis has been described in writings of a number of ancient societies as early as the 5th century. However, the discovery of hepatitis viruses began in the mid-1960s with the hepatitis B virus (HBV) and continued with the hepatitis A virus (HAV) (Krugman et al., 1967) and hepatitis D virus (HDV) (Rizzetto et al., 1977) in the 1970s. Advances in molecular techniques in the late 1980s, especially the polymerase chain reaction (PCR), ushered in the discovery of two new viruses, hepatitis C virus (HCV) (Choo et al., 1989) and hepatitis E virus (HEV) (Wong et al., 1980); earlier referred to as ( non A non B hepatitis) NANBH viruses. The most common cause of viral hepatitis are hepatitis A, B, C, and E. Recently, the discovery of another hepatitis virus referred to as GB-C or HGV, has been reported (Simons et al., 1995; Yoshiba et al., 1995).

Hepatitis A is an acute self-limiting disease resulting from infection with Hepatitis A virus (HAV). HAV is a small non-enveloped RNA virus belonging to Picornaviridae family. This viral RNA encodes a large polyprotein from which non-structural proteins are subsequently cleaved (Wheeler et al., 1986). Hepatitis B virus (HBV) is a member of Hepanaviridae. HBV has a small circular DNA genome that is
partially double-stranded and replicates through an RNA intermediate that is transcribed with a viral reverse transcriptase enzyme (Tiollias et al., 1981). Infection with HBV can produce a chronic infection that might lead to primary hepatocellular carcinoma in some patients. HDV, also called the Delta virus, is an "incomplete" virus and the only viroid known to infect man. It contains a small single-stranded RNA, a unique internal protein, the delta agent and an outer coat of hepatitis B surface antigen. HDV requires an active HBV infection to replicate (Rizzetto et al., 1980).

The recent identification of two new hepatitis viruses has almost completely characterized the disease entity previously designated non-A, non-B hepatitis, which was defined by the serological exclusion of known causes of viral hepatitis and their epidemiological characteristics. Hepatitis C virus infection is responsible for the majority of cases of parenterally transmitted non-A, non-B hepatitis and produces a persistent infection that is often associated with chronic liver disease (Houghton et al., 1991). HCV is a single-stranded positive sense RNA virus with a genome of 9.4 kb, having a single open reading frame. The 5' end of the genomic RNA encodes the nucleocapsid and envelope proteins, followed by non-structural proteins that extend to the 3' end of the genome. HCV has been classified as a separate genus in family Flaviviridae (Houghton et al., 1991). Like HBV, HCV is also a major contributing cause of hepatocellular carcinoma (Kiyosawa et al., 1990).

In search for new agents responsible for viral hepatitis, two groups independently reported the isolation of blood-borne virus, as hepatitis G virus (HGV) (Simons et al., 1995) and GB virus C (GBV-C) (Yoshiba et al., 1995). These were later shown to be isolates of the same virus. GBV-C or HGV is an enveloped RNA virus belonging to the Flaviviridae family. GBV-C/HGV is transmitted by contaminated blood or blood products, intravenous drug use, from mother to child, sexually and possibly through close social contacts (Abe, 2001; Aikawa et al., 1996; Linnen et al., 1996). Several reports indicate a high prevalence of GBV-C/HGV viremia (1-4%) within healthy populations in Europe and North America (Alter et al., 1997a; Alter et al., 1997b), and even higher prevalence (10-33%) among residents in South America and Africa (Abe, 2001). GBV-C/HGV has been suggested to be the causative agent for non-A non-E hepatitis.
Another possible hepatitis virus, TT virus, was discovered in the blood of a patient with post-transfusion non-A non-E hepatitis (Nishizawa et al., 1997). By using PCR primers designed to overcome the high nucleotide sequence divergence, TT virus was found to be ubiquitous with a world wide distribution. A disease association is thus unlikely (Okamoto and Mayumi, 2001). Most recently a DNA virus designated as SEN-V has been implicated as a major cause of non-A non-E hepatitis. Based on limited data available to researchers, SEN-V is the most likely agent (Bowden, 2001).

2.2. Hepatitis E:

Large epidemics of viral hepatitis have been reported since the mid-1950s from Asia, mainly the Indian subcontinent. These viral epidemics were first thought to be due to Hepatitis A virus. After the development of serological tests for the diagnosis of HAV infection in the 1970s, retrospective examination of acute and convalescent serum specimens from these early outbreaks indicated that the vast majority of cases were not attributed to HAV infection (Wong et al., 1980). These observations defined a new form of non-A, non-B hepatitis that was transmitted via the fecal-oral route. Subsequently, numerous outbreaks of this enterically transmitted non-A, non-B hepatitis were identified in many parts of the world because of the widespread use of IgM testing for the diagnosis of acute hepatitis A. In addition, this new disease had several distinct epidemiological features that differentiated it from hepatitis A, including a high attack rate among adults and an unusually high mortality rate among pregnant women (Joshi et al., 1985; Khuroo et al., 1981). This form of viral hepatitis was initially designated epidemic non-A, non-B hepatitis and then enterically transmitted non-A, non-B (ET-NANB) hepatitis until an etiologic agent was fully characterized in the late 1980s. Afterwards, it was designated hepatitis E.

2.2.1. Epidemiology of HEV:

Hepatitis E has been reported to cause large epidemics in several countries on all continents as shown in next page (Figure 2.2.1). It was first identified after an outbreak in New Delhi, in 1955, in which twenty-nine thousand cases of hepatitis were identified, following wide-spread fecal contamination of the city’s drinking water (Khuroo, 1980). Similar outbreaks have been reported as sporadic infections from different parts of the
Review of Literature

world including India (Khuroo et al., 1993; Jameel et al., 1992; Khuroo, 1980), Nepal (Kane et al., 1984), Myanmar (Shwe and Soc, 1985), Pakistan (De Cock et al., 1987), Mexico (Velazquez et al., 1990), Soviet Union (Balayan et al., 1983), Africa (Sarhou et al., 1986; Belabbes et al., 1985), North America (Fortier et al., 1989), Europe (Jardi et al., 1993; Lavanchy et al., 1994) and Australia (Moaven et al., 1995). The disease is found most frequently in developing and underdeveloped countries where fecal contamination of drinking water is common (Jameel, 1999).

**Geographic distribution of human HEV**

(CDC data)

Fig 2.2.1 CDC image showing global distribution of HEV.

Acute hepatitis E has also been reported from developed countries such as Australia (Moaven et al., 1995), France (Coursaget et al., 1994), Israel (Karenyi et al., 1995), The Netherlands (Zaalijer et al., 1993), Spain (Jardi et al., 1993), UK (Mutimer et al., 1995; Sallie et al., 1994) and USA (CDC, 1993). These are mainly associated with travel to endemic areas. However, rare cases of acute hepatitis E infection, not acquired from HEV endemic areas, have also been reported from United States (CDC, 1993) and Europe (Coursaget et al., 1993; Zaalijer et al., 1993; Wang et al.,
About 1.5% healthy adults showing anti-HEV positive and cases of fulminant hepatic failure associated with HEV in Europe (Sallie et al., 1994) indicates that the prevalence of infection is higher than expected. The age-specific prevalence of anti-HEV antibodies have also been studied in endemic areas (Arankalle et al., 1995; Mushawar et al., 1993). These increase till the third decade of life and then stay constant. This pattern does not suggest a unique outbreak of HEV in the population in the recent past, but sporadic transmission that accumulates over age (Arankalle et al., 1995; Mushawar et al., 1993; Schlauder et al., 1993).

The mortality rate of HEV infection is 1-2% that is 10 times more than that of HAV. Incidentally it is even higher in pregnant women reaching up to 10-20% (Kuru et al., 1993; Bradley et al., 1990; Purcell and Ticehurst, 1988).

2.2.2. Mode of Transmission:

Hepatitis E is a water-borne disease. It is mainly transmitted feco-orally through contaminated drinking water (Naik et al., 1992; Ramalingaswami and Purcell, 1988; Belabbes et al., 1985; Sreenivasan et al., 1978; Vishwanathan 1957). Other modes of transmission like person to person contact (Coursaget et al., 1994; Velazquez et al., 1990) and parenteral transmission (Wang et al., 1993; Sarthou et al., 1986) have also been reported. Vertical transmission of HEV has been reported in humans with infants showing clinical, serological or virological evidence of HEV infection at birth (Nanda et al., 1995). However no documentation of neonatal HEV infection in infected pregnant rhesus monkey has been found (Tsarev et al., 1995; Arankalle et al., 1993).

2.2.3. Clinical features of HEV:

The clinical features of hepatitis E are typical of acute viral hepatitis. Patients with symptomatic HEV infection manifest abdominal pain, anorexia, joint-pain (arthralgia), clay-colored stool, dark colored urine, diarrhea, fever, hepatomegaly, jaundice, nausea, rash, splenomegaly and vomiting. Cholestasis (bile-stasis) is predominant in about 20-57% of the patients (Khuroo, 1980).

