

Hepadnaviruses: Variations on the Retrovirus Theme



CHAPTER

- * The virion and the viral genome
- * The viral replication cycle
- * The pathogenesis of hepatitis B virus
- * A plant “hepadnavirus”: cauliflower mosaic virus
- * The evolutionary origin of hepadnaviruses
- * QUESTIONS FOR CHAPTER 21

It is appropriate to end the molecular biology of virus replication with a brief description of a virus group that combines a complex, even bizarre, replication strategy with a very small and compact genome. Add to this complexity the fact that the hepadnaviruses are clearly related to retroviruses, and you have quite a finale!

The hepadnaviruses are named after the propensity, illustrated by the human member—hepatitis B virus, to infect and damage the liver. The name also reflects the fact that they contain a DNA genome in mature capsids. Their relationship to retroviruses is that they encode and incorporate reverse transcriptase into the virion, and the genome is replicated by transcription of the viral genome followed by its conversion into DNA as the virion matures. It has been a difficult technical task to set up systems where hepadnaviruses replicate effectively in cultured cells and they establish persistent (even lifelong) infections in their hosts. The lack of good cell culture models for their study makes them a difficult research subject, but the availability of duck and woodchuck hepatitis virus model systems alleviates this problem to some degree.

The virion and the viral genome

Hepatitis B virus virions, also called *Dane particles* after the investigator who first described their characteristic appearance in the electron microscope, are small (35–45 nm) enveloped icosahedrons. The envelope contains three membrane-associated polypeptides, while the capsid (or core) is composed of a single core capsid protein, the HBc or “core” antigen, along with reverse transcriptase—termed “P”.

The viral genome is 3.2 kb long, making it one of the smallest replication-competent virus genomes known. The virion DNA is *partially* double stranded (ds). As shown in Fig. 21.1, virion DNA contains a full-length negative-sense strand that is complementary to the viral mRNAs, and a partially completed positive-sense strand, the single molecule of P protein in the core is covalently bound to the 3′ end of this partially completed strand. Remember that retroviruses contain

