Molecular Biology of the Hepatitis B Virus for Clinicians

Sibnarayan Datta*, Soumya Chatterjee*, Vijay Veer*, Runu Chakravarty**

*Defence Research Laboratory Tezpur, Tezpur, Assam and **ICMR Virus Unit, Kolkata, West Bengal, India

Hepatitis B virus (HBV) infection is one of the major global health problems, especially in economically underdeveloped or developing countries. HBV infection can lead to a number of clinical outcomes including chronic infection, cirrhosis and liver cancer. It ranks among the top 10 causes of death, being responsible for around 1 million deaths every year. Despite the availability of a highly efficient vaccine and potent antiviral agents, HBV infection still remains a significant clinical problem, particularly in those high endemicity areas where vaccination of large populations has not been possible due to economic reasons.

Although HBV is among the smallest viruses in terms of virion and genome size, it has numerous unique features that make it completely distinct from other DNA viruses. It has a partially double stranded DNA with highly complex genome organization, life cycle and natural history. Remarkably distinct from other DNA viruses, it uses an RNA intermediate called pregenomic RNA (pgRNA) and reverse transcriptase for its genome replication. Genome replication is accomplished by a complex mechanism of *primer shifting* facilitated by direct repeat sequences encoded in the genome. Further, the genome has evolved in such a manner that every single nucleotide of the genome is used for either coding viral proteins or used as regulatory regions or both. Moreover, it utilizes internal *in-frame* translation initiation codons, as well as different reading frames from the same RNA to generate different proteins with diverse functions. HBV also shows considerable genetic variability which has been related with clinical outcomes, replication potential, therapeutic response etc. This review aims at reviewing fundamental events of the viral life cycle including viral replication, transcription and translation, from the molecular standpoint, as well as, highlights the clinical relevance of genetic variability of HBV. (J Clin Exp Hepatol 2012;2:353–365)

Globally, hepatitis B virus (HBV) is a major health problem, with almost 2 billion infected subjects, more than 350 million of whom have chronic hepatitis B infection (CHB).[1] CHB has been associated with 100-fold increase in risk for development of hepatocellular cancer (HCC).[2,3] It is estimated that 15–40% of CHB patients develop severe liver complications such as cirrhosis or HCC in their lifetimes, contributing to more than 1 million deaths every year. The burden of the disease is increasing as almost 45% of the global population lives in economically developing regions having high prevalence (>8%) of chronic HBV infection, that is, in sub-Saharan Africa and eastern Pacific regions, particularly East Asia. Ironically, even though a highly effective prophylactic vaccine has been available for over three decades, HBV still remains among the 10 leading causes of deaths worldwide.

Although, immediately after its discovery, numerous epidemiological studies strongly established the relation between liver complications and HBV, studies on HBV life cycle were greatly hampered because of its narrow host range, unavailability of HBV supporting cell lines and convenient animal models. Nevertheless, with the advent of modern biomolecular tools, major hurdles have been overcome in the recent decades, which have led to an in depth understanding of HBV virology.[4] This article aims at reviewing the key facets of HBV molecular biology with special emphasis on virus replication, transcription, translation and genetic variability.

BRIEF HISTORY OF INFECTIOUS HEPATITIS

The earliest description of '*epidemic hepatitis*', which may or may not have been due to HBV, is credited to Hippocrates (ca 450 BC), almost 2500 years ago. In modern times

Keywords: *Hepadnavirus*, rcDNA, pgRNA, HBsAg, genotype
Received: 17.7.2012; *Accepted:* 12.10.2012; *Available online: 22.10.2012*
Address for correspondence: Dr. Sibnarayan Datta, Scientist 'C', Biodefence & Biodiversity Group, Defence Research Laboratory Tezpur (DRDO), Post Bag No. 2, Tezpur Post office, Tezpur 784 001, Assam, India. Tel.: +91 94353 81750; fax: +91 3712 258534
E-mail: sndatta1978@gmail.com
Abbreviations: HBV: hepatitis B virus; HCC: hepatocellular cancer; DR: direct repeat; ORF: open reading frames; RT: reverse transcriptase; TP: terminal protein; LHB: large envelope protein; pHSA: poly-human serum albumin; IL: interleukin; WHV: woodchuck hepatitis virus; rcDNA: relaxed circular DNA; cccDNA: covalently closed circular; BCP: basal core promoter; ER: endoplasmic reticulum; pgRNA: pregenomic RNA; TNF-α: tumor necrosis factor-α; TGF-α: transforming growth factor-α; dGMP: deoxyguanosine monophosphate; PC: precore; LEF: liver enriched factors; MHR: major hydrophilic region; EN: enhancer; EBP: enhancer binding protein; CHB: chronic hepatitis B infection; MHBs: middle hepatitis B surface antigen; SHBs: small hepatitis B surface antigen; AUG: translation start codon
http://dx.doi.org/10.1016/j.jceh.2012.10.003

