View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Universiti Putra Malaysia Institutional Repository

UNIVERSITI PUTRA MALAYSIA

DISPLAY OF THE PRES REGIONS OF THE SURFACE ANTIGEN OF HEPATITIS B VIRUS ON M13 AND T7 BACTERIOPHAGES

KOK WAI LING

FSAS 2001 42

DISPLAY OF THE PRES REGIONS OF THE SURFACE ANTIGEN OF HEPATITIS B VIRUS ON M13 AND T7 BACTERIOPHAGES

By

KOK WAI LING

Thesis Submitted in Fulfilment of the Requirement for the Degree of Master of Science in the Faculty of Science and Environmental Studies Universiti Putra Malaysia

June 2001

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

DISPLAY OF THE PRES REGIONS OF THE SURFACE ANTIGEN OF HEPATITIS B VIRUS ON M13 AND T7 BACTERIOPHAGES

By

KOK WAI LING

June 2001

Chairman: **Tan Wen Siang, Ph.D.**

Faculty: **Science and Environmental Studies**

Hepatitis B virus (HBV) is the prototype of the family *Hepadnaviridae,* which causes liver disease in humans, mammals and birds. The envelope of HBV contains three related surface antigens (termed L-, M- and S-HBsAg) produced by alternative initiation of translation in a single coding region. These polypeptides harbour a common 226 amino acids at their C-terminus, which is also the entire length of the S-HBsAg. The M-HBsAg contains an N-terminal extension of 55 amino acids known as the PreS2 region. The longest of the three, L-HBsAg, has the PreS1 region of 108 or 119 amino acids (depending on serotype) followed by the PreS2 and the S regions. The PreS domain is believed to be involved in virion assembly and attachment to a hepatocyte receptor during infection. In order to study the functions of this region, the PreS and PreS1 domains were fused to the g3p protein of bacteriophage M13 and 10B protein of bacteriophage T7,

respectively, that allow the fusion proteins to be displayed. The PreS-g3p fusion protein produced in a suppressor strain of *Escherichia coli* was detected by the anti-E tag antibody with a size of approximately 66 kDa on a Western blot. In a non-suppressor strain of *E. coli*, the soluble PreS protein was detected in the medium, periplasm and cytoplasm with a molecular mass of approximately 22 kDa. Meanwhile in the T7 system, the first and second halves of PreS1 were detected by the T7 Tag antibody on a Western blot with a size of around 50 kDa. The functional display of the PreS region would provide an alternative means to study its interactions with the nucleocapsid and hepatocytes. Precise definition of the regions and specific amino acids in L-HBsAg that are required for efficient interaction with the nucleocapsid and hepatocytes, may help to identify lead compounds for therapeutic agents based upon inhibition of viral morphogenesis.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

PERSEMBAHAN ANTIGEN PERMUKAAN KAWASAN PRES VIRUS HEPATITIS B PADA BAKTERIOFAJ M13 DAN T7

Oleh

KOK WAI LING

Jun 2001

Pengerusi: **Tan Wen Siang, Ph.D.**

Fakulti: **Sains dan Pengajian Alam Sekitar**

Virus Hepatitis B (HBV) adalah prototaip famili *Hepadnaviridae* yang menyebabkan penyakit hati pada manusia, mamalia dan burung. Penyalut HBV terdiri daripada tiga jenis antigen permukaan yang saling berkaitan (dikenali sebagai L-, M- dan S-HBsAg) hasil daripada tanslasi pemulaan yang berlainan dalam satu kawasan pengkodan. Kesemua polipeptida ini mengandungi 226 asid amino pada terminal-C yang juga merupakan keseluruhan polipeptida S-HBsAg. M-HBsAg mengandungi 55 asid amino terlunjur dari terminal-N yang dikenali sebagai kawasan PreS2. L-HBsAg yang merupakan polipeptida terpanjang antara ketiga-tiga polipeptida tersebut, mengandungi kawasan PreS1 yang terdiri daripada 108 atau 119 asid amino (bergantung pada serotaip), diikuti oleh kawasan PreS2 dan kawasan S. Kawasan PreS adalah dipercayai terlibat dalam pembentukan virus dan pelekatan pada reseptor hepatosit semasa jangkitan. Demi mengkaji fungsi

