Establishment of Molecular Genotyping Methods for Hepatitis B Virus

By
Leena Hussein M. Bajrai
(B.Sc.)

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Biochemistry

Department of Biochemistry
Faculty of Science
King Abdulaziz University
JEDDAH

1429 H/ 2008 G
استحداث طريقة لتحديد أنماط التهاب فيروس الكبد الوبائي ب باستخدام تقنية الأحياء الجزيئية

إعداد
لينا حسين باجري

قدمت هذه الرسالة استكمالاً لمتطلبات درجة الماجستير في العلوم

قسم الكيمياء الحيوية
كلية العلوم
جامعة الملك عبد العزيز - جدة

١٤٢٩ - ٢٠٠٨ م
مختصر

يعتبر فيروس التهاب الكبد الوبائي ب أحد أهم المسببات المعروفة عالميا لأمراض الكبد الحادة والمزمنة والمسنود على وفاة ملايين المصابين به. هناك ثمانية أنماط من فيروس التهاب الكبد الوبائي ب، من H إلى A، المصنف على أساس التشابه الجيني التسلسلي لهذا الفيروس.

وفي هذه الدراسة البحثية، تم تكبير منطقة S من هذا الجينوم عن طريق استخدام تفاعلات البلمرة التسلسلي (Nested PCR) لاستمداد الطريقيين الجزيئيين المعتمدة على تعدد شكل طول جزء الحمض (Restriction Fragment Length Polymorphism)، حيث أن هاتين الطريقيين المستخدمتين قد استعملتا لمعرفة الأنواع الثمانية لفيروس التهاب الكبد الوبائي ب.

وقد تم جمع عينات دم من المتبرعين لفيروس التهاب الكبد الوبائي B وذلك من بنك الدم بمستشفى حراء العام مكة المكرمة ومستشفى جامعة الملك عبد العزيز. حيث تم تجميع 80 عينة دم من المتسكعون الذين تم فحص عيناتهم بشكل روتيني ببنوك الدم المذكورة وتبين اصابتهم بفirus HBsAg بالمرض من خلال النتائج الإيجابية لإختبار الـ DNA. من تلك العينات لإجراء الاختبارات الجزيئية لتحديد الأنواع الثمانية لفيروس التهاب الكبد الوبائي B باستخدام (Nested-PCR) لاستمداد تطبيق الطريقيين المذكورة أعلاه لتحديد أنواع الفيروس.

وبناءً على نتائج هذه الدراسة، تبين أن عدوى فيروس التهاب الكبد الوبائي B منتشر بشكل واضح بثلاثة أنواع رئيسية (D, E, A) بنسبة متتالية (5, 6.3, 2.7) بالإضافة إلى وجود أنواع مختلطة وهي B+D و B+G. وبالتالي تعتبر النمط الفيروسي D الأكثر انتشارًا في شريحة المتبترعين بالدم المشمولين بالدراسة بمنطقة الشرق الأوسط ومنطقة البحر الأبيض المتوسط.
Abstract

Hepatitis B virus (HBV) is one of the major causative agents of acute and chronic liver disease worldwide and is believed to be responsible for a million deaths annually. Eight genotypes of HBV, A to H, have been described on the basis of similarity of the complete genome sequence. In this study, the S regions of HBV were amplified by nested-PCR to establish molecular detection methods based on Restriction Fragment Length Polymorphism (RFLP) and DNA sequencing. The established methods were used to determine the HBV genotypes among infected donors from two blood transfusion centers within the western region of Saudi Arabia (General Hera'a Hospital, Makkah, King Abdulaziz University Hospital, Jeddah). HBV genotypes were determined in 80 blood donors who had HBsAg positive employing two different nested PCR based molecular methods. The result showed that HBV infection in blood donors are attributed predominantly to viral genotypes (D, E, A) that constituted (83.7%, 6.3%, 5%) respectively of the total infection. In addition, there was mixed B+G (2.5%) and B+D (1.3%) among the studied group. This study indicates that HBV genotypes D predominant in Westren of Saudi Arabia as it does through out the Middle East and Medetrianen area.
Table of Contents

Acknowledgments…………………………………………………… ii
Abstract………………………………………………………… iv
Arabic Abstract………………………………………………… v
Table of Contents……………………………………………… vi
List of Abbreviations and Symbols……………………………. x
List of Figures………………………………………………… xiii
List of Tables………………………………………………… xv

