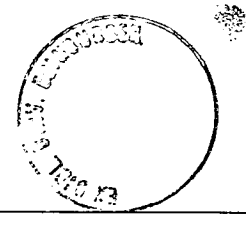
Interactions Between the Surface and Core Antigens of Hepatitis B Virus

by
Wen Siang Tan

A Thesis Presented for the Degree of Doctor of Philosophy
Institute of Cell and Molecular Biology
University of Edinburgh
November 1997



Dedicated to

Bee Kee and Jean Yin,

who always remind me to change nappies

DECLARATION

I hereby declare that the work presented in this thesis is the result of my own independent investigation unless otherwise stated and that this thesis has been composed by myself. The experiments presented were devised in collaboration with my supervisor Prof. Sir Kenneth Murray. The work has not been or is not being concurrently submitted for any other degree or professional qualification.

Edinburgh, November 1997

(Wen Siang Tan)

ACKNOWLEDGEMENTS

Over the past three years, from the first day my wife and I arrived in Edinburgh to the completion of my study, we met many wonderful people, who have been so helpful and caring. Their kindness will always deeply remain in our hearts.

First, I wish to express my most sincere acknowledgement to my supervisor, Prof. Sir Kenneth Murray for his valuable guidance, support and freedom throughout the entire process of research and writing. His passion in experiments had accompanied me through many quiet weekends and public holidays. Indeed, he helped to shape my career as a scientist. On behalf of my family, once again I would like to thank him for his care, concern and love.

I am greatly indebted to Dr. Mike Dyson, a humorous and talented young man, for his encouragement, advice and jokes whenever I faced difficulties in the experiments. I am also grateful to Dr. Graeme Reid for being my second supervisor and for helpful discussion. I wish to thank everyone in Prof. Murray's group: Mrs. Fiona Gray, Mrs. Sandra Bruce, Mrs, Jeanette Goman, Mrs. Suling Zhao, Ms. Helen Williamson, Ms. Nicola Preston, Dr. Volker Germaschewski and Dr. Sergei Petukhov for making my time in the laboratory an enjoyable one.

Many thanks to the following scientists whose efforts and generosity allowed me to employ a variety of different techniques to explore the secret of hepatitis B virus: Prof. Malcolm Walkinshaw and Ms. Jacqui Dornan (Department of Biochemistry, University of Edinburgh) for showing me the basic techniques of light scattering and crystallisation; Dr. Andy Cronshaw (Welmet Protein Characterisation Facility) for performing protein sequence analysis; Dr. Kenneth Zahn (Rockefeller University) and Dr. Arthur Landy (Brown University) for providing us with plasmids bearing the T4 AGA tRNA gene and; Dr. David Dryden (ICMB) for allowing me to use the GraFit™ software.

I am thankful to Mr. Frank Johnston and Mr. Graham Brown not only for their great photographic skill but also their jokes and laugh. I wish also to acknowledge everyone in the media room and washing up kitchen, particularly Ms. Jean Ramsay, Mrs. Joan Smail, Mrs. Aileen Greig and Mrs. Joan Davidson, and all the members of ICMB for creating such a pleasant environment to work in.

Special thank to my wife, Bee Kee, for her sacrifices, patience, motivation and love. I owe her more than she knows. I extend my thanks to our parents, brothers, and sisters, and Ms. Evelyn Lau for their constant love and support.

Finally, I would like to thank the Universiti Putra Malaysia for financing my study and family from September 1994 to September 1997, and the Roseanne Campbell Trust for Hepatitis Research for supporting me until I complete my thesis.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

CHAPTER 2: MATERIALS AND METHODS

CHAPTER 3: HBcAg PARTICLES: YIELD IMPROVEMENT, HETEROGENEITY AND FRACTIONATION

CHAPTER 4: MUTAGENESIS AND *IN VITRO* EXPRESSION OF SURFACE ANTIGEN

CHAPTER 5: CHARACTERISATION OF THE INTERACTION BETWEEN SURFACE AND CORE ANTIGENS

LIST OF ABBREVIATIONS AND SYMBOLS

A

α	alpha
A	Adenine
Å	Ångstrom unit (10^{-8} cm)
Amp	ampicillin
ATP	adenosine triphosphate

B

β	beta
bp	basepair
BSA	bovine serum albumin

C

C	cytosine / core
°C	degrees centigrade
Cam	chloramphenicol
CAPS	3-(cyclohexylamino)-1-propanesulfonic acid
ccc	covalently closed circular
cm	centimetre
Ci	curies
CIP	calf intestinal alkaline phosphatase
CITE	cap-independent translation enhancer
cpm	counts per minute
C-terminus	carboxy terminus
CTLs	$CD8^+$ cytotoxic T lymphocytes

D

ddNTP	dideoxy-nucleoside triphosphate
ddATP	2',3'-dideoxy-adenosine-5'-triphosphate
ddCTP	2',3'-dideoxy-cytidine-5'-triphosphate
ddGTP	2',3'-dideoxy-guanosine-5'-triphosphate
ddTTP	2',3'-dideoxy-thymidine-5'-triphosphate
DHBV	duck hepatitis B virus
DNA	deoxy-ribonucleic acid
dNTP	deoxynucleoside triphosphate
dATP	2'-deoxy-adenosine-5'-triphosphate
dCTP	2'-deoxy-cytidine-5'-triphosphate
dGTP	2'-deoxy-guanosine-5'-triphosphate
dTTP	2'-deoxy-thymidine-5'-triphosphate
dUTP	2'-deoxy-uridine-5'-triphosphate
D_T	translational diffusion coefficient
dUMP	deoxy-uridine-monophosphate