kawasan ini, kawasan PreS dan PreS1, masing-masing, digabungkan dengan protein g3p pada M13 dan protein 10B pada T7 yang membenarkan gabungan protein tersebut dipersembahkan. Gabungan protein PreS-g3p yang dihasilkan oleh strain tertindas *Escherichia coli* dapat dikesan dengan antibodi anti-E tag dengan saiz lebih kurang 66 kDa pada pemblotan Western. Pada strain tak-tertindas *E. coli*, protein PreS terlarut dikesan di dalam medium, periplasma dan sitoplasma dengan jisim molekul lebih kurang 22 kDa. Sementara itu, dalam sistem T7, separuh pertama dan kedua PreS1 dapat dikesan dengan antibodi T7 Tag pada pemblotan Western dengan saiz berukuran sekitar 50 kDa. Persembahan berfungsi kawasan PreS akan membekalkan satu cara alternatif untuk mengkaji interaksinya dengan nukleokapsid dan hepatosit. Definisi yang tepat terhadap kawasan terlibat serta asid amino spesifik pada L-HBsAg yang berkesan dalam interaksi nukleokapsid dan hepatosit mungkin dapat membantu dalam penentuan sebatian untuk dijadikan sebagai agen terapeutik berdasarkan kepada perencatan morfogenesis virus.

ACKNOWLEDGEMENTS

There are a lot of wonderful people whom I would like to acknowledge. First and foremost, I wish to convey my most sincere gratitude to my friendly and helpful supervisor, Dr. Tan Wen Siang for teaching me the ABC in molecular biology. I have indeed gained tremendously from his constant guidance, invaluable advice and great motivation throughout the period of this study, subsequently, bringing this project into existence.

I am also very grateful to my other two supervisors, Associate Professor Datin Dr. Khatijah Yusoff and Associate Professor Dr. Sheila Nathan for their helpful discussions and constructive suggestions. Special thanks to the staff of the department and members of theVirology laboratory: Subha, Pria, Kok Lian, Chiew Ling, Chui Fung, Sing King, Tang, Amir, Rebecca, Swee Tin, Lau, Wawa and Sharifah for making my time in the laboratory joyful and pleasant with all their jokes and funny gestures. I also wish to thank Mr. Majid Eshaghi for his suggestions and advise.

Last but not least, I am greatly indebted to my parents, sister and brothers for their love, support and encouragement. A special thanks is also due to John Hun for his love, support and motivation.

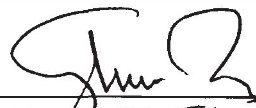
I certify that an Examination Committee met on 6th June 2001 to conduct the final examination of Kok Wai Ling on her Master of Science thesis entitled "Display of the PreS Regions of the Surface Antigen of Hepatitis B Virus on M13 and T7 Bacteriophages" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Michael C.V.L. Wong, Ph.D.
Lecturer
Faculty of Food Science and Biotechnology
Universiti Putra Malaysia
(Chairman)

Tan Wen Siang, Ph.D.
Lecturer
Department of Biochemistry and Microbiology
Universiti Putra Malaysia
(Member)

Datin Khatijah Yusoff, Ph.D.
Associate Professor
Department of Biochemistry and Microbiology
Universiti Putra Malaysia
(Member)

Sheila Nathan, Ph.D.
Associate Professor
Faculty of Science and Technology
Universiti Kebangsaaan Malaysia
(Member)



MOHD. GHAZALI MOHAYIDIN, Ph.D.
Professor/Deputy Dean of Graduate School,
Universiti Putra Malaysia.

Date: 28 JUN 2001

This thesis submitted to the Senate of Universiti Putra Malaysia has been accepted as fulfilment of the requirement for the degree of Master of Science.

AINI IDERIS, Ph.D.
Professor
Dean of Graduate School,
Universiti Putra Malaysia.

Date:

DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations, which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UPM or other institutions.