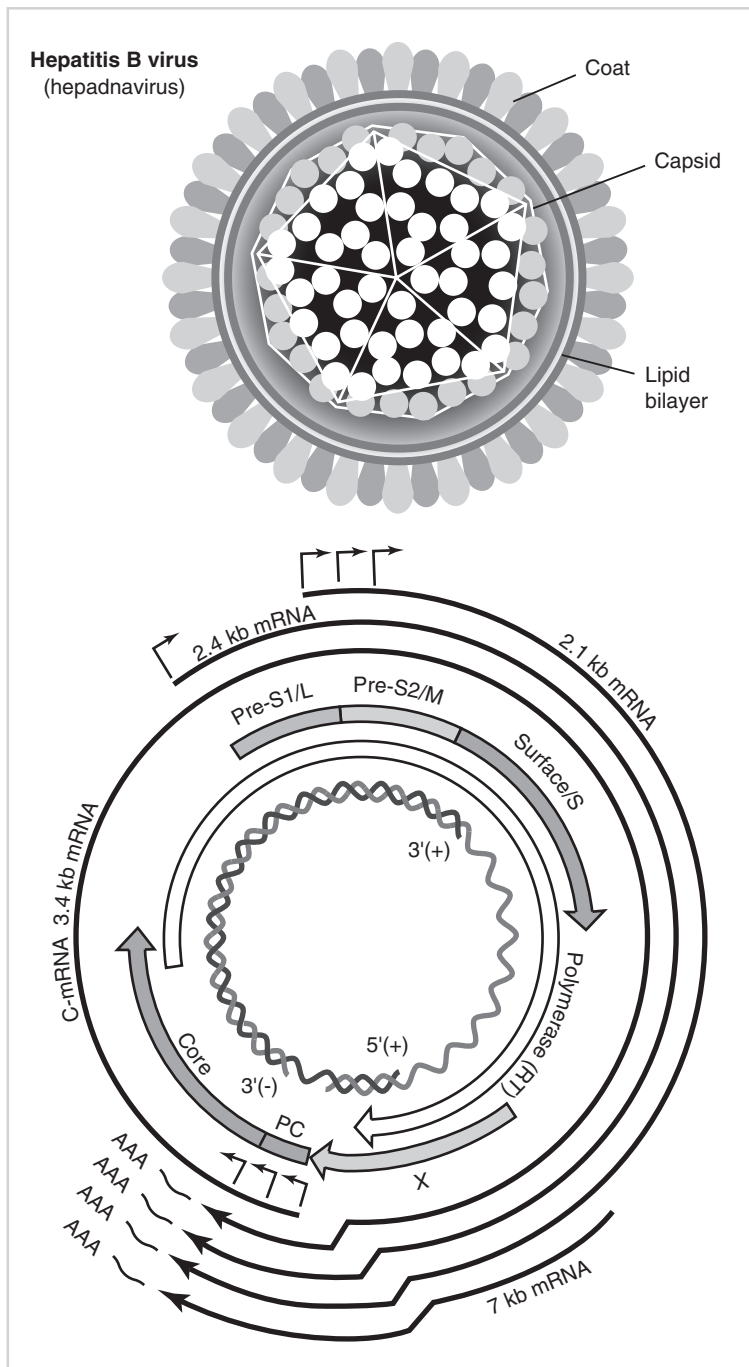


Fig. 21.1 A diagram of the genomic and genetic map of human hepatitis B virus. The virion DNA is partially double stranded and was derived by incomplete reverse transcription of full-length virion RNA transcribed during infection. The genetic map shows the compressed arrangement of the hepatitis virus genes. The bottom panel shows the transcripts expressed from the hepatitis B virus genome. The genome contains three specific promoters. All transcripts terminate at the same site in the interior of the core protein gene.

between 50 and 100 copies of reverse transcriptase in their mature capsids (Chapter 20). The fact that P is actually bound to the partially formed second strand of virion DNA probably is a factor in this difference. The virion DNA is linear but arranged as a circle with a specific gap or nick in the negative-sense strand.

The genetic map of the virus (also shown in Fig. 21.1) is complex. A region at the gap in the negative-sense strand has two sets of repeated sequences, and this region encodes a polyadenylation site and a potential promoter sequence. Reading the sequence in a clockwise manner, the four encoded

translational reading frames are C–core, S (envelope proteins), P–polymerase, and X protein. The core and polymerase open reading frames lie in the same orientation as *gag* and *pol* in a retrovirus genome, but the S translational reading frame overlaps these reading frames. This occurs because the encoded protein information is in a different reading frame.

The viral replication cycle

Following receptor-mediated adsorption, penetration by membrane fusion, and partial uncoating, the partially double-stranded virion DNA is completed by virion reverse transcriptase. The genome migrates to the infected cell's nucleus where the free ends are ligated (probably by cellular enzymes), and the small dsDNA molecule becomes associated with cellular histones to become an episome, or mini-chromosome.

Note, unlike the replication process of retroviruses, viral genomes never become integrated into the cellular genome, and, unlike the episomal DNA of some papovaviruses, the viral genome cannot be replicated by cellular DNA polymerase (see Chapter 17).

Cellular enzymes interacting with virion promoters transcribe four partially overlapping and unspliced mRNAs 3.4, 2.4, 2.1, and 0.7 kb. These transcripts all have distinct 5' ends, but terminate at the same polyadenylation site on the viral genome (see Fig. 21.1).

The largest mRNA (also called C-mRNA), which is longer than the DNA template from which it is expressed by virtue of the location of the polyadenylation signal, has sequences repeated at both ends (like retroviral genomic RNA). It serves as a precursor to virion DNA, when it is encapsidated into immature cores or capsids. The 3.4 kb C-mRNA also encodes the core protein and the P protein from an internal translation initiator. Newly synthesized P can reverse transcribe the 3.4 kb mRNA in the cytoplasm and some of this cDNA generates full ds cDNA, which migrates back to the nucleus of the infected cell where further transcription can take place. Unlike reverse transcription with retroviruses, that mediated by the P protein does not require a tRNA primer—the protein itself serves as a primer.

The S proteins, which are progenitors to the envelope proteins, are expressed by the two smaller transcripts. These transcripts have no obvious counterpart in retroviruses. The 0.7 kb mRNA expresses the X protein.