Chapter 1:  

1. Introduction  

1.1. History of Viral Hepatitis  

1.2. Structure of Hepatitis B Virus  

1.3. Hepatitis B Virus Antigens  

1.4. Genome of Hepatitis B Virus  

1.5. Replication of Hepatitis B Virus  

1.6. Hepatitis B Virus Genotypes Classification  

1.7. Viral Gene Expression  

1.7.1. Amplification of cccDNA and Production of Infectious Virus  

1.8. Pathogenesis of Human Infection with Hepatitis B Virus  

1.9. Clinical Aspects of Hepatitis B Virus Infection  

1.9.1. Acute Infection  

1.9.2. Chronic Infection  

1.10. Epidemiology of Hepatitis B Virus Infection  

Pages  

1  

1  

2  

5  

8  

11  

14  

14  

16  

17  

18  

19  

22  

24
Chapter 2:

2. Materials and Methods

2.1. Materials

2.1.1. Equipments
2.1.2. Chemicals, Reagents and Kits
2.1.3. Clinical Samples for Sero-prevalence and Genotyping Study

2.2. Methods

2.2.1 DNA Extraction
2.2.2. Polymerase Chain Reaction
2.2.3. Identification of amplified fragment
2.2.4. Genotyping Methods

2.2.4.1. Restriction Fragments Length Polymorphism (RFLP)
2.2.4.2. Sequencing Method
2.2.4.3. Reverse Hybridization for HBV genotyping

2.3. Statistical Analysis

Chapter 3:

3. Results

3.1. Introduction

3.2. HBV genotyping results as revealed by RFLP

3.2.1. Sample Collection
3.2.2. Results of the nested PCR
3.2.3. Results of RFLP
3.2.4. Comparison of HBV genotypes determined by INNO-LiPA and RFLP

3.3. HBV genotypes by DNA Sequencing

3.3.1. DNA Sequencing of HBV blood donor samples in Makkah

3.3.2. DNA Sequencing of HBV blood donor samples in Jeddah

3.3.3. Comparison of HBV genotypes between Makkah and Jeddah blood donor samples by RFLP and Sequencing methods

Chapter 4:

4. Discussion

Appendices

References

Summary in Arabic
## Abbreviations and Symbols

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALT</td>
<td>Alanine aminotransferase</td>
</tr>
<tr>
<td>ABI</td>
<td>Applied Biosystems</td>
</tr>
<tr>
<td>Anti-HBe</td>
<td>Hepatitis B e Antibody</td>
</tr>
<tr>
<td>Anti-HBs</td>
<td>Hepatitis B Surface Antibody</td>
</tr>
<tr>
<td>AST</td>
<td>Aspartate aminotransferase</td>
</tr>
<tr>
<td>Au</td>
<td>Australia</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>CAE</td>
<td>Capillary Array Electrophoresis</td>
</tr>
<tr>
<td>cccDNA</td>
<td>Covalently Closed Circular Double-Strand DNA</td>
</tr>
<tr>
<td>CHV</td>
<td>Crane Hepatitis B Virus</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>ddNTPs</td>
<td>Dideoxynucleotides</td>
</tr>
<tr>
<td>DHV</td>
<td>Duck Hepatitis B Virus</td>
</tr>
<tr>
<td>dNTPs</td>
<td>Deoxynucleoside triphosphate</td>
</tr>
<tr>
<td>DR1</td>
<td>Direct Repeats1</td>
</tr>
<tr>
<td>DR2</td>
<td>Direct Repeats2</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic Reticulum</td>
</tr>
<tr>
<td>EN1</td>
<td>Enhancer1</td>
</tr>
<tr>
<td>EN2</td>
<td>Enhancer2</td>
</tr>
<tr>
<td>EIA</td>
<td>Enzyme Immunoassay</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
</tbody>
</table>
GHV  Goose Hepatitis B Virus
GSHV  Ground Squirrel Hepatitis B Virus
HAV  Hepatitis A Virus
HBV  Hepatitis B Virus
HBcAg  Hepatitis B Core Antigen
HBeAg  Hepatitis B e Antigen
HBIG  Hepatitis B Immunoglobulin
HBsAg  Hepatitis B Surface Antigen
HCC  Hepatocellular Carcinoma
HCV  Hepatitis C Virus
HHV  Heron Hepatitis B Virus
HIV  Human Immunodeficiency Virus
IFN-α  Interferon-alpha
IgM  Immunoglobulin M
LHBs  Large Surface Protein
MgCl₂  Magnesium Chloride
MHBs  Middle Surface Protein
MSM  Multisexual Men
mRNA  Messenger ribonucleic acid
PC  Precore
PCR  Polymerase chain reaction
PET  Positive effector of transcription
pgRNA  Pregenomic ribonucleic acid
rcDNA  Relaxed circular DNA
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFLP</td>
<td>Restriction Fragment Length Polymorphism</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RIAs</td>
<td>Radioimmunoassays</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-boric acid-EDTA</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra violet</td>
</tr>
<tr>
<td>WHV</td>
<td>Woodchuck Hepatitis Virus</td>
</tr>
<tr>
<td>YMDD</td>
<td>Tyrosine-Methionine-Aspartate-Aspartate</td>
</tr>
</tbody>
</table>
# List of Figures