KOK WAI LING

Date: 28 JUNE 2001

TABLE OF CONTENTS

				Page
ABSTRACT				ii
ABSTRAK				iv
ACKNOWLEDGEMENTS				vi
APPROVAL				vii
DECLARATION				ix
LIST OF TABLES				xii
LIST OF FIGURES				xiii
LIST OF ABBREVIATIONS				xv
CHAPTER				
1	INTRODUCTION			1
2	LITERATURE REVIEW			5
	2 1	Hepatitis B virus		5
		2 1 1	Morphology and Genome Structure	6
		2 1 2	Genomic Organisation and Viral Transcripts	7
	2 2	Surface Antigens		10
	2 3	Biology of Filamentous Phage		14
		2 3 1	Morphology and Structure	14
		2 3 2	Life Cycle of Filamentous Phage	15
	2 4	Biology of Bacteriophage T7		17
	2 5	What is Phage Display?		19
		2 5 1	Types of Phage-Display Systems	20
		2 5 2	Affinity Selection	23
		2 5 3	Applications of Phage Display	25
3	MATERIALS AND METHODS			30
	3 1	Media, Buffers and Solutions		30
	3 2	Bacterial Strains and Plasmid		31
	3 3	Display of PreS Regions on Bacteriophage M13		32
		3 3 1	Amplification of the PreS Coding Regions with PCR	32
		3 3 2	Purification of DNA	33
		3 3 3	Restriction Enzyme (RE) Digestion of the Purified PCR Product	34
		3 3 4	Ligation to the pCANTAB5E Phagemid Vector	35
		3 3 5	Preparation and Maintenance of *Escherichia coli* Stocks	35
		3 3 6	Preparation of Competent Cells and Transformation	36
		3 3 7	Screening of Recombinant Phagemid	37
		3 3 8	Phage Rescue with M13KO7 Helper Phage	38

3 3 9 Expression and Detection of Phage Displayed Protein 38
3 3 10 Expression of Soluble Proteins 40
3 4 Display of PreS Regions on Bacteriophage T7 43
3 4 1 Cloning in T7 Vectors 43
3 4 2 *In vitro* Packaging of T7 DNA 44
3 4 3 Plaque Assay 45
3 4 4 Preparation of Phage Lysate 46
3 4 5 Amplification of Phage DNA 47
3 4 6 Sequencing of the Putative T7 Recombinants 47
3 4 7 Protein Expression in T7 49

4 RESULTS AND DISCUSSION 51
4 1 Display of PreS Regions on M13 51
4 1 1 Cloning into the pCANTAB 5E Phagemid Vector 54
4 1 2 Gene Expression and Analysis of Phage-Displayed and Soluble Proteins 59
4 2 Display of PreS Regions on T7 66
4 2 1 Cloning in T7 Select 415-1b Vector 68
4 2 2 Protein Expression in T7 Select415-1 Vector 75

5 SUMMARY AND CONCLUSION 78

REFERENCES 83

APPENDICES 92
A1 pCANTAB 5E vector map 92
A2 T7 Select415-1b vector map 93
A3 pCR2 1-TOPO vector map 94

BIODATA OF THE AUTHOR 95

LIST OF TABLES

Table		Page
1	Media, buffers and solutions	30
2	Bacterial strains and plasmid	31
3	Recipes for SDS-PAGE [15% (w/v)] gel preparation	39
4	DNA fragments generated by PCR	44
5	Oligonucleotides for amplification of S1-47 and S1-50	44

LIST OF FIGURES

Figure		Page
2 1	A schematic representation of the structure of HBV	6
2 2	The structural and functional features of HBV DNA	9
2 3	Schematic representation of the HBV envelope proteins	10
2 4	Topologies of S-, M- and L-HBsAg in the ER membrane	12
2 5	A schematic diagram of the filamentous bacteriophage particle	14
2 6	Life cycle of M13	16
2 7	The lytic infection cycle of bacteriophage T7	18
2 8	Types of phage display systems	22
2 9	A simple diagram showing the procedure in "biopanning"	24
4 1	The recombinant phage construct	53
4 2	Amplification of the PreS coding regions by PCR	54
4 3	RE analysis to confirm the presence of insert in the recombinant phagemid	57
4 4	Schematic diagram showing the map of the regions amplified by oligonucleotides pCANTAB 5E-S1/pCANTAB 5E-S6 and TWS1/TWS3	57
4 5	PCR verification of the putative recombinant plasmids	58
4 6	Chromatogram showing the sequencing result of recombinant plasmid pCANTAB 5E-PreS	59
4 7	SDS-PAGE of PreS-g3p fusion protein	61
4 8	Western blot of fusion phage protein	62
4 9	SDS-PAGE [15% (w/v)] of soluble PreS-E tag fusion protein prepared from periplasm, cytoplasm and medium	65