The incubation period varies from 15-50 days. In case of a human volunteer, clinical symptoms developed 32 days after oral-intake of infected fecal samples (Chauhan et al., 1993). The pre-icteric phase lasts for 1-10 days with nonspecific symptoms like epigastric pain, nausea and vomiting. Malaise is the most common
characteristic complain followed by fatigue and loss of appetite. The icteric phase begins with appearance of jaundice and clay-colored stool due to hyper bilirubinemia. Peak infectivity occurs during the last asymptomatic days of the incubation period and early days of acute symptom coinciding with appearance of virus in bile and feces during the same time. Liver function tests are indicative of hepatic necrosis. Laboratory findings include elevated serum bilirubin, alanine aminotransferase (ALT), alkaline phosphatase and gamma-glutamyl transferase. Elevation of ALT levels occur as single peak preceding or coinciding with the onset of jaundice, which is similar to most other forms of viral hepatitis (Balyan, 1990; Khuroo, 1980). Figure 3 shows the correlation between clinical disease, viremia and antibody response.

Fulminant hepatic failure is an uncommon consequence of acute hepatitis with mortality reaching up to 90% with the onset of encephalopathy (O'Grady et al., 1989; Bermuau et al., 1986). During HEV epidemics unusually high rates of fulminant hepatitis leading to about 20% mortality rate, have been observed in pregnant women (Joshi et al., 1985; Khuroo et al., 1981). The highest rates of fulminant hepatitis and death have occurred during the 20th to 32nd week of gestation, as well as during labour (Khuroo et al., 1981). Death is usually due to encephalopathy, haemorrhagic diathesis or renal failure. Intrauterine infections have also been reported (Khuroo et al., 1995).

Although the mechanisms are not known, a hypothesis has been proposed to explain the pathogenesis of fulminant hepatitis E in pregnant women (Purcell and Ticehurst, 1988). This suggests that the virus damages the sinusoidal cells, particularly the Kupffer cells. This in turn diminishes their ability to protect the hepatocytes against endotoxins originating from Gram-negative bacteria in the intestinal tract. Direct injury of hepatocytes by endotoxins and secondary injury, mediated through the release of eicosanoids, leads to prostaglandin-mediated chemotactic attraction of neutrophils and leukotriene-mediated oedema and cholestasis. The enhanced sensitivity of pregnant women to such an endotoxin-mediated effect is well recognized and may explain the striking mortality rate of hepatitis E in pregnancy (Purcell and Ticehurst, 1988). The validity of this hypothesis, however, has not been confirmed.
Most Hepatitis E outbreaks have been reported in young to middle age adults (15-50 years of age) (Khuroo, 1991). Humans and experimentally infected primates demonstrate an early acute-phase response characterized by the presence of both IgM and IgG antibodies. As shown in the figure (Fig. 2.2.3), IgM titers peak with maximal ALT activity and decline within 5-6 months. IgG response begins shortly after IgM and remains high from 1-4.5 years after acute phase of disease (Khuroo et al., 1993). The long-term persistence of IgG anti-HEV is further suggested by its detection in asymptomatic persons living in endemic areas. Seroconversion may be critical for early clearance of the virus as in about 10% of the patients, in absence of seroconversion, extended viremia has been observed (Nanda et al., 1995).

2.2.4. Classification of HEV:

Based on the morphological features, HEV was initially placed in the family Caliciviridae under the genus Hepesvirus (Miller, 1995). The structural organization of HEV with non-structural protein at the 5' end and structural protein at the 3' end of the genome and subgenomic mRNA detected in the tissue was superficially suggestive of a relationship with calciviruses (Tam et al., 1991). However, the order of genes is not
identical to that of a typical calcivirus. There have also been suggestions that HEV may be non-enveloped alpha like viruses because of homologous regions in the genome and sub-genomic transcripts (Purdy et al., 1993). Analysis of HEV nonstructural genes (RNA helicase and RNA-dependent RNA polymerase) and other positive stranded RNA viruses show that HEV forms a distinct phylogenetic group close to Rubella virus (Family Togaviridae) and beet necrotic yellow vein virus, a plant furovirus, rather than to members of Calciviridae (Fry et al., 1992). In addition, HEV codon usage resembles Rubella virus closely (Tam et al., 1991). It has been proposed that based on genetic relatedness and genome organization, the taxonomy of positive stranded RNA viruses should be reorganized to include rubella virus, beet necrotic yellow vein virus and HEV in separate but related families (Koonin and Dolja, 1993). New recommendations from the International Committee on Taxonomy of Viruses places HEV into a separate family called HEV-like viruses.

2.2.5. Pathology and histopathology:

HEV like other hepatitis viruses, causes liver damage, however it is not cytopathic to hepatocytes. Following acute liver injury, the outcome and clinical manifestations of hepatitis are determined by the host immune system. As the pathogenesis of HEV is not well understood, the primary site of replication, following entry of HEV in host via oral route, is predicted to be in the intestinal tract. It is still not clear how the virus reaches the liver however, HEV replicates in the cytoplasm of hepatocytes (Krawczynski and Bradley, 1989). In cynomolgus macaques inoculated with HEV intravenously, expression of HEV antigens in hepatocytes indicative of viral replication, is seen about day 7 post infection and begins to decrease after the maximum ALT activity. HEV antigen has been detected simultaneously in hepatocyte cytoplasm, bile and feces during the second or third week after inoculation and before and concurrently with the onset of ALT elevation and histopathologic changes in the liver. These findings suggest that HEV maybe released from hepatocytes into bile before the peak of morphologic changes in the liver, during the highly replicative initial phase of infection. The onset of ALT elevations and histopathologic changes in the liver generally correspond with the anti-HEV in the serum and with decreasing levels of HEV antigens in hepatocytes. In addition infiltrating lymphocytes in the liver have been found to have a Cytotoxic/suppression
immunophenotype. These findings suggest that liver injury maybe largely immune-mediated and that both a cell-mediated immune mechanism and humoral immunity are necessary for development of liver lesions. In a separate study, in situ hybridization and immunohistochemistry were done to detect the HEV genome and gene products in liver tissue of two patients with fulminant Hepatitis E (Lau et al., 1995). Both HEV RNA and proteins were detected exclusively in cytoplasm of hepatocytes.

Generally hepatitis E is self-limiting. Hence, liver samples are rarely available for histopathological analysis. The most extensive study of changes in the liver cells was carried out on biopsy specimen obtained during an outbreak in New Delhi in 1962. Two types of histopathologic changes were observed in the specimens from New Delhi outbreak: Cholestatic or “obstructive” type changes found in 58% of the patients and non-cholestatic or “classic” type change identified in 42% of the cases (Krawczynski et al., 1996; Bradley et al., 1994; Asher et al., 1990; Purcell and Ticehurst, 1988). The cholestatic type of changes is characterized by bile-stasis in canaliculi and gland-like formation of parenchyma cells. Degenerative changes of hepatocytes including acidophilic bodies and focal necrosis are less frequent. Lobules are inflammed and contain macrophages and polynuclear leukocytes as predominant population. Portal ducts are enlarged with inflammatory infiltrate consisting of polymorphonuclear leukocytes, lymphocytes and macrophages. Kupffer cells are prominent containing PAS-positive, diastase resistant cytoplasmic granules. Non-cholestatic or “classic type” of acute hepatitis was marked by focal necrosis and accumulation of mononuclear macrophages, lymphocytes and activated Kupffer cells. Hepatocytes appeared ballooned (swollen) and acidophilic degeneration of hepatocytes and acidophlic body formation is also observed. Experimentally infected primates also show acidophilic type of degeneration, focal necrosis with prominent accumulation of mononuclear macrophage and activated Kupffer cells (Krawczynski et al., 1996; Bradely et al., 1994). To summarize hepatitis E is associated with necrotic and degenerative changes in both humans and primates.

2.2.6. Animal models and in-vitro culture:

HEV transmission studies have been carried out on a number of primate species like chimpanzees, macaques, African green monkeys marmosets, owl monkeys and squirrel monkeys (Krawczynski, 1993; Tsarev et al., 1993; Ray et al., 1991; Ticehurst,
Review of Literature

1991; Panda et al., 1989; Bradley et al., 1987). However HEV infection in cynomologus macaques has been the most reproducible and widely used experimental model and has provided the most detailed data reflecting the pathogenetic events of the infection (Tsarev et al., 1994; Krawczynski, 1993; Tsarev et al., 1993; Longer et al., 1993; Bradley et al., 1987). In experimentally infected cynomolgus macaques, the average incubation period for acute hepatitis E is about 21 days; the acute phase of the disease extends from a couple of days before the peak of ALT activity to around 7 days after the peak of ALT activity (Krawczynski, 1993; Tsarev et al., 1993; Ticehurst et al., 1992; Balayan, 1990; Bradley et al., 1987). Similarly, single bile specimens obtained before the elevation of liver enzyme activity in the infected rhesus monkeys were found to be more frequently positive for HEV RNA by RT-PCR than were specimens obtained during the peak of transaminase levels (Nanda et al., 1994; Krawczynski and Bradley, 1989).