Lurman, in a classic epidemiologic study published in 1885, first described 191 cases of '*serum hepatitis*' among 1289 shipyard workers in Bremen, Germany, who received small pox vaccine fortified with human lymph.[5] Large hepatitis epidemics were documented during wars, especially the American Civil War (1861–65), the Franco–Prussian War (1870), World War I (1914–18) and World War II (1939–45).[6,7] The causative agent for serum hepatitis remained a medical enigma until a geneticist, a hematologist and an airline pilot serendipitously found the '*Australia antigen*' (*Au*) in sera of Australian aborigines in the 1960s and identified it as the hepatitis B virus surface antigen (HBsAg) in 1965.[8] In 1967, Krugman and colleagues recognized the parenterally transmitted nature of serum hepatitis, which led to the formulation of hygienic measures to prevent serum hepatitis.[9] The discovery of the Australia antigen was the first step in the development of hepatitis screening assays which dramatically reduced cases of post-transfusion hepatitis and became the basis of a highly effective vaccine that not only controlled hepatitis but also HCC. In 1976, Baruch S. Blumberg was awarded the Nobel Prize in Physiology or Medicine for this landmark discovery of the *Au* antigen.

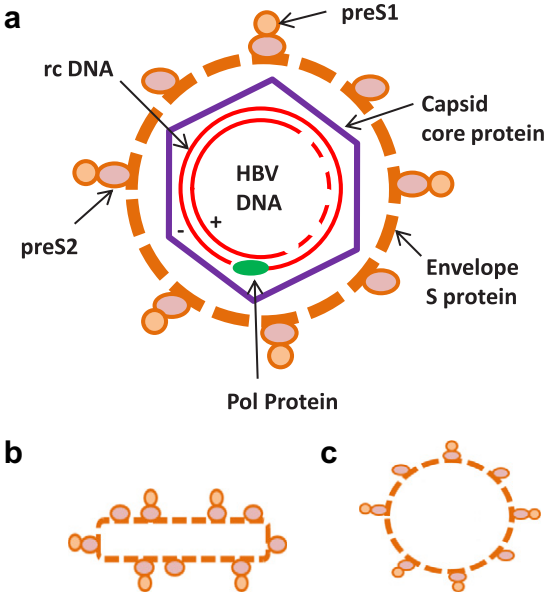

Figure 1 Schematic representation of different forms of infectious and non-infectious hepatitis B virus particles. (a) Complete viral particle, (the 42–47 nm Dane particle), (b & c) two species (filamentous and spherical respectively) of non-infectious 20 nm surface antigen particles. 'Pol' indicates HBV polymerase protein. Figure not according to the scale.

BASIC VIROLOGY

Taxonomic Classification

Unique structural features of the HBV set it apart from other families of animal DNA viruses and have led to its classification under a new family—'*Hepadnaviridae*' (hepatotropic DNA viruses). *Hepadnaviridae* family has two genera, the *Orthohepadnaviruses* that infect mammals (human, woodchucks, ground squirrel etc.), and the *Avihepadnaviruses* that infect birds (ducks, wild herons etc.). Human HBV is the prototype member of the *Hepadnaviridae* family.

Viral Morphology

HBV is unique in different aspects of morphology, genetic material, genome organization, genome replication and genetic control.

Unlike other viruses, HBV produces 3 different types of virus-related particles. As visualized by examining preparations from HBV infected persons under the electron microscope, these are—i) spherical, double shelled particles, 42–47 nm in diameter, ii) spherical particles, 20 nm in diameter, and iii) filaments 20 nm in diameter and of variable length (Figure 1). The 42–47 nm double shelled particles, also called *Dane particles* after DS Dane who first described them in 1970, are the actual infectious particles.[10] The number of Dane particles, which determine infectivity of the sample since only they contain the replication competent HBV genome,[11] may range from as low as 10^2 particles mL^{-1} of serum in occult and asymptomatic infections to more than 10^8 particles mL^{-1} during the active replicative phase of infection. Dane particles incorporate 25–27 nm icosahedrally symmetrical nucleocapsids containing the viral nucleic acid, viral polymerase and associated proteins.[12] The 20 nm spherical particles are produced in excess, up to 1000-fold more than the Dane particles, while the 20 nm filamentous particles are produced in lesser amounts.[4,13] The spherical and filamentous 20 nm particles are mainly composed of the highly immunogenic viral surface glycoprotein, HBsAg, but lack the nucleocapsid with its viral nucleic acid and polymerase activity, so that they are non-infectious.

Viral Nucleic Acid

Another unique feature of HBV is its genome. The HBV DNA is a relaxed, circular, partial double strand, approximately 3.2 kb long. The two strands are asymmetric, a feature exclusive to the *Hepadnaviruses* (Figure 2). The minus strand is complete but contains a 'nick' at a unique site, while the plus strand is incomplete.[12,14,15] The genome sequence has termini with cohesive ends that match the distinctively located 5′ ends of the two strands, and maintain the circular configuration of the DNA.[16] The 5′ ends of both the strands incorporate direct repeats (DRs), regions of short repeat sequences, ~11 nucleotides long, which are crucial in priming viral replication.[17] The 5′ end of the negative DNA strand encodes the first DR, termed 'DR1', while the positive DNA strand starts with another direct repeat, 'DR2'. The negative strand also has a terminal protein, which is a part of the viral polymerase, covalently linked to its 5′ end. On the other hand the plus