Expression of viral proteins also leads to the encapsidation of C-mRNA. These immature cores then proceed to generate a complete negative-sense cDNA copy of the encapsidated RNA by the action of encapsidated reverse transcriptase. Reverse transcriptase RNase H activity degrades encapsidated template RNA and partial replication of positive-sense DNA occurs using the negative-sense cDNA as a template while the capsid matures and becomes encapsidated.

The pathogenesis of hepatitis B virus

Hepatitis B virus is only one of a number of viruses targeting the liver. General aspects of the differential pathogenesis of these various hepatitis viruses were discussed in Chapter 4.

Although the hepatitis B virus genome does not integrate into its host cell, the episomal viral transcription unit survives for a long time in infected liver cells. Continued production of new virus, perhaps modulated by the X protein, leads to a persistent infection. Immune-competent individuals who were infected as healthy adults usually can clear the virus after a long recovery period, although permanent liver damage can occur.

Such recovery does not always happen, however, and hepatitis B virus infections of adults can have serious and fatal sequelae. Individuals infected as infants or young children often do not clear the virus efficiently and become chronically infected. A further complication results from infection of immunocompromised adults where high mortality rates from acute hepatic failure are not uncommon. Since the virus is spread by injection of contaminated blood, hepatitis B infections are a

significant danger to medical personnel, especially those treating chronic intravenous drug users and AIDS patients.

It is notable that chronic hepatitis B virus infections acquired in early childhood rather convincingly are statistically correlated with the subsequent development of **hepatocellular carcinomas (HCCs)**. Chronic hepatitis is endemic in areas of Southeast Asia, regions that also demonstrate a high occurrence of fatal liver cancer. The mechanism by which chronic hepatitis infections lead to hepatic carcinoma is not fully understood, but the fact that a very high proportion of infected woodchucks develop HCC when experimentally infected with the related woodchuck hepatitis C virus (HCV) adds strong support to the epidemiological evidence linking HCV and cancer.

Several models of oncogenesis are currently under investigation. It could be that the occasional integration of viral DNA into an infected cell genome leads to the interaction between a viral protein and cellular control circuits in a manner somewhat analogous to human papillomavirus (HPV)-induced carcinomas. The only virus gene product that currently is thought to be a potential candidate for having a direct role in induction of tumors is X protein, which has been shown to have some regulatory and transcriptional stimulatory activities in the laboratory. This model suffers, however, from the fact that a significant portion of cancer cells isolated from HCC patients do not contain any evidence of integrated HCV DNA, and many of those that have viral DNA integrated have extensive rearrangements and deletions in the X protein-encoding sequences.

Another plausible current model is that the continued destruction of liver tissue due to chronic infection leads to abnormal cell growth by a mechanism similar to that seen with chronic human T-cell leukemia virus (HTLV) infections of lymphatic tissue (see Chapter 20). Chronic tissue damage leads to a proliferation of cytokines produced to encourage tissue regeneration, but this eventually leads to mutational damage to cells and ultimately, cancer.

A plant “hepadnavirus”: cauliflower mosaic virus

The known dsDNA viruses of plants include the caulimoviruses (cauliflower mosaic virus) and the badnaviruses (rice tungro bacilliform virus). Like the hepadnaviruses, each of these groups of plant viruses replicate their DNA genomes through a single-stranded RNA intermediate using a reverse transcriptase activity. Since cauliflower mosaic virus has been most widely studied, a few features of this agent are examined herein, and its genome is shown in Fig. 21.2.

Genome structure

The genome of cauliflower mosaic virus is a circular, dsDNA of about 8 kbp that contains three single-strand breaks: one in the transcribed or negative strand and two in the positive strand. The DNA is packaged into a 54 nm icosahedral particle assembled from a single coat protein. The genome itself has six major open reading frames transcribed in a single direction, as well as a large intergenic region.

Viral gene expression and genome replication

The virus's genome is uncoated and delivered to the cell's nucleus where the strand breaks are repaired and a mini-chromosome formed after association with cell histones. Transcription takes place from this template. The genome is transcribed by host RNA polymerase into two species: a 35 s RNA that is slightly longer than the genome and containing a short overlap, and a 19 s RNA. These two transcripts are exported to the cytoplasm where they are translated.

The 35 s RNA is the message for five of the six proteins encoded. Exactly how these proteins are translated is not known, although some form of internal ribosome initiation is proposed. This set of proteins includes the coat protein as well as the viral reverse transcriptase.