Pigs (Meng et al., 1997) and mice (Karetnyi et al., 1993) have also been reported to be susceptible to infection with HEV. Unlike primates, pigs develop jaundice showing that the disease is severe (Meng et al., 1998). Experimentally infected rats show histopathological changes and/or viral antigens in duodenum, spleen, mesenteric lymph nodes and peripheral blood mononuclear cells which suggests that these tissues might be infected, besides liver (Maneerat et al., 1996). It has been reported that anti-HEV antibodies are acquired naturally in primates and swine suggesting that these species have been exposed to HEV or a related agent, and that hepatitis E might be a zoonotic disease (Meng et al., 1999). The role of swine in HEV transmission is not clear, although domestic swine were reported to be susceptible to infection with a human HEV strain (Halbur et al., 2001). A novel virus designated swine hepatitis E virus, was identified in pigs (Meng et al., 1997). Swine HEV cross reacts with antibodies to human HEV capsid antigen. The putative capsid gene (ORF2) of swine HEV shared about 79-80% sequence homology at nucleotide level and 90-92% identity at the amino acid level with human HEV strains. The discovery of swine HEV not only has implications for HEV vaccine development, diagnosis, and biology, but also raises a potential public health concern for zoonosis or xenozoonosis following xenotransplantation with pig organs.

There have been reports of replication of HEV in FRhK (Rhesus kidney cells) cells for a Russian isolate of HEV (Kazachkov et al., 1992) and in 2BS diploid human
embryonic lung cells (Huang et al., 1992) and A549 human diploid cells (Huang et al., 1995) for a Chinese isolate. The Russian isolate was recovered by cocultivation of FRhK cells with primary cynomolgus monkey kidney cells, obtained during the acute phase of HEV infection in an experimentally infected cynomolgus. The virus did not produce any cytopathic effect (CPE), but chronic infection of the cells was detected by dot blot hybridization and immunofluorescence with convalescent sera from hepatitis E cases (Kazachkov et al., 1992). The Chinese strain 87A was recovered in 2BS diploid human embryonic lung cells and produced CPE (Huang et al., 1992). The virus was neutralized with acute phase sera but not convalescent sera from cases of hepatitis E (Kazachkov et al., 1992). Subsequently, the same 87A strain has also been propagated in human kidney and A549 cell lines (Huang et al., 1995; Li et al., 1995). A marked CPE appeared in the infected cells. The size of this virus was found to be 30 nm and HEV RNA was detected in culture supernatants by reverse transcriptase polymerase chain reaction (RT-PCR) suggesting the production and secretion of HEV particles in culture (Huang et al., 1995; Li et al., 1995).

Propagation of HEV in vitro in tissue culture system has been reported (Tam et al., 1996). This system utilizes hepatocytes isolated from cynomolgus macaques infected with Burma strain and these cells can be maintained in long-term culture for production of HEV. The yield is however, low. Strand-specific RT–PCR was used to detect positive sense genomic and negative sense replicative strands of HEV RNA in hepatocytes. Genomic RNA was also detected in culture medium, again providing evidence for secretion of HEV particles. Immune Electron Microscope (IEM) studies later confirmed this finding (Tam et al., 1996).

Recently, Emerson et al. (2004) have shown the replication of Hepatitis E virus (HEV) RNA in seven primate cell cultures transfected with in vitro transcripts of an infectious cDNA clone. As per their observation, Cell-to-cell spread did not occur in cell cultures, but rhesus monkeys inoculated with lysates of HEV-transfected PLC/PRF/5 and Huh-7 cells became infected with HEV. However, the replication efficiency was found to be inefficient since viral gene products could only be detected by immunofluorescence assay (Emerson et al., 2004).
2.2.7. Morphology of HEV/Genome Organization:

HEV was first identified by IEM in stool of patient with enterically transmitted non-A, non-B hepatitis. The virus is a non-enveloped, spherical virus with icosahedral symmetry and 27-34nm diameter (Kane et al., 1984; Balayan et al., 1983). The virus is unstable in harsh conditions like high salt concentration. Several geographically distinct isolates have been sequenced and have shown a high degree of nucleotide and amino-acid conservation (Panda et al., 1995; Yin et al., 1994; Aye et al., 1993; Tsarev et al., 1992; Huang et al., 1992; Aye et al., 1992; Tam et al., 1991). HEV has a positive polarity genome of ~7.5 kb with a short 5' and 3' non-coding region (Fig. 2.2.7).

![HEV Genome Organization Diagram](image)

Fig 2.2.7 Genome organization of HEV.

The coding region consists of three open-reading frames (ORFs). Of these ORF1 (~5kb) encodes a nonstructural protein, ORF2 (2kb) encodes the major viral capsid protein and ORF3 encodes a small protein of undefined function (Bradley and Purdy, 1994). All three proteins are expressed during infection as observed by antibodies against these proteins, found in infected human and experimental animals (Panda et al., 1995; Khudyakov et al., 1994).
Viral RNA also contains short 5' and 3'-untranslated regions (UTRs) of 26 and 68 nucleotides, respectively (Tam et al., 1996). These have the potential to fold into conserved stem-loop and hairpin-like structures. Such secondary structures are also found in a conserved 58-nucleotide region of ORF1 (Tam et al., 1996). These regions, together with a region showing homology to the junction sequence of Sindbis virus are postulated to be important for HEV RNA replication (Purdy et al., 1993).

**ORF1**: ORF1 encodes the nonstructural protein of HEV. Based on homology with protein motifs found in other positive-stranded viruses, ORF1 encodes a methyltransferase, a papain-like-cystein protease, a helicase and a RNA-dependent RNA polymerase (RdRp) (Koonin et al., 1992). It still remains to be determined if ORF1 encodes for a functionally active polyprotein or it undergoes processing to yield biochemically active individual units. Expression of complete ORF1 coding for 1,693 amino acid residues in vitro, in *Escherichia coli* and HepG2 cells, resulted in a large 186 kDa protein that was not processed automatically (Ansari et al., 2000). In another study, prolonged in vivo expression yielded an N-terminal 78 kDa protein and a C-terminal 107 kDa fragment (Ropp et al., 2000). Transfection of in vitro synthesized RNA, consisting of the complete HEV genome, to HepG2 cells resulted in synthesis of ORF1, ORF2 and ORF3 products as well as release of small amount of infectious virus into the culture medium (Panda et al., 2000). However, no 186 kDa protein was detected in pulse-chase experiments. Instead, region specific anti-sera precipitated smaller 35 to 40 kDa polypeptides.

The presence of a methyltransferase motif in ORF1 suggests that HEV may have a capped RNA genome, since this enzyme is generally responsible for methylating the 5'-terminal guanosine in the cap to produce the structure m\(^7\)G(5')ppp(5')X. It was shown that monoclonal antibodies to 2, 2, 7-trimethyl guanosine (m\(^3\)G) which cross-reacted with m\(^7\)G structures, could bind to HEV RNA (Lazizi et al., 1999). Moreover, the binding could be competitively inhibited by the cap analog m\(^7\)G (5')ppp(5')G, confirming that binding was specific for methylating cap structure (Lazizi et al., 1999). It was thus postulated that a virus-encoded enzyme might synthesize the cap. Recently, cDNA encoding putative methyltransferase was expressed in insect cells as a 110 kDa protein which was proteolytically processed to an 80 kDa protein (Magden et al., 2001). This
protein was shown to have virus specific methyltransferase and guanylytransferase activities similar to those of alphavirus replicase protein (Ahola and Kaariainen, 1995; Scheidel et al., 1989). Both activities are required in the capping of viral mRNAs (Ahola et al., 2000). Immunopurified P110 catalyzed transfer of a methyl group from S-adenosyl methionine (AdoMet) to GTP and GDP to yield m\(^7\)GTP or m\(^7\)GDP (Magden et al., 2001). This transfer of methyl group from AdoMet to GTP to yield m\(^7\)GTP has been shown to be virus specific enzyme activity (in alphaviruses) and results in covalent complex non-structural protein and m\(^7\)GMP, the last step in capping of mRNA (Ahola and Kaariainen, 1995).

In the positive-stranded RNA viruses the initiation of replication requires interaction of the 3' end of the genome with its RNA dependent RNA polymerase and possibly host derived co-factors, for the synthesis of negative strand replicative intermediate. Secondary structure predictions of the conserved 3' end of the infectious HEV genome were carried out to identify possible stem-loop structures necessary for RNA-protein interaction and the model was confirmed by structure probing experiments. Electrophoretic mobility shift assays showed specific binding assays of purified and refolded recombinant HEV RdRp protein to the 3' end of its RNA genome containing the poly A stretch (Agarwal et al., 2001).