4 10 Western blotting of soluble PreS-E tag fusion protein in different cellular compartments 66

4 11 A schematic diagram of the T7 Select415-1b phage display vector 67

4 12 Amplification of the coding fragments of S1-47 and S1-50 by PCR 69

4 13 RE digestion of the coding fragments of S1-47 and S1-50 69

4 14 A diagram showing the map of the regions amplified by the T7 SelectUP and T7 SelectDOWN primers 71

4 15 PCR amplification of plaques to check for recombinant carrying the coding fragments of S1-47 and S1-50 71

4 16 S1-47 and S1-50 recombinant plasmids in pCR2 1 TOPO vector before and after being linearised with *Kpn*I 73

4 17 A schematic diagram showing the construction of the first and second half of PreS1 in T7 cloned into pCR2 1 TOPO vector 74

4 18 PCR amplification to confirm the presence of inserts in pCR2 1 TOPO vector 74

4 19 Sequencing result of S1-47 (a) and S1-50 (b) recombinant plasmids 75

4 20 Detection of proteins by SDS-PAGE [15% (w/v)] 76

4 21 Western blot of S1-47 and S1-50 fusion proteins displayed on T7 phage 76

ABBREVIATIONS

A_{600}	absorbance at wavelength 600 nm
ATP	adenosine triphosphate
β	beta
bp	base pair
BSA	bovine serum albumin
cDNA	complementary DNA
CITE	cap-independent translation enhancer
C-terminus	carboxy terminus
dNTP	deoxyribonucleotide phosphate
DHBV	duck hepatitis B virus
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DR	direct repeat
DTT	1, 4-Dithiothreitol
EDTA	ethylenediamine tetraacetic acid
ELISA	enzyme-linked immunosorbent assay
ER	endoplasmic reticulum
g3p, g6p, g7p, g8p, g9p	products of M13 genes 3, 6, 7, 8, 9, respectively
GSHV	ground squirrel hepatitis virus
HBcAg	hepatitis B core antigen
HBeAg	hepatitis B e antigen
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HBxAg	hepatitis B x antigen
HIV	human immunodeficiency virus
IgG	immunoglobulin G
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase
kDa	kilodalton
λ	lambda
LB	Luria broth
L-HBsAg	large surface antigen
M	Molar
mA	milliampere (10^{-3} A)
M-HBsAg	medium surface antigen
mRNA	messenger RNA
nm	nanometer (10^{-9} m)
N-terminus	amino-terminus
OD	Optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PEG	polyethylene glycol
pfu	plaque forming unit

pH	*Puissance hydrogene*
PreS1	N-terminal region of L-HBsAg comprising 108/119 amino acids
PreS2	region of L- and M-HBsAg comprising 55 amino acids
RF	replicative form
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
S-HBsAg	small surface antigen
TBS	tris-buffered saline
TCA	trichloroacetic acid
TE	tris-EDTA buffer
TEMED	tetramethyl ethylenediamine
TES	N-tris-(hydroxymethyl)-methylaminoethanesulfonic acid
TSS	transformation and storage solution
U	unit
UV	ultraviolet
vol	volume
v/v	volume/volume
WHV	woodchuck hepatitis virus
w/v	weight/volume
$\times g$	centrifugal force
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
YT	yeast tryptone

CHAPTER 1

INTRODUCTION

Hepatitis B viruses are a group of small enveloped hepatotropic partially double-stranded DNA viruses that cause acute and chronic infections in humans, mammals and birds (Mason and Seeger, 1991). Chronic infection of HBV is estimated to occur in ~300 million people world-wide, these people have an increase risk of developing cirrhosis and hepatocellular carcinoma. Thus, hepatitis B is a major world-wide health problem today (Hildt *et al*., 1996).