## Chapter 1

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Genome of hepatitis B virus</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>Transcriptional and translation map of HBV</td>
<td>10</td>
</tr>
<tr>
<td>1.3</td>
<td>Replication steps of hepadnavirus</td>
<td>12</td>
</tr>
<tr>
<td>1.4</td>
<td>Serological profile of acute, resolving hepatitis B virus</td>
<td>20</td>
</tr>
<tr>
<td>1.5</td>
<td>Serological profile of chronic hepatitis B with seroconversion</td>
<td>22</td>
</tr>
</tbody>
</table>

## Chapter 2

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Illustration of nested PCR</td>
<td>46</td>
</tr>
<tr>
<td>2.2</td>
<td>Diagrammatic representation of the position of PCR primers and RFLP analyses for HBV genotyping using the 2nd round PCR product</td>
<td>48</td>
</tr>
</tbody>
</table>

## Chapter 3

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Gel electrophoresis for the products obtained by nested PCR</td>
<td>56</td>
</tr>
<tr>
<td>3.2</td>
<td>Distribution of HBV genotypes between Saudi and non-Saudi blood donors by RFLP</td>
<td>59</td>
</tr>
<tr>
<td>3.3</td>
<td>RFLP of HBV- genotypes E, B and C or D samples</td>
<td>60</td>
</tr>
<tr>
<td>3.4</td>
<td>RFLP of HBV- genotypes A, E or G samples</td>
<td>61</td>
</tr>
<tr>
<td>3.5</td>
<td>RFLP of HBV- genotypes D or E samples</td>
<td>62</td>
</tr>
<tr>
<td>3.6</td>
<td>RFLP of HBV- genotypes B, D or E samples</td>
<td>62</td>
</tr>
<tr>
<td>3.7</td>
<td>HBV genotyping by INNO-LiPA</td>
<td>64</td>
</tr>
</tbody>
</table>
3.8. Distribution of HBV genotypes by Sequencing among blood donors 71

3.9. Distribution of HBV genotypes between Saudi and non-Saudi blood donors as determined by Sequencing 72
# List of Tables

## Chapter 1

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1.</td>
<td>Prevalence of hepatitis B virus</td>
<td>21</td>
</tr>
<tr>
<td>1.2.</td>
<td>Interpretation of available serologic test results for HBV</td>
<td>27</td>
</tr>
</tbody>
</table>

## Chapter 3

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.</td>
<td>Distribution of HBV genotypes by RFLP in Makkah and Jeddah blood donors</td>
<td>58</td>
</tr>
<tr>
<td>3.2.</td>
<td>Comparison of HBV genotypes of blood donors as determined by LiPA and RFLP</td>
<td>65</td>
</tr>
<tr>
<td>3.3.</td>
<td>Distribution of HBV genotypes from Makkah blood donors as determined by Sequencing</td>
<td>67</td>
</tr>
<tr>
<td>3.4.</td>
<td>Distribution of HBV genotypes from Jeddah blood donors as determined by Sequencing</td>
<td>69</td>
</tr>
</tbody>
</table>
1. Introduction

1.1 History of Viral Hepatitis

Hepatitis is an inflammation of the liver characterized by presence of inflammatory cells in the liver tissue. It is caused by many factors including autoimmune disorders, physical, chemical or biological agents (Bernard et al., 1996).