**ORF2:** ORF2 has been suggested to be the major capsid protein of Hepatitis E virus. The 660 amino acid protein is positively charged at its N-terminal. About 10% of the protein consists of Arginine residues which impart it a high iso-electricpoint of ~10.3. This protein thus forms an ideal candidate to encapsidate the negatively charged viral RNA genome (Tam et al., 1991).

Based on comparative analysis of homologous domains and consensus sequence of proteins, certain predictions were made for ORF2. The hydropathy profile of ORF2 showed the presence of a highly hydrophobic stretch, rich in leucines, at its N-terminus followed by turn inducing and positively charged amino acids. Such an N-terminus stretch is characteristic of signal sequence of proteins translocated across cellular membranes (Schatz and Dobberstein, 1996). A search for potential N-linked glycosylation sites within consensus N-X-S/T revealed the presence of three such sites within ORF2: Asn-137, Asn-310 and Asn-562. These features were found in the deduced
amino acid sequence of ORF2 from all HEV isolates sequenced so far (Panda et al.,
1995; Yin et al., 1994; Aye et al., 1993; Tsarev et al., 1992; Huang et al, 1992; Aye et
al., 1992; Tam et al., 1991).

When expressed in COS-1 cells, ORF2 showed two forms, one of predicted size of
\(-74 \text{ kDa}\) and other of larger size of \(-88 \text{ kDa}\). The 88 kDa form was shown to be
glycosylated ORF2 (gpORF2) through tunicamycin inhibition and endoglycosidase H
digestion experiments (Zafarullah et al., 1999). Tunicamycin is an inhibitor of N-linked
glycosylation and endoglycosidase H specifically cleaves high-mannose residues from
glycoproteins, both modifications known to occur within the endoplasmic-riculum(ER)
(Tarentino et al., 1989). Further analysis with drugs Brefeldin A and monensin indicated
that the ER may be the major site of accumulation of gpORF2 (Zafarullah et al., 1999).

The sites of glycosylation on ORF2 were analyzed through site-directed
mutagenesis. Asparagine residues within the three conserved N-X-S/T sequences at
amino acid positions 137, 310 and 562 were changed to alanine. The single, double and
triple mutants were then analyzed for glycosylation. While all the single mutants, ORF2
[137], ORF2 [310] and ORF2 [562] showed the presence of gpORF2, double mutants
ORF2 [137, 310] and ORF2 [310, 562] showed only non-glycosylated ORF2 whereas the
double mutant ORF2 [137, 510] expressed protein with both glycosylated and non-
glycosylated forms. Predictably, triple mutant ORF2 [137, 310, 562] only expressed non-
glycosylated form of ORF2. Theses results suggest that Asn-310 is the major site, though
not the only site for glycosylation on ORF2 (Zafarullah et al., 1999).

Pulse-chase and deletion mutations have shown that ORF2 carries an N-terminal
signal sequence which is responsible for its transit across the ER, where it gets
glycosylated (Zafarullah et al., 1999). This viral protein was also found to be expressed
on the cell-surface. It was further shown that translocation into the ER, but not
glycosylation is prerequisite for cell surface expression of ORF2 (Zafarullah et al., 1999).
Recently it has been suggested that non-glycosylated cystolic form of ORF2, which
accumulates in mammalian cells and not the membrane associated glycosylated form,
takes part in the capsid assembly (Torressi et al., 1999). It has also been suggested that
the non-glycosylated form is more stable than the glycosylated form. The precise role of
the glycosylated form of ORF2 in the hepatocyte remains unclear.
Expression of ORF2 has been carried out through various expression systems including E.coli, baculovirus, vaccinia virus and animal cells. When expressed through the baculovirus expression system, ORF2 was shown to assemble into virus-like particles (VLPs) which remained cell-associated (Li et al., 1997). The primary translational product with a molecular mass of 58 kDa was produced in the insect cells and further processed to 50 kDa form. This form was found to be secreted into the culture medium. Electron microscopic studies of the culture medium revealed that the 50 kDa form self-assembled to form VLPs (Li et al., 1997). Antigen derived from the insect cell culture systems has the potential to be a good diagnostic reagent.

**ORF3:** The ORF3 of HEV encodes a small protein of ~13.5 kDa designated as ORF3 protein. A number of studies have established that ORF3 carries an immunodominant epitope at its C-terminal end, antibodies to which are universally present in infected humans and animals (Panda et al., 1995; Khudyakov et al., 1994; Li et al., 1994; Khudyakov et al., 1993). However, no information exists regarding the possible role(s) of ORF3 protein during the viral life cycle and disease pathogenesis. An analysis of the ORF3 protein showed two highly hydrophobic domains in its N-terminal half (Fig 1.6). Domain 1, encompassing amino acids 15 to 31 is cysteine-rich, while Domain II encompassing amino acids 38 to 62 is rich in aliphatic and aromatic hydrophobic residues. The C-terminal half of the protein is rich in proline residues with two regions that have the propensity to fold into polyproline helices. Further, the extreme C-terminus of ORF3, comprising amino acids 92 to 123, is highly immunodominant (Fig. 1.6) as evident from western blotting and peptide blocking assays (Panda et al., 1995). Though its primary amino acid sequence contains a N-terminal hydrophobic region suggestive of a signal sequence and N linked glycosylation motifs, no experimental proof exist for the same.

In our laboratory, ORF3 was expressed by transient transfection in COS-1 and Huh-7 cells and has been experimentally proven to be a phosphoprotein, which is modified at a serine residue. Deletion and site-directed mutagenesis established Ser-80 as the phosphorylation site. Ser-80 is present in all the HEV isolates including swine HEV (Panda et al., 1995; Yin et al., 1994; Aye et al., 1993; Tsarev et al., 1992; Aye et al., 1992; Tam et al., 1991), but is absent from the most divergent Mexican isolate (Huang et
This residue is present within a conserved primary sequence that showed consensus sites for phosphorylation by proline directed kinases such as p34\(^{cdk2}\) kinase (CDK1) and mitogen-activated protein kinase (MAPK). In vitro experiments with hexahistidine-tagged ORF3 expressed either in E. coli or in COS-1 cells showed an efficient phosphorylation profile with MAPK, which was identical to that observed in vivo. In its primary sequence, ORF3 possesses two highly hydrophobic N-terminal domains. Subcellular fractionation and deletion analysis proved that ORF3 protein associates with the cytoskeletal fraction through the hydrophobic domain I (amino acid residues 1 to 32) (Zafarullah et al., 1997).

Recently, Korkaya et al., have used in vitro binding assays to show that ORF3 binds to a number of cellular proteins such as Src, Hck, Fyn, the p85\(^{\alpha}\) subunit of phospholipase C\(\gamma\), and the adaptor protein Grb2. A yeast two-hybrid assay was used to confirm the ORF3-Grb2 interaction. The binding involves the proline-rich region in ORF3 and the src homology 3 (SH3) domains of Grb2. Competition assays and computer-assisted modeling was used to evaluate the binding surfaces and interaction energies of the ORF3-SH3 complex. In ORF3-expressing cells pp60\(^{src}\) was found to associate with an 80 kDa protein, but no activation of the src kinase was observed in these cells. However, there was increased activity and nuclear localization of extracellular signal-regulated kinase (ERK) in ORF3-expressing cells. These studies suggested that ORF3 is a viral regulatory protein involved in the modulation of cell signaling (Korkaya et al., 2001). Recently, it has been proved that the mechanism of ORF3 mediated MAPK activation underlies its ability to directly bind and inhibit the activity of MAPK phosphatase Pyst1 (Kar-roy et al., 2004).

Further, by screening a human liver cDNA library using ORF3 protein as bait, our laboratory has identified alpha 1 microglobulin and bikunin protein as ORF3 binding proteins. Experiments done in Huh7 human hepatoma cells indicate that ORF3 protein expedites the antergrade transport of alpha 1 microglobulin protein from the golgi region (Tyagi et al., 2004). However, physiological significance of all these phenomena during the natural course of infection remains to be understood.

It is evident that the studies on biology of HEV have been hampered because it has not been successfully cultured. Two decades later, HEV still remains a mystery with
reference to its replication strategy. A replication model of HEV replication has been proposed (Reyes et al., 1993), more on the basis of similarities to other positive-stranded RNA viruses, rather than direct experimental evidence. On entry into a permissive cell, viral genomic RNA is translated in the cytoplasm of the infected cell to produce non-structural ORF1 encoded polyprotein. Presence of viral replicase (RdRp) suggests replication of genomic positive-strand into a negative-strand replication intermediate. Such replicative intermediates of sub-genomic (~3.7 kb and ~2 kb) and genomic (~7.5 kb) RNAs have been shown in experimentally infected cynomolgus macaques (Tam et al., 1991). HEV positive and negative stranded RNA has also been shown in liver of rhesus macaques' model (Nanda et al., 1994). Based on homology to junction sequences present in Sindbis alphavirus, these replicative intermediates are postulated to act as templates for synthesis of additional copies of the genomic positive-strand as well as sub-genomic positive strands (Li et al., 1994). The sub-genomic positive-strand HEV RNA is postulated to be translated into structural protein(s) at late stage of viral replication. These proteins encapsidate the viral genomic RNA producing progeny virions.