Despite the widespread use of effective vaccines based on surface antigens (HBsAg) derived from human plasma (Szmuness *et al*., 1980) or produced in yeast resulting from recombinant DNA approaches (Valenzuela *et al*., 1982; Murray *et al*., 1984), HBV infections are responsible for 1-2 million deaths annually (Mahoney, 1999). To date, only interferon alfa and lamivudine monotherapy have been approved in many countries for the treatment of chronic HBV infection (Lok *et al*., 1993). Moreover, the success rate for these treatments is less than 50%, therefore, development of an improved vaccine, a rapid and sensitive diagnostic test and antiviral compounds are greatly in need to control the disease.

The causative agent of hepatitis B, the Dane particle, consists of an inner nucleocapsid, comprising the core protein (HBcAg), viral polymerase and viral

DNA, surrounded by a membranous envelope containing viral surface antigens (HBsAg). There are three surface antigens: large (L), medium (M) and short (S), which share a common C-terminal region but have different N-termini, arising from variable use of different initiation triplets within a continuous open reading frame (Ganem and Varmus, 1987). The L-polypeptide consists of PreS1, PreS2 and S regions, whereas the M-polypeptide consists of PreS2 and S regions (Heermann *et al*., 1984).

All these surface antigens were shown to elicit virus-neutralizing and protective antibodies (Itoh *et al*., 1986). Antibodies directed against PreS region correlate with viral clearance and recovery from acute HBV infection and the antibodies response to PreS can overcome nonresponsiveness to S region in the currently available vaccine (Milich *et al*., 1985). It has been reported that inclusion of the PreS region into the commercially available vaccines could result in enhanced antibody response to S region and thus provide a more effective vaccine for HBV (Neurath *et al*., 1989). In addition, the PreS1 and PreS2 regions have been shown to play an important role in the attachment to HBV hepatocyte receptor (Neurath *et al*., 1986; Pontisso *et al*., 1989).

Earlier work has shown that L-HBsAg binds to the core protein *in vitro* (Dyson and Murray, 1995). Furthermore, using phage display libraries, peptides were selected, in which they bound to the core protein and blocked the binding of L-HBsAg to the core protein (Dyson and Murray, 1995). Different mutagenesis experiments

demonstrated that virion morphogenesis required the 17 C-terminal amino acids of the PreS1 region and the 5 N-terminal amino acids of the PreS2 domain of the L-HBsAg, exposed at the cytosolic face of the ER membrane (Le Seyec *et al*., 1998; 1999). Recently, mutagenesis studies have confirmed that the interaction between L-HBsAg with HBcAg is mediated through two distinct sites, one element contains the PreS domain and the other is composed of about the first two third of the S-HBsAg (Tan *et al*., 1999). Besides, point mutation studies showed that Arg 92 in the PreS domain plays a pivotal role in the interaction (Tan *et al*., 1999).

The broad diversity of peptides displayed on the surface of bacteriophage has made phage display a powerful tool in drug discoveries by affinity selection of specific ligands that interact with a particular target (Wilson and Finlay, 1998). Phage display differs from conventional expression systems, in that the foreign gene sequence is inserted into the gene encoding one of the phage coat proteins, so that the foreign amino acid sequence is genetically fused to the endogenous amino acids of the coat protein to make a hybrid "fusion protein" (Smith and Petrenko, 1997). The hybrid coat protein is incorporated into phage particles as they are released from the cell, so that the foreign peptide or protein domain is displayed on the outer surface.

The most commonly utilized vectors for phage display are based on filamentous bacteriophage (M13, fd, f1), where the foreign peptide can be fused to the amino terminus of the coat proteins 3 or 8 (Scott and Smith, 1990). Molecules that have

been displayed either fused to the g3p or g8p of the filamentous phage include peptides (Smith and Scott, 1993), constrained peptides (McLafferty *et al*., 1993), antibody-like molecules (Fab and single-chain; Winter *et al*., 1994), enzymes (Corey *et al*., 1993), enzyme inhibitors (Roberts *et al*., 1992) and products of cDNA libraries (Hottiger *et al*., 1995).