The name of viral hepatitis is based on "hepato" (liver) "tropic" (replicates and causes infection in the liver). Hepatitis B Virus (HBV) belongs to *Hepadna* virus family which considered as the smallest DNA genome of animal DNA viruses (Bernard et al., 1996). The family consists of genus *Orthohepadnavirus*, which is made up of *hepadnaviral* species infecting mammals, and genus *Avihepadnavirus*, which consists of *hepadnaviruses* infecting birds. Hepatitis B virus, the smallest of all human DNA viruses, is a member of this family of highly species-specific enveloped DNA viruses which replicate by reverse transcription of a viral RNA, the pregenome (Summers and Mason, 1982).

In 1947, the terms, hepatitis A and hepatitis B, were first introduced by MacCallum in order to categorize infectious (epidemic) and serum hepatitis (MacCallum, 1947). Thus, these terms were eventually adopted by the World Health Organization Committee as Viral Hepatitis (WHO, 1973). Before isolating, transmission of viruses that caused hepatitis was differentiated on the basis of epidemiological observations. Hepatitis A virus was considered predominantly transmitted via the fecal-oral route while type B hepatitis was believed to be primarily transmitted parenterally (Blumberg et al., 1967).
Most of acute hepatitis is caused by one of hepatotropic viruses A, B, C, D, E, F, or G which are classified into different viral hepatitis families. However, chronicity is common with hepatitis B and C (Blumberg et al., 1967).

In 1963, Blumberg discovered a previously unknown protein in the blood of an Australian antigen (Blumberg et al., 1967). This protein was denoted as the Australia (Au) antigen. It became apparent that this protein was related to hepatitis B, then by 1968, Okochi and Murakami had established that the Au antigen (now known as the hepatitis B surface antigen) is usually found in the serum of hepatitis B infected patients (Prince, 1968; Okochi and Murakami, 1968).

1.2 Structure of Hepatitis B Virus

Viruses are the smallest infectious agents (0.01-0.4 µm) and so able to pass bacterial filters. It contains either DNA or RNA, hence they classified to DNA viruses and RNA viruses. Viruses are unable to form its own energy, metabolic intermediates and replication (non-viable outside host cell) so, they are obligatory intracellular (Summers et al., 1975).

As shown in Figure 1.1, the hepadnavirus genome is a partially double-stranded circular DNA structure that is encapsidated within the enveloped viral particle (Summers et al., 1975). Upon infection of hepatocytes, viral DNA is transported to the nucleus, where it is converted to a covalently closed, circular, double-stranded DNA (cccDNA). This DNA is essentially an episome that does not replicate but functions solely as the template for all viral transcription (Summers et al., 1975).

HBV is a member of the Hepadnaviridae family of enveloped DNA viruses (Ganem and Schneider, 2001). It is the smallest DNA virus known: it has only 3,200 bp in its circular pattern genome that is uniquely organized in a partly double stranded (Figure 1.1). It is enveloped and replicates in the liver and causes hepatic
dysfunction (Lau and Wright, 1993). The natural host for HBV is humans, although similar viruses have been isolated from apes, woodchucks (woodchuck hepatitis virus [WHV]), squirrels (ground squirrel hepatitis virus [GSHV]), herons (heron hepatitis B virus [HHV]), ducks (duck hepatitis B virus [DHV]), geese (goose hepatitis B virus [GHV]), and cranes (crane hepatitis B virus [CHV]) (Summers et al., 1975; Ganem and Schneider, 2001).
Figure 1.1  Genome of Hepatitis B Virus.
Adapted from (David and Frank, 1994)
1.3 Hepatitis B Virus Antigens

The major (or small) surface protein is 226 amino acids long and found in non-glycosylated (p24) and glycosylated (gp27) forms. It is encoded in the thirty half of the surface open reading frame (ORF) and translated from the third of three in-phase initiation codons (Zuckerman et al., 2004). Larger, pre-S proteins are translated utilising the two upstream initiation codons; translation from the second results in two intermediate-sized glycoproteins (gp33 and gp36) with a glycosylated 55 amino acid N-terminal extension, the pre-S2 domain (Zuckerman et al., 2004). These middle surface proteins are minor components of virions and subviral particles. Translation of the entire ORF (pre-S1+pre-S2+S) gives rise to the large surface proteins (p39 and gp42) which are found predominantly in virions and perhaps also the tubular, 22nm forms. A domain within the pre-S1 region seems to be responsible for the attachment of the virus to a receptor on the hepatocyte. Synthesis of the pre-S1 protein also may act as a signal for virion assembly in the infected cell (Zuckerman et al., 2004).