It is abundantly clear that the biology of HEV warrants further investigation especially with regards to its replication strategy and pathogenesis. For understanding a disease, it is prerequisite that the life cycle of the organism causing it is understood. However, as disease causing pathogens, viruses cannot be studied in isolation without studying their life cycle inside the host. In the absence of a model infection system, mechanism of HEV pathogenesis has been restricted to expressing individual gene products inside cell lines and characterizing their property. Although, it may not be relevant to extrapolate the data to a natural infection condition, however, the data obtained can be compared to similar conditions demonstrated in other viruses and thus may be helpful in understanding the biology of HEV in an indirect way. Keeping this in mind, we have characterized the functional property of major capsid (ORF2) protein in a human liver cell line. A brief review of some of the characteristic properties demonstrated by capsid proteins of other RNA viruses is described below.
2.3 Genomic RNA binding by the capsid protein in positive strand RNA viruses.

Interaction of genomic RNA with the capsid protein is commonly observed among RNA viruses. Although, the nature and region of interaction defines the functional significance of that interaction; in general, capsid protein and genomic RNA interaction serves two crucial functions in the viral life cycle.

(I) Packaging of the genome into the viral capsid.

(II) Efficient replication and translation of the viral RNA.

(I) Role of capsid protein and genomic RNA interaction in packaging of genome into the viral capsid.

Once the replication is complete and abundant amount of viral RNA and proteins have been synthesized, it is time for the new viral gene products to assemble into virus particle and bud off from the infected cell. In positive strand RNA viruses, capsid protein usually forms oligomer which binds to the genomic RNA and packages it inside the capsid so that after exit from the infected cell, the viral genome is protected from the harsh conditions of external environment. In the genomic RNA, the genome incorporation signal is normally present at the 5' end so that only full length genomic RNA (not the subgenomic RNAs) is packaged into the capsid. On the other hand, the RNA binding region of the capsid protein is reach with basic amino acids like arginine. Hence, a strong charge based interaction occur between the positively charged cluster of amino acid residues and the negatively charged nucleic acid (RNA) leading to formation of the RNA-protein complex. Subsequently, the linear genome is folded into a compact structure so that it can be completely concealed inside the capsid.

Essential role played by the basic amino acid rich region of capsid protein in genome packaging and virus assembly has been investigated in detail in Cowpea chlorotic mottle virus (CCMV). When the arginine rich motif was mutated in the capsid protein of CCMV, packaging of the genomic RNA into the capsid and assembly of progeny virus particle was found to be significantly inhibited (Annamalai et al., 2005). Similar arginine rich motif is present in the capsid protein of most RNA viruses. Thus,
the above observations prove that interaction between capsid protein and genomic RNA is physiologically essential for efficient packaging of the genome and assembly of the progeny virus.

Although physical interaction between genomic RNA and capsid protein can be readily detected in vitro in a wide group of viruses, however additional regulatory steps are present in some cases that coordinate packaging of genome to replication and translation of viral proteins. An elegant study done in Polio virus demonstrates that there exist a functional coupling between replication and packaging of polio virus RNA (Nugent et al., 1999).

Further, post translational modification of capsid protein also regulates its association with genomic RNA and virus assembly. For example, in Rubella virus, only the dephosphorylated capsid protein can interact with RNA and in conditions where virus assembly was inhibited, capsid protein was found to be hyperphosphorylated. On the other hand, this hyperphosphorylated form was observed to enhance the genome replication efficiency of the virus (Law et al., 2003). Thus, at a physiological level, capsid protein interaction with genomic RNA is temporally regulated by post translational modification. However, the above fact might be a specific property of the Rubella virus only. Nonetheless, all the above examples support the hypothesis that interaction between genomic RNA and capsid protein is essential for efficient packaging of the genome.

(2) Role of capsid protein and 3' UTR of genomic RNA interaction in replication and translation of viral genome.

Circularization of the genome of RNA viruses plays an essential role in the replication and translation of viral RNAs. For many positive-strand RNA viruses, circularization seems to be mediated by protein bridges between the 5' and 3' termini of the viral RNA. On entry into the host cell, the genome of positive-strand RNA viruses has to serve as messenger for the translation of viral proteins including replication factors. A number of viral RNAs resemble cellular mRNAs by having a 5' -cap structure (m7GpppN) and 3'- poly(A) tail. In cellular mRNAs, these structures permit the formation of a protein bridge by interactions between eIF4G and eIF4E from the cap-
binding complex of translation initiation factors and the poly(A) binding protein (PABP) associated with the poly(A) tail (Gallie, 1991, 1998). Picorna viruses do not have a cap structure, but the internal ribosome entry site in the 5' untranslated region (UTR) of the RNA binds initiation factors and a poly(rC) binding protein that mediate the interaction with PABP. Circularization of Poliovirus RNA play a role both in translation and replication of the viral RNA (Gamarnik and Andino, 2000; Herold and Andino, 2001). On the other hand, genomic RNAs of many viruses lack a poly(A) tail, and elements in the 5' and/or 3' UTR are believed to compensate for the absence of a poly(A) tail (Gallie and Kobayashi, 1994; Wells et al., 1998; Qu and Morris, 2000; Timmer et al., 1993; Lipzig et al., 2001). For example, efficient translation of AMV RNAs in plant cells depends on the ability of capsid protein to bind to the 3' end of its own messenger RNA in cis or in trans. Thus, the binding of CP to the 3' termini of AMV RNAs may be functionally equivalent to the binding of PABP to the poly(A) tail of cellular mRNAs (Neelman et al., 2001). Similarly, study done by Neelman et al. proves that efficient translation of alfa mosaic virus RNAs requires the binding of coat protein dimers to the 3' termini of the viral RNAs. The 3' untranslated region (UTR) of the AMV subgenomic CP messenger RNA 4 contains at least two CP binding sites. They found that a CP binding site in the 3'-terminal 112 nucleotides of RNA 4 was required for efficient translation of the RNA whereas an upstream binding site was not. Binding of CP to the AMV 3' UTR induces a conformational change of the RNA but this change alone was not sufficient to stimulate translation. Further, CP mutant R17A was unable to bind to the 3' UTR and translation in vivo of RNA 4 encoding this mutant occurred at undetectable levels (Neelman et al., 2004).

Thus, the above examples demonstrate that capsid protein binding to genomic RNA serves multiple purposes during the viral life cycle. However, the exact mechanism varies from virus to virus and it may be possible that some other viral gene products cooperate with each other to synergize the process in an in vivo context.

Besides their role in assembly of the virus, capsid proteins of many viruses have been shown to significantly perturb the cellular homeostasis. Since capsid proteins are generally synthesized as glycoprotein, they have a tendency to accumulate in the endoplasmic reticulum (ER) resulting in ER stress.
2.4. Endoplasmic Reticulum (ER) Stress

The endoplasmic reticulum (ER) is a principal site for folding and maturation of transmembrane, secretory and ER resident proteins. Perturbations that alter ER homeostasis can lead to ER stress. ER stress has been broadly categorized into three different types: 1) the presence of mis or unfolded proteins in the organelle, 2) the starvation of cholesterol and 3) the overloading of the ER with correctly folded proteins. Each stress situation triggers a unique cellular response, using various signal transduction pathways to induce specific response.

2.4.1 The unfolded protein response UPR

The unfolded protein response (UPR), as the name suggests, is activated by the presence of mis- or unfolded proteins in the ER (table 1) (Pahl, 1999). Of the three pathways, it is the only one conserved, at least in part, between humans and yeast. The more recently discovered sterol regulatory element binding protein (SREBP) pathway regulates cholesterol synthesis in the ER membrane in response to sterol plasma levels. In both of these pathways, novel mechanisms of transcription factor activation have been described, which may prove to be paradigms for processes widely used in other pathways. The most recently described ER-nuclear signal transduction pathway, ER-overload response (EOR), a stress triggered by congestion of the organelle with too many proteins, activates a well-characterized transcription factor, nuclear factor κB (NFκB). However, it uses a novel pathway in the process.
Fig 2.4.1 Schematic representation of UPR signaling.