Recently, a number of other display systems have been developed using non-filamentous phage. These include the display of fusions to the bacteriophages lambda (Mikawa *et al*., 1996; Santi *et al*., 2000), T4 (Efimov *et al*., 1995) and T7 (Houshmand *et al*., 1999).

In order to study the development of a more effective vaccine and the receptor-ligand interaction, it would of useful to obtain high-level expression of the PreS region in bacteria but has not yet been reported. Therefore, the objective of this study was to clone the coding fragments of PreS and PreS1 regions and subsequently express them in a phagemid and T7 vector, respectively. The functional display of the PreS or PreS1 regions could serve as a useful tool to investigate further their interactions with the viral nucleocapsid or hepatocyte receptor. Identification of the lead compounds for therapeutic agents based upon inhibition of viral morphogenesis is a quantum leap towards the control of hepatitis B.

CHAPTER 2

LITERATURE REVIEW

2.1 Hepatitis B Virus

Hepatitis B virus (HBV) is a major cause of chronic inflammatory liver disease and liver cirrhosis associated with the development of hepatocellular carcinoma. HBV is an enveloped DNA virus of the hepadnavirus family (Ganem and Varmus, 1987). The other identified members of this family are the woodchuck hepatitis virus (WHV), ground squirrel hepatitis virus (GSHV), duck hepatitis virus (DHBV) and heron hepatitis virus (HHBV) (Seeger *et al*., 1991).

The host range of HBV is very restricted and it only infects human and chimpanzees (Gust *et al*., 1986). The routes of infection occur by both vertical and horizontal transmission of the virus (Gust *et al*., 1986). Vertical transmission occurs by the passage of HBV from an infected mother to her offspring while horizontal transmission happens during sexual contact and percutaneous contact or other close contact between the virus and man.

HBV infection is highly polymorphic, ranging from inapparent forms to acute hepatitis and severe chronic liver disease. The pathological consequences of the viral infection are unpredictable and the mechanism through which HBV enters

hepatocytes has not been resolved despite considerable understanding of the details of hepadnaviral genome replication (Ganem and Varmus, 1987; Nassal and Schaller, 1993; 1996).

2.1.1 Morphology and Genome Structure

Serum from individuals infected by HBV contains distinct forms of viral particles. Most of them are spherical or filamentous particles of about 22 nm in diameter (Bayer *et al*., 1968; Gerin *et al*., 1969). These subviral particles consist of a single viral envelope and are therefore not infectious. The infectious agents also known as Dane particles (Dane *et al*., 1970), are spherical 42 nm double-shelled particles. The Dane particle (Figure 2.1) consists of a nucleocapsid, comprising the core protein (HBcAg), viral polymerase and viral DNA, surrounded by a membranous envelope containing viral surface antigens (HBsAg).