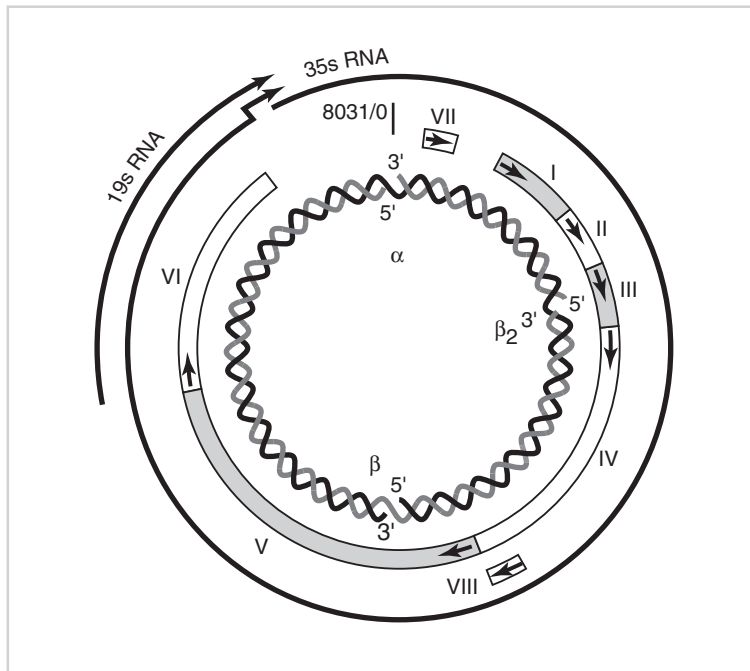


Fig. 21.2 The genome of cauliflower mosaic virus. The three breaks in the genome are indicated by Greek letters. The translational reading frames are indicated by Roman numerals. The two viral mRNAs are also indicated.

The 19 s RNA is translated into the sixth viral protein, a regulatory protein. Genome replication takes place by reverse transcription of the 35 s RNA in the infected cell's cytoplasm. The primer for this replication is a host tRNA specific for methionine and synthesis probably takes place within a precursor particle that will ultimately mature into a virion.

The evolutionary origin of hepadnaviruses

The obvious relationship between the hepadnaviruses, caulimoviruses, badnaviruses, and retroviruses leads to a question that illustrates both the strengths and weaknesses of applying molecular genetics to questions of virus origins. Retroviruses (and their relatives, retrotransposons) are widespread, if not ubiquitous, throughout the plant, animal, and bacterial kingdoms. The conservation of critical sequences of reverse transcriptase suggests that it is an enzyme whose appearance in the biological world was a unique event. It is not too much of a jump to suggest that its appearance occurred early in the evolutionary scene. Indeed, retroviruses may well have had a major role in some types of evolutionary processes, as they display the capacity to mediate horizontal transfer of genes or groups of genes among organisms.

The relationship between the replication strategy and gene order of the DNA viruses related to retroviruses is strong evidence these derive from an ancestral retrovirus, but when this occurred is not known. It can be debated that the presence of hepadnaviruses in avian as well as mammalian species argues for a time of appearance before divergence of their hosts, but it is just as likely that a successful hepadnavirus progenitor arose in mammals and then spread to birds (or vice versa). The same argument holds for looking for predecessors of plant and animal retro-related viruses.

The point is that there is no real way to tell. Even if a clear picture of both bird and mammal or plant and animal ancestry vis-à-vis each other were established, it would not guarantee a resolution of the puzzle. Perhaps detection of the genetic remnants of an ancestral virus in the genomes of the various hosts could shed light on the question, but the fact that the viruses do not integrate as part of their productive life cycle makes the likelihood of finding such a remnant somewhat remote.

QUESTIONS FOR CHAPTER 21

- 1 Compare translation of the hepadnavirus core and polymerase proteins with the translation of the retroviral *gag:pol* region (described in the previous chapter).
- 2 Describe the role of the viral reverse transcriptase in the replication of the DNA genome of hepadnavirus.
- 3 What is the best current model for how hepadnavirus causes hepatic carcinoma?
- 4 Justify the following statement: Hepadnaviruses evolved from an ancestral retrovirus.