TABLE 1. Conditions that induce the unfolded protein response

<table>
<thead>
<tr>
<th>Drugs perturbing ER function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tunicamycin</td>
</tr>
<tr>
<td>2-Deoxyglucose</td>
</tr>
<tr>
<td>Brefeldin A</td>
</tr>
<tr>
<td>Castanospermine</td>
</tr>
<tr>
<td>Glucosamine</td>
</tr>
<tr>
<td>Thapsigargin</td>
</tr>
<tr>
<td>ALF42</td>
</tr>
<tr>
<td>Glucose starvation</td>
</tr>
<tr>
<td>Reducing agents</td>
</tr>
<tr>
<td>2-Mercaptoethanol</td>
</tr>
<tr>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>Heavy metals</td>
</tr>
<tr>
<td>Cobalt</td>
</tr>
<tr>
<td>Nickel</td>
</tr>
<tr>
<td>Overexpression of mutant proteins that cannot fold correctly</td>
</tr>
<tr>
<td>Calcium ionophores</td>
</tr>
<tr>
<td>A-23187</td>
</tr>
<tr>
<td>Ionomycin</td>
</tr>
</tbody>
</table>
Viral gene products (e.g. HCV core gene product)

The UPR is an integrated intracellular signaling pathway that transmits information about the protein folding status in the ER lumen to the cytoplasm and the nucleus (Fig 2.4.1). The UPR includes transcriptional induction of UPR genes (red arrows), translational attenuation of global protein synthesis (black arrows) and ER-associated degradation (ERAD) (green arrows). These divergent outputs provide adaptive responses for survival. If the protein-folding defect is not corrected, cells undergo apoptosis (light-blue arrows). The three major transducers of the UPR are PERK, IRE1 and ATF6. PERK is an ER transmembrane protein kinase that phosphorylates the α subunit of translation initiation factor 2 (eIF2α) in response to ER stress. Phosphorylation of eIF2α reduces the formation of translation initiation complexes, which leads to reduced recognition of AUG initiation codons and therefore general translational attenuation. This translational control provides an efficient mechanism to reduce the number of unfolded proteins in the ER. Paradoxically, the translation of selective mRNAs that have a lower requirement for eIF2 and the translation initiation complex is enhanced, such as the mRNA encoding the activating transcription factor ATF4. GADD34 transcription is induced by the UPR through ATF4, and the protein product recruits protein phosphatase 1 (PP1) to dephosphorylate eIF2α-P and reverse the translational attenuation. ATF6 is an ER transmembrane activating transcription factor. Upon ER stress, ATF6α and ATF6β transit to the golgi compartment where they are cleaved by S1P and S2P proteases to yield a cytosolic fragment. The free ATF6 fragment migrates to the nucleus to activate transcription. IRE1 is an ER transmembrane glycoprotein and it contains both kinase and RNase activities in the cytoplasmic domain. ER stress leads to its autophosphorylation and the subsequent activation of its RNase activity. The substrate of IRE1 in mammals, XBP1 mRNA, encodes a basic leucine -zipper containing transcription factor. Splicing of XBP1mRNA is initiated by the RNase activity of IRE1 to generate mature XBP1mRNA. Whereas the ATF6 and PERK pathways are not conserved in lower eukaryotes, the IRE1 signaling pathway is conserved in all known eukaryotic cells. The signalling from downstream effectors of IRE1, PERK and ATF6 merges in the nucleus to activate transcription of UPR target genes. The mammalian ER stress element
Review of Literature

(ERSE) is present in the promoter regions of many, but not all, UPR target genes. XBP1, ATF6 and the CAAT-binding factor (CBF), all of which bind to ERSE, along with ATF4, activate transcriptional induction of target genes. ATF6 also induces XBP1 transcription, providing a positive feedback for the UPR. In particular, upregulation of molecular chaperones and folding catalysts increases the folding capacity of the ER, providing a protective effect for cell survival. In addition, activated Ire1p in yeast induces transcription of genes, such as \textit{INO1}, that mediate phospholipids biosynthesis to increase the ER volume.

The UPR also induces transcription of genes encoding proteins that remodel the secretory pathway to decrease the concentration of UPS. BiP, the ER chaperone, is the master regulator of the activation of the three proximal ER stress transducers – IRE1, PERK and ATF6. All transducers contain a luminal domain that interacts with BiP. Under normal conditions, BiP serves as a negative regulator of IRE1, PERK and ATF6 activation. Upon ER stress, BiP binds to UPS, thereby allowing BiP release from the transducers. BiP release from IRE1 and PERK permits their homodimerization and activation. BiP release from ATF6 permits its transport to the golgi compartment for regulated intramembrane proteolysis. This BiP regulated activation provides a direct mechanism to sense the folding capacity of the ER. Prolonged UPR activation leads to apoptotic cell death, in which IRE1 serves a proapoptotic function. Activated IRE1 recruits Jun N-terminal inhibitory kinase (JNK) and TRAF2 to activate apoptosis-signaling kinase 1(ASK1), which in turn activates JNK and mitochondria/Apaf1-dependent caspases. Procaspare-12 (pCSP-12) is an ER-associated proximal effector of apoptosis. TRAF2 release from pCSP-12 permits the clustering and activation of CSP-12. Activated CSP-12 activates CSP-9, which in turn activates CSP-3, leading to apoptosis. Upon ER stress, activated CSP-7 can cleave pCSP-12 to generate active CSP-12. In addition, UPR activation induces CHOP/GADD153 expression through the PERK and ATF4 pathways. CHOP is a pro-apoptotic transcription factor that potentiates apoptosis. Finally, in response to prolonged ER stress, attenuation of cyclin D1 translation through PERK leads to cell cycle arrest during G1 phase. This provides an ER checkpoint to prevent cells from progressing through the cell cycle (Chuan, 2003).
2.4.2 The ER associated degradation (ERAD)

When proteins remain misfolded for a prolonged period in the ER, they are destined for degradation by host quality control machinery. This alleviates ER stress and normal metabolism is restored. Although some misfolded proteins are transported to and degraded in the lysosome (Hong et al., 1996; Chang and Fink, 1995) most of them are degraded in the cytosol by the ubiquitin–proteasome system after direct retrotranslocation from the ER. This process has been termed as ER associated degradation (ERAD) (Tsai et al., 2002).

The retro-translocation pathway is divided into four different steps: substrate recognition and targeting to the retro-translocation machinery; protein transport across the ER membrane; release of the substrate from the ER membrane into the cytosol; and degradation of the substrate. It is unlikely that all substrates actually use the same pathway. However, the general mechanism is schematically in the ensuing diagram (Fig 2.4.2). As illustrated in the figure, in step 1, misfolded lumenal proteins are recognized and unfolded by endoplasmic reticulum (ER) chaperones and targeted to the translocation channel. Bacterial toxins probably use the same mechanism for unfolding. Misfolded membrane proteins might be targeted to the channel by diffusion in the plane of the membrane (dashed arrow). In step 2, polypeptides are moved through the channel. Polyubiquitylation (Ub) occurs when the polypeptide chain becomes accessible in the cytosol. In step 3, polyubiquitylated polypeptides are released from the ER membrane into the cytosol. In step 4, polypeptides are degraded by the proteasome.
Fig 2.4.2 Schematic representation of ERAD signaling.

2.4.3 ER-Nucleus Signaling by SREBP

Cholesterol, an essential cell membrane component, is made available to cells in two ways. One is the uptake of cholesterol in food. After resorption in the small intestine, Cholesterol is transported in the bloodstream bound in low-density lipoproteins (LDL). The LDL receptors on cell surfaces bind and internalize LDL from the plasma. Alternatively, almost all cell types can synthesize cholesterol de novo (Brown et al., 1986; Goldstein, 1990). Both cholesterol uptake and synthesis are subject to feedback repression. Transcription of the genes encoding the LDL receptor, lipoprotein lipase, as well as cholesterol and fatty acid biosynthetic enzymes are inhibited by high intracellular sterol levels and induced upon sterol depletion. Regulation is mediated at the level of transcription. This requires a 10-bp DNA sequence, the sterol regulatory element 1 (SRE-1), in the promoters of sterol-regulated genes (Fig 2.4.3) (Smith et al., 1990). Basic-helix-loop-helix-leucine zipper (bHLH-LZ) protein, SREBP-1 and SREBP-2 have been shown to activate transcription via the SRE-1 (Yokoyama et al., 1993). Both SREBP-1 and the homologous SREBP-2 (Hua et al., 1993) are synthesized as large precursor molecules of; 1,150 amino acids, which are inserted into the ER/nuclear membrane (Wang et al., 1994).
The NH2 terminus is cytoplasmic and contains the bHLH-LZ motif. The middle of the protein contains two transmembrane domains, which are inserted into the ER or nuclear membrane, causing the intervening 30 amino acids to form a "luminal loop" that protrudes both into the ER and into the nuclear envelope (Hua et al., 1995). The COOH terminus again faces the cytoplasm (Sato et al., 1994). Upon sterol depletion, SREBP precursors are cleaved by two sequential proteolytic steps (Sakai et al., 1996), releasing a 500-aminoacid fragment that contains the basic DNA-binding domains as well as the leucine zipper dimerization motif. SREBP cleavage is mediated by SREBP cleavage-activating protein (SCAP). Cleaved SREBP, which contains the bHLH-LZ motif and apparently some of the transmembrane domain, translocates to the nucleus where it activates transcription of 3 groups of target genes: 1) genes encoding enzymes involved in cholesterol biosynthesis, such as HMG-CoA synthase; 2) genes encoding proteins involved in cholesterol and fatty acid uptake from the plasma, for example, the LDL receptor and lipoprotein lipase; and 3) genes encoding proteins mediating fatty acid synthesis, for example, the fatty acid synthase (FAS) gene. The importance of this signaling pathway becomes apparent in patients with Niemann-Pick disease.