The subviral particle and the virion surface are composed of hepatitis B surface antigen (HBsAg) anchored in a lipid bilayer derived from the endoplasmic reticulum of the host cell. The major antigenic determinant on the particles is the common, group-specific antigen, a, which is believed to form a ‘double-loop’ structure (aa 124–137 and aa 139–147) on the surfaces of the virions and subviral particles (Zuckerman et al., 2004). The formation of anti-A antibodies following vaccination seems to be sufficient to confer protective immunity. The major HBsAg protein also carries a pair of mutually exclusive subdeterminants, d or y and w or r, which in each case seem to correlate with variation at single amino acid positions (aa 122 and 160, respectively). Thus, four principal phenotypes of HBsAg are recognised—adw, adr,
ayw and, more rarely, ayr—and these show differing geographical distribution. For example, in Northern Europe, the Americas and Australia subtype adw predominates, whilst ayr occurs in a broad zone that includes Northern and Western Africa, the Eastern Mediterranean, Eastern Europe, Northern and Central Asia, and the Indian subcontinent. Both adw and adr are found in Malaysia, Thailand, Indonesia and Papua New Guinea, whereas subtype adr predominates in other parts of south-east Asia, including China, Japan and the Pacific Islands. The subtypes provide useful epidemiological markers of HBV (Zuckerman et al., 2004). Unusual variants which lack the group specific antigen, a, may be selected by antibody in immunized infants infected perinatally and in persistently infected individuals following treatment with hepatitis B immunoglobulin or a natural antibody response. Surface variants of HBV were first described in Italy, infecting children and adults in the presence of specific hepatitis B surface antibodies (anti-HBs) after several months of successful immunisation with two generally licensed hepatitis B vaccines given with and without hepatitis B immunoglobulin (Carman et al., 1990).

The core protein (p22) is the major component of the nucleocapsid and has an arginine-rich domain at its carboxyl terminus, which presumably interacts with the viral nucleic acid. Hepatitis B core antigen (HBcAg) is translated from the second initiation codon in the core ORF (Zuckerman et al., 2004). Translation from the upstream initiation codon yields a precursor protein (p25) which is processed to yield hepatitis B e antigen (HBeAg). The precore region, between the two initiation codons, encodes a signal sequence which directs p25 to the endoplasmic reticulum, where it is cleaved by a cellular signal peptidase (Zuckerman et al., 2004). HBeAg is secreted following further proteolysis, which removes the carboxyl-terminal domain. HBeAg is also expressed on the surface of the infected hepatocyte and is a major
target for the cellular immune system. HBeAg is not an essential protein of the virus. Variants of HBV with mutations in the precore region (precore mutants), and which are defective for the synthesis of HBeAg, are discussed below. The polymerase (P) ORF, which overlaps the other three, encodes the viral polymerase (Zuckerman et al., 2004). This enzyme has both DNA- and RNA-dependent activities and the predicted aa sequence has been shown to have homology with retroviral reverse transcriptases. The polymerase protein also acts as the primer for minus strand DNA synthesis and has an ‘RNase H’ activity which degrades the RNA pregenome during minus strand synthesis. The fourth open reading frame has been termed X because the function of its product was originally obscure. It is now known that this protein acts as a transcriptional transactivator and may enhance the expression of the other viral proteins (Zuckerman et al., 2004).