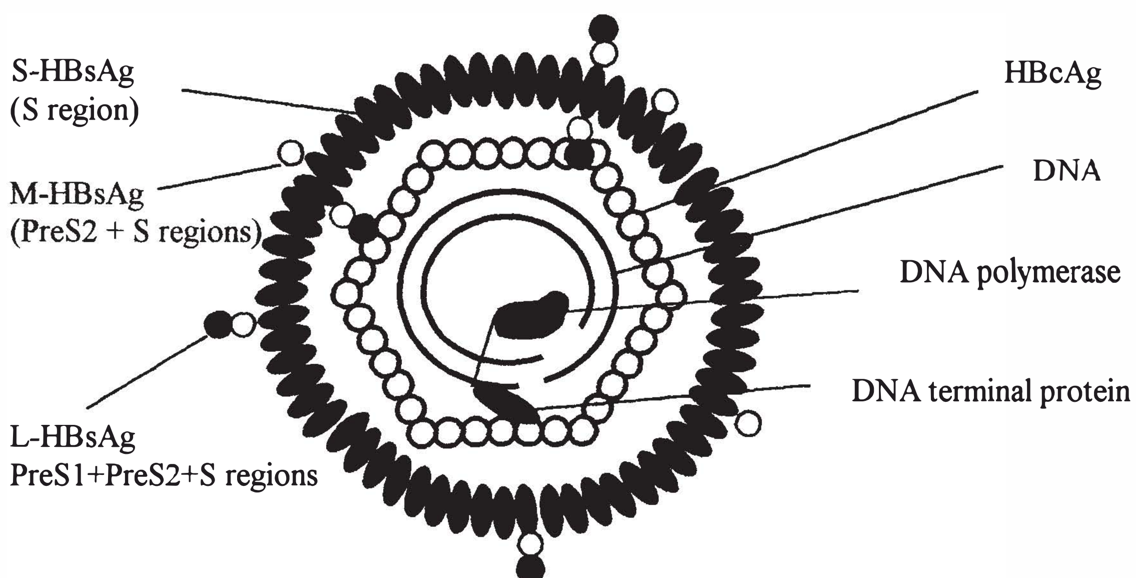


Figure 2.1: A schematic representation of the structure of HBV (modified from Nassal and Schaller, 1993).

The viral envelope contains three different, but related, HBsAg polypeptides, known as short (S), medium (M) and long (L) polypeptides, which are encoded in a single open reading frame of the viral genome by using three different in–frame start codons and a common stop codon (Heermann *et al*., 1984). The genome is a partially double-stranded circular DNA of only 3.2 kilobases (kb) in length, the smallest of any animal DNA virus yet encountered (Figure 2.2). The negative strand is linear and of fixed length of about 3.2 kb while the positive strand is of variable length. The maintenance of the circular structure of the genome is assured by base-pairing of the 5' ends of the two strands containing the viral direct repeat (DR) sequences. At both sides of the cohesive ends, there are 11-base pair direct repeats (DR1 and DR2), which are critically involved in the initiation of viral DNA synthesis (Ganem and Varmus, 1987).

2.1.2 Genomic Organisation and Viral Transcripts

The negative strand transcript of HBV contains four major open reading frames (ORFs): S, C, P and X (Figure 2.2). The coding region for HBsAg (ORF S) proved to be the 3' portion of a larger coding region: upstream of ORF S is an in-phase reading frame (ORF PreS) with two conserved in-phase ATG codons that can direct the synthesis of additional HBsAg related proteins which subdivide the PreS region into two functional subregions, termed PreS1 and PreS2. The coding region for HBcAg (ORF C) is also preceded by a short upstream in-phase ORF (termed ORF PreC), which produces a hydrophobic polypeptide bearing hepatitis B e

antigen (HBeAg) and could have a role in the attachment of core to the viral envelope. ORF X encodes a pleiotropic transcriptional activator of yet undefined function for the viral life cycle. Overlapping these coding regions, ORF P, is believed to encode the viral polymerase (Ganem and Varmus, 1987; Nassal and Schaller, 1993).

Four species of mRNA transcripts (Figure 2.2) are produced from a corresponding set of promoters on a covalently closed circular DNA utilizing the transcription machinery of the host cell. There are two classes of viral RNAs, genomic and subgenomic. The genomic RNAs are approximately 3.5 kb in length and their synthesis is controlled by the core (C) promoter (Yaginuma and Koike, 1989). These RNAs are bifunctional, serving not only as the mRNA for the precore, core and P proteins, but also as the template for reverse transcription in the DNA genome. The subgenomic RNAs include 2.4, 2.1 and 0.9 kb in length. The PreS1 promoter regulates the transcription of the 2.4 kb mRNAs which serve as a template for the L-HBsAg (Will *et al*., 1987). The 2.1 kb mRNAs encoded the M- and S-HBsAg which are the most abundant in acutely infected livers. Synthesis of the 2.1 kb mRNAs is controlled by the PreS2/S promoter (Raney *et al*., 1989). The smallest transcript, 0.9 kb, gives rise to the X protein, HBxAg. All these transcripts terminate at a common position, some 20 nucleotides downstream of the conserved hexanucleotide TATAAA (Ganem and Varmus, 1987).