Fig 2.4.3 Schematic representation of SREBP signaling.
2.4.4 The ER overload response (EOR)

Overexpression of various membrane proteins, which accumulate in the ER (Table 2), activates a signaling cascade between ER and nucleus, which is termed as ER overload response. Enhanced activity of NF-κB (Meyer et al., 1992; Pahl et al., 1995; Pahl et al., 1996) transcription factor is a hallmark of ER overload response. Also, perturbation of ER function with a variety of agents that either inhibit N-glycosylation (tunicamycin and 2-deoxyglucose) or protein transport by causing the formation of a mixed ER-Golgi compartment (brefeldin A) strongly induce the transcription factor (Pahl et al., 1995b) NF-κB. This demonstrates that a number of the stimuli that activate NF-κB (i.e. EOR) also induce the UPR (see Table 1). For example, both NF-κB and the UPR are activated by overexpression of the μ-heavy chain, tunicamycin, 2-deoxyglucose, brefeldin A, or thapsigargin (Pahl et al., 1995; Shamu et al., 1994) (Tables 1 and 2). However, several pharmacological and biological inducers distinguish between NF-κB induction and UPR activation, thus supporting the hypothesis that NF-κB functions in a novel signal transduction pathway between the ER and the nucleus.

<table>
<thead>
<tr>
<th>TABLE 2. ER stress conditions that activate nuclear factor κB</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER overload by protein overexpression</td>
</tr>
<tr>
<td>Hepatitis B virus truncated middle HB surface antigen</td>
</tr>
<tr>
<td>Influenza virus hemagglutinin</td>
</tr>
<tr>
<td>Immunoglobulin m-heavy chain</td>
</tr>
<tr>
<td>MHC class I</td>
</tr>
<tr>
<td>Adenovirus E3/19K</td>
</tr>
<tr>
<td>EPO receptor</td>
</tr>
<tr>
<td>Drugs perturbing ER function</td>
</tr>
<tr>
<td>Tunicamycin</td>
</tr>
<tr>
<td>2-Deoxyglucose</td>
</tr>
<tr>
<td>Monensin</td>
</tr>
<tr>
<td>Brefeldin A</td>
</tr>
<tr>
<td>Thapsigargin</td>
</tr>
<tr>
<td>Cyclopiazonic acid</td>
</tr>
<tr>
<td>MHC, major histocompatibility</td>
</tr>
</tbody>
</table>

34
The ER-overload hypothesis has been investigated using the adenovirus E3/19K protein as a model (Pahl et al., 1996). Wild-type E3/19K resides in the ER, where it binds to MHC class I molecules, thereby preventing their transport to the cell surface. The viral protein possesses a COOH-terminal retention signal sequence, which causes the protein to be continuously retrieved to the ER from post-ER compartments. As expected, expression of even small amounts of E3/19K strongly activates NF-κB. Because the sequence requirements of E3/19K for MHC class I binding and ER retention are precisely known, the NF-κB activating ER signal could be investigated using point mutants that abolish either property. Two point mutants that no longer bind MHC class I molecules activate NF-κB as effectively as the wild type protein. Thus the interaction between E3/19K and another protein, MHC class I, is not necessary for NF-κB activation. Titration experiments confirmed that the NF-κB-activating effects of overexpressing E3/19K and MHC class I are additive, not synergistic. However, there is a stringent requirement of ER retention for NF-κB activation. Two mutant proteins, which are equally well expressed as the wild type, but escape ER retention and are expressed on the cell surface, no longer activate the transcription factor even when highly
overexpressed. Hence, the sole determinant for the NF-κB-activating effects of E3/19K is its ER retention. The di-lysine ER retrieval motif in the cytoplasmic tail of E3/19K was not required; since proteins that carry mutations in this motif but nonetheless accumulate in the ER were equally potent in NFκB activation. These studies suggest that the NF-κB activating signal is generated by the accumulation of proteins within the ER membrane, a process we refer to as ER overload.

Physiological Role of the EOR

(A) THE ANTIVIRAL RESPONSE.

Upon viral infection, cells become programmed to produce large amounts of viral capsid proteins that are processed through the ER. This may elicit an EOR just as seen upon expression of individual viral membrane proteins using transient transfection methods. Transient expression of three unrelated viral proteins, influenza hemagglutinin (Pahl et al., 1995), hepatitis B virus MHBS (Meyer et al., 1992), and adenovirus E3/19K (Pahl et al., 1996), to levels barely detected by immunofluorescence staining activates NFκB by an EOR (Table 2). Because the NF-κB-activating stimulus appears to simply be the accumulation of membrane proteins, we expect that a large variety of membrane proteins from unrelated viruses can elicit this signal. Several NFκB target genes encode important antiviral defense proteins, for example, the cytokine β-interferon and proteins involved in viral peptide presentation to T cells, such as the proteasome subunit LMP2, the TAP1 peptide transporter, MHC class I molecules, and β2-microglobulin (Baeuerle et al., 1994). By inducing their expression via NFκB, the EOR may elicit a fast, nonspecific and therefore broad antiviral response. Direct evidence for this role of the EOR remains to be obtained.

(B) DISEASES OF PROTEIN FOLDING.

Many diseases are caused by mutations that cause misfolding of proteins such as a1-Antitrypsin deficiency, Alzheimer's disease, Cystic fibrosis etc. These mutant proteins have been shown to accumulate in the ER. Interestingly, several of these conditions are
accompanied by inflammation, of which the pathomechanism remains unknown. It has been hypothesized that the accumulation of wild-type or mutant proteins in the ER may lead to NF-κB activation. The activated transcription factor would lead to expression of its target genes, many of which encode inflammatory cytokines such as interleukin (IL)-1 and TNF-α and chemokines such as IL-8, monocyte chem. attractant protein-1 (MCP-1), monocyte inhibitoryprotein-1a (MIP-1), and RANTES. Secretion of these proinflammatory mediators would result in neutrophil, macrophage, and ultimately T-cell recruitment. These cells would become activated and in turn themselves release inflammatory cytokines, stimulating and sustaining the inflammatory process. This model could explain how the accumulation of a misfolded or a wild-type protein may lead to inflammation.

2.4.5 The UPR and viral pathogenesis

The two major mediators of the IFN-induced arm of the innate immune response are evolutionarily related to IRE1 and PERK. The kinase/endoribonuclease domain of IRE1 is homologous to RNaseL, and the protein kinase domain of PERK is related to the double-stranded RNA–activated (dsRNA-activated) eIF2α protein kinase PKR. RNaseL and PKR mediate the IFN induced antiviral response of the host, which is required to limit viral protein synthesis and pathogenesis. As part of the innate immune response to viral infection, RNaseL and PKR are activated by dsRNAs produced as intermediates in viral replication. In contrast to activation by dsRNA, IRE1 and PERK are activated by ER stress, which can be induced by high-level viral glycoprotein expression. All enveloped viruses produce excess glycoproteins that could elicit PERK andIRE1 activation to meet the need for increased folding and secretory capacity. More studies will be required to elucidate the role of the UPR in various viral diseases.

Among the positive-stranded RNA viruses causing hepatitis, Hepatitis C virus (HCV) demonstrates a classic example of virus induced UPR activation. HCV encodes a single polyprotein. Polyprotein cleavage generates at least ten polypeptides, including two glycoproteins, E1 and E2. A large amount of E1 forms disulfide–cross-linked aggregates with E2 in the ER (Sharp, 1971). Since the accumulation of misfolded α1-PI
elicits UPR activation, with subsequent hepatocyte death and hepatocellular carcinoma, it is possible that the aggregated E1/E2 complexes in the HCV-infected hepatocyte also contribute to hepatitis and hepatocellular carcinoma.

2.4.6 Viral exploitation of ERAD pathway

Some viral gene products have been shown to exploit the ERAD pathway towards their benefit. The expression of certain viral proteins leads to the selective degradation of normal cellular proteins that are required for the immune defense of the host. For example, the expression of either US2 or US11 — two proteins of the human cytomegalovirus (HCMV), targets the newly synthesized major histocompatibility (MHC) class I heavy chain for degradation (Kopito, 1997; Wiertz et al., 1996). This single-spanning protein is recognized by the virally encoded proteins, exported from the ER into the cytosol and degraded by the ubiquitin–proteasome system. The Vpu protein of the human immunodeficiency virus (HIV) also triggers ER retention and degradation of the CD4 receptor, and the MK3 protein of the murine γ-herpesvirus-68 leads to the selective degradation of MHC class I heavy chains by the ubiquitin–proteasome system (Boname et al., 2001). The same pathway of retro-translocation might also be used by certain toxins, including cholera toxin, enterotoxin and ricin (Sandvig and van Deurs, 2000; Lord and Roberts, 1998). Cholera toxin, for example, is taken up by endocytosis and transported backwards through the secretory pathway until it arrives in the ER lumen. A fragment of the toxin is then released from the rest of the molecule and is retrotranslocated across the ER membrane into the cytosol (Lencer et al., 1999; Schmitz et al., 2000). In contrast to the other substrates discussed above, the toxin fragment is not degraded, at least not completely. It becomes an active enzyme that modifies a heterotrimeric G protein, which, in turn, leads to the opening of a chloride channel with consequent massive chloride and water secretion.