Orthohepadnaviruses express three envelope components called (S), (M) and (L). The smallest component, S (226 amino acids long), defines the S domain. The extra domain of the middle component (M) is known as pre-S2, while the domain unique to the largest one L is called pre-S1 (Dane et al., 1970). All three envelope components are glycosylated, type II transmembrane proteins that can form multimers stabilized by disulfide bridges formed by cysteine residues present in the S domain. They are all found, S, L and M, as components of the 42-nm-diameter infectious viral particles, also known as Dane particles (Dane et al., 1970). The largest (L) and the meddil (M) components are present in roughly equal amounts in Dane particles and together constitute approximately 30% of the envelope protein content (Heermann et al., 1987). The smallest one (S) by itself, and together with the larger envelope proteins, also forms filamentous and spherical "surface antigen" particles that are secreted from infected cells in at least 100-fold excess over virions.
These spheres and filaments can accumulate to concentrations of several hundred micrograms per milliliter in the blood of HBV-infected patients. Complexes of these particles with their cognate antibodies are probably responsible for the immune complex syndromes that sometimes occur during transient infections (Lambert et al., 1991; Yuasa et al., 1991).

Antibodies to surface antigen particles composed of S protein alone are sufficient to provide protection against HBV infection. However, there is good reason to believe that the pre-S1 domain is, at least in part, the substrate for the still elusive viral receptor. Epitopes in pre-S1 displayed on the outside of surface antigen can also elicit virus-neutralizing antibodies and alter the host range of the virus upon genetic recombination (Lambert et al., 1991; Yuasa et al., 1991; Ishikawa and Ganem, 1995).

1.4 Genome of Hepatitis B Virus

As shown in Figure 1.1, the minus strand of the DNA (cccDNA) is almost a complete circle (Summers and Mason, 1982; Lau and Wright, 1993). It consists of two ORF in the form of overlapping genes that encode both structural proteins (pre-S, surface and core) and replicative proteins (polymerase and X protein). The plus strand of the DNA is shorter and variable in length (Lau and Wright, 1993).

Four mRNA transcripts of known function have been identified as being involved in HBV transcription and translation, which utilize the same polyadenylation signal and transcribed from cccDNA (Summers and Mason, 1982; Lau and Wright, 1993). The pregenomic RNA (pgRNA), the longest transcript (3.5 kb), is the template for genome replication and the expression of precore/core and polymerase proteins. Another transcript (2.4-kb) encodes pre-S1, pre-S2 and HBsAg, and transcript (2.1-kb) encodes only pre-S2 and HBsAg, while the smallest transcript
(0.7 kb) encodes the X protein (Summers and Mason, 1982; Gerlich and Bruss, 1993).

As shown in Figure 1.2, the inner circle depicts the rcDNA with the reverse transcriptase attached to the 5’ end of the complete minus-strand DNA and a capped RNA oligomer attached to the 5’ end of the incomplete plus-strand DNA. The positions of the direct repeats, DR1 and DR2, as well as the positions of the two enhancers, EN1 and EN2, are indicated (Christoph and William, 2000). The outer circle depicts the three major viral RNAs; the core (C) or pgRNA, the pre-S (L) mRNA, and the S mRNA. The common 3’ ends of the three mRNAs are indicated by the letters A. Not shown in the Figure 1.2, is the putative X mRNA that spans the X coding region and terminates at the site indicated for the other three mRNAs. The four protein-coding regions are shown between the inner and outer circles. They include the precore (PC) and core genes, the polymerase gene, and the X gene. In Figure 1.2, the envelope genes pre-S1 (L), pre-S2 (M) and surface (S) overlap with the polymerase open reading frame (Christoph and William, 2000).

The surface/pre-S gene encodes for the virus envelope. The major protein that forms the HBsAg particles is the smallest gene product (SHBs). The middle protein (MHBs), which contains the pre-S2 component, and the large surface protein (LHBs), which contains pre-S1, are also incorporated into HBsAg particles but are found in larger proportions in the intact virus particles. The pre-S proteins play an important role in the attachment of HBV to hepatocytes (Gerlich and Bruss, 1993). Liver-specific attachment sites have been identified in vitro for pre-S1 and pre-S2 (Neurath et al., 1986; Pontisso et al., 1989; Gerlich et al., 1993). Pre-S2 was found to attach to artificially polymerized human serum albumin (Pontisso et al., 1989). Since binding of this albumin has also been observed on hepatocytes, it has been
hypothesized that binding of HBV to its host cell might be mediated by a bridge of modified albumin (Komani and Peeples, 1990; Leenders et al., 1990).

**Figure 1.2** Transcriptional and translational map of HBV. Adapted from (Christoph and William, 2000)