2.5. Viral modulation of NF-κB signaling pathway
The main function of the immune system is to recognize and deal with pathogens. Since NF-κB proteins are important components of signaling pathways that regulate the development of immune responses and associated effector functions; thus they play a major role during infection by pathogens. NF-κB is required for resistance to a variety of viral, bacterial, and parasitic infections, and this is clearly illustrated by gene deletion studies (Caamano and Hunter, 2002). On the other hand, pathogens too have evolved strategies to modulate the cellular NF-κB activity towards their benefit.

2.5.1. Components of the NF-κB signaling complex

In mammals the NF-κB family of transcription factors contains five members: NF-κB1 (p105/p50), NF-κB2 (p100/p52), RelA (p65), RelB, and c-Rel. NF-κB1 and NF-κB2 are synthesized as large polypeptides that are post translationally cleaved to generate the DNA binding subunits p50 and p52, respectively. Members of the NF-κB family are characterized by the presence of a Rel homology domain, which contains a nuclear localization sequence and is involved in sequence-specific DNA binding, dimerization, and interaction with the inhibitory IκB proteins (Gewirtz et al., 2001). The NF-κB members dimerize to form homo- or heterodimers, which are associated with specific responses to different stimuli and differential effects on transcription. NF-κB1 (p50) and NF-κB2 (p52) lack transcriptional activation domains, and their homo dimers are thought to act as repressors. In contrast, Rel-A, Rel-B, and c-Rel carry transcriptional activation domains, and with the exception of Rel-B, they are able to form homo- and heterodimers with the other members of this family of proteins. The balance between different NF-κB homo- and heterodimers will determine which dimers are bound to specific κB sites and thereby regulate the level of transcriptional activity. In addition, these proteins are expressed in a cell- and tissue-specific pattern that provides an additional level of regulation. For example, NF-κB1 (p50) and RelA are ubiquitously expressed, and the p50/RelA heterodimers constitute the most common inducible NF-κB binding activity. In contrast, NF-κB2, Rel-B, and c-Rel are expressed specifically in lymphoid cells and tissues. In unstimulated cells, NF-κB dimers are retained in the cytoplasm in an inactive
form as a consequence of their association with members of another family of proteins called IkB (inhibitors of kB). IkB family of proteins includes IkBa, IkBβ, IkBγ, Bcl-3, and the carboxyl-terminal regions of NF-κB1 (p105) and NF-κB2 (p100). Recent studies have also identified a novel family member, IkBe that is thought to act in the nucleus (Yamazaki et al., 2001). IkB proteins bind with different affinities and specificities to NF-κB dimers. Thus, not only are there different NF-κB dimers in a specific cell type, but the large number of combinations between IkB and NF-κB dimmers illustrates the sophistication of the system.

Although several non-receptor-mediated pathways (such as oxidative stress or UV irradiation) lead to activation of NF-κB, it is the receptor-mediated events which result in activation of these transcription factors that have been best characterized (Fig. 2.5.1.). The binding of a ligand (e.g., tumor necrosis factor alpha [TNF-α], interleukin 1 [IL-1], CD40L, lipopolysaccharide [LPS]) to its receptor triggers a series of events involving protein kinases that result in the recruitment and activation of IkB kinases (IKKs) that phosphorylate IkB. There are at least three components of this signal; some complex—IKKα, IKKβ, and NEMO/IKKγ—which together provide an additional level of regulation that controls gene transcription. The phosphorylation of two serine residues at the NH2 terminus of IkB molecules, for example, Ser32 and Ser36 in IkBa, leads to the polyubiquitination on Lys21 and Lys22 of IkBa and subsequent degradation of the tagged molecule by the 26Sproteasome (Karin and Ben-Neriah, 2000). The degradation of IkB exposes the nuclear localization sequence and allows NF-κB dimers to translocate to the nucleus, bind to κB motifs present in the promoters of many genes, and regulate transcription. As part of an autocrine loop, the binding of NF-κB will induce transcription of IkB genes and so provide a mechanism for limiting the activation of NF-κB activity (Brown et al., 1993). In this system, the activation of NF-κB is independent of de novo protein synthesis and so allows a rapid response to appropriate stimuli. Recent studies have also shown that while IKKβ is required for inducible phosphorylation-dependent degradation of IkB, IKKα is not. Instead, it appears that IKKα preferentially phosphorylates NF-κB2, and this is required for the processing of the p100 NF-κB2 precursor (Chen et al., 2001).
2.5.2. Viral exploitation of NF-κB transactivation property

Since NF-κB activation is an immediate early response of the host defense mechanism during viral infection, it has been an attractive strategy for the pathogens to manipulate the transactivation property of NF-κB towards their benefit.

Some pathogens have developed strategies to interfere with host NF-κB responses. The African swine fever virus, which replicates in macrophages, encodes an IkB-like protein which can bind to RelA and interferes with NF-κB activation (Revilla et al., 1998). Many of the ortho poxviruses also interfere with the regulation of NF-κB, and cowpox virus is capable of inhibiting the degradation of phosphorylated IkB, which may contribute to the pathogenesis of this virus (Oie and Pickup, 2001). Other examples link
the virulence of different bacteria with the ability to inhibit activation of NF-κB. *Mycobacterium ulcerans* causes a progressive necrotizing lesion associated with a lack of an immune response and, lipoprotein preparations from this pathogen can inhibit TNF-α induced activation of NF-κB (Pahlevan et al., 1999). It has been suggested that the ability of this pathogen to subvert TNF-α induced signaling may contribute to the persistence of this infection and the chronic inflammation it causes. Virulent strains of *Yersinia enterocolitica* suppress cellular activation of NF-κB, and as a consequence, the expression of TNF-α by infected macrophages is blocked, resulting in cells undergoing apoptosis (Ruckdeschel et al., 1998). The mechanism that underlies this inhibitory effect is due to the ability of these organisms to inhibit mitogen-activated protein kinase kinase activity and subsequent activation of IκB kinase (Orth et al., 1999; Schesser et al., 1998), likely through the proteolytic activity of YopJ (Orth et al., 2000). In contrast, invasive shigella activates NF-κB in epithelial cells, whereas noninvasive strains do not (Philpott et al., 2000). It appears that nonpathogenic strains of salmonella inhibit the ubiquitination of IκB and thereby prevent its degradation and so inhibit activation of NF-κB. As a result, commensal bacteria in the gut do not stimulate an inflammatory response and are able to survive in this environment (Neish et al., 2000).

While the activation of NF-κB is generally associated with the development of protective immunity against infection, there are cases where pathogens use these events to their advantage. The activation of NF-κB is required for the ability of several viruses to express genes and replicate. The identification of two NF-κB binding sites in the enhancer region of the promoter of the long terminal repeat (LTR) gene of human immunodeficiency virus type 1 (HIV-1) led to studies which investigated the role of these elements in the pathogenesis of AIDS. Stimuli such as IL-1, TNF-α, and LPS enhanced transcription of the HIV-1 LTR through the induction of NF-κB DNA binding activities (Roulston et al., 1995). Thus, the activation of factors that control the expression of immune response genes enhances the replication of HIV. In addition, in vitro studies indicate that HIV-infected cells have a constitutively activated IKK complex and that the presence of IκBα in the nucleus helps to maintain NF-κB–DNA complexes and so enhance NF-κB activity (Deluca et al., 1999). Thus, it appears that multiple mechanisms
underlie the NF-κB-mediated transcription of HIV. Several other viruses have also devised strategies that take advantage of NF-κB to regulate their replication, and hepadna viruses express a protein which activates NF-κB by interacting with IκB α as well as p105, the NF-κB1 precursor (Su and Schneider, 1996).

In vivo correlates of situations in which NF-κB activation is required for viral survival are rare, but studies which used HIV transgenic mice showed that removal of the NF-κB sites in the LTR resulted in decreased rates of proviral gene expression during an inflammatory stimulus (Gazzinelli et al., 1996). Another example is provided by studies with murine encephalo myocarditis virus (EMCV) in which NF-κB1 -/- mice infected with this virus is more resistant than wild-type littermates to infection, and this resistance was initially attributed to enhanced production of the type I interferons (Sha et al., 1995). Subsequent studies confirmed a role for the type I IFN in resistance against EMCV and revealed that NF-κB1 -/- cells undergo rapid apoptosis when infected with EMCV, so compromising the ability of EMCV to replicate (Schwarz et al., 1998). Thus, in normal mice, NF-κB1 protects infected cells from apoptosis and allows the virus to replicate. Therefore, NF-κB signaling pathway is critical for infection as well as for survival.