DR	direct repeat
DTT	1,4-Dithiothreitol
dsDNA	double-stranded DNA

E

ε	encapsidation signal
EDTA	ethylenediamine tetraacetic acid
EGTA	ethylenebis(oxyethylenenitrilo) tetraacetic acid
ELISA	enzyme-linked immunoabsorbent assay
EMC	encephalomyocarditis virus
ER	endoplasmic reticulum

F

Fig.	figure

G

g	gram
G	Guanosine
GSHV	ground squirrel hepatitis virus

H

h	hour
HBcAg	hepatitis B core antigen
HBeAg	hepatitis B e antigen
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HBxAg	hepatits B x antigen
HCC	hepatocellular carcinoma
HIV	human immunodeficiency virus

I

ICMB	Institute of Cell and Molecular Biology, University of Edinburgh
IPTG	isopropyl-β-D-thiogalactopyranoside

K

Kan	kanamycin
kb	kilobase
kDa	kilodalton
K_d	dissociation constant
K_d^{rel}	relative dissociation constant
kg	kilogram (10^3 g)

L

λ	lambda
l	litre

LB	Luria broth
L-HBsAg	large surface antigen
LMP	low melting point

M

η	viscosity
µg	microgram (10^{-6} g)
µl	microlitre (10^{-6} l)
µM	micromolar (10^{-6} M)
M	Molar
mA	miliampere (10^{-3} A)
MAb	monoclonal antibody
mg	miligram (10^{-3} g)
M-HBsAg	medium surface antigen
MHC	major histocompatibility complex
min	minute
ml	mililitre (10^{-3} l)
mm	milimetre (10^{-3} m)
mM	milimolar (10^{-3} M)
M_r	molecular weight
mRNA	messenger RNA

N

N	nonspecific base
NET-gel	sodium-Tris-EDTA-gelatin buffer
ng	nanogram (10^{-9} g)
nm	nanometre (10^{-9} m)
NMR	nuclear magnetic resonance spectroscopy
NP-40	nonidet P40
N-terminus	amino-terminus

O

OD	Optical density
ORF	open reading frame

P

P	polymerase protein
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate saline buffer
PEG	polyethylene glycol
pfu	plaque forming unit
pH	*Puissance hydrogene*
pmole	picomole
PreS1	N-terminal region of L-HBsAg comprising 108 or 119 amino acids
PreS2	region of L- and M-HBsAg comprising 55 amino acids
PVDF	polyvinylidene difluoride

R

RF	replicative form
RNA	ribonucleic acid
RNAsin	RNase inhibitor
rNTP	ribonucleoside triphosphate
rpm	revolutions per minute

S

s	second
s	Svedberg unit
S	surface protein
SDS	sodium dodecyl sulphate
S-HBsAg	small surface antigen
sos	sum of square
ssDNA	single-stranded DNA
STE	Sodium-Tris-EDTA buffer
SV40	simian virus 40

T

T	thymine / triangulation number
TBE	tris-buffered EDTA solution
TBS	tris-bufferd saline
TCA	trichloroacetic acid
TE	tris-EDTA buffer
TEMED	tetramethyl ethylenediamine
TES	N-tris-(hydroxymethyl)-methyl-2-aminoethanesulfonic acid
Tet	tetracycline
TP	terminal protein
tRNA	transfer RNA

U

U	unit
UV	ultraviolet

V

V	volt
vol	volume
v/v	volume/volume

W

W	Watt
WHV	woodchuck hepatitis virus
w/v	weight/volume

X

X	X protein
x*g*	centrifugal force
x-gal	5-bromo-4-chloro-3-indol-β-D-galactopyranoside

Y

Y	fraction bound
YT	yeast tryptone

AMINO ACID ABBREVIATIONS

Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cysteine	C	Cys
Glutamic acid	E	Glu
Glutamine	Q	Gln
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

ABSTRACT

The core antigen (HBcAg) of hepatitis B Virus (HBV) can be expressed in *Escherichia coli* where it assembles into icosahedral particles of two sizes containing 240 or 180 subunits. These particles could be fractionated by sucrose gradient centrifugation and light scattering showed their size distribution to be essentially monodisperse. HBcAg is 183 amino acids long and highly enriched in arginine residues in the C-terminal region. Around 50% of these residues are encoded by the rare triplet AGA in *E.coli*. Supplementation of the level of AGA tRNA in the cell with plasmids expressing the T4 AGA tRNA gene significantly enhanced the yield of HBcAg. SDS polyacrylamide gel electrophoresis of the two kinds of particles showed that around half of their subunits were smaller than the full length HBcAg and varied in size. N-terminal sequence analysis revealed that these smaller species were heterogeneous at the extremely basic C-terminal end.

In the virus the icosahedral nucleocapsids are surrounded by an envelope consisting of cellular lipids and three related surface antigens (L-, M-, and S-HBsAg) which result from alternative translation initiation of a common reading frame. The L-HBsAg is believed to mediate the contact between the envelope and nucleocapsid. The N- and C-termini of this protein were shortened in order to define the minimum stretch of amino acids that contains the exact contact residues. To determine which residues are directly involved in the interaction, single and multiple mutations were generated by site-directed mutagenesis. The resulting mutated proteins were expressed in rabbit reticulocyte lysates and their ability to interact with HBcAg was examined with an immunoprecipitation assay and a newly established equilibrium binding assay in solution which allows the determination of relative dissociation constants with Scatchard and non-linear regression analyses.

Binding of each type of HBcAg particles by L-HBsAg displayed two widely differing dissociation constants, suggesting the existence of two binding sites between the molecules. It appears that amino acids 24 to 191 of the L-HBsAg N-terminus form one intact binding domain and a second binding domain between residues 191 and 322 is required for the protein to bind with high affinity to HBcAg. Disruption of either one of these domains resulted in a weak interaction showing

only one dissociation constant. Mutation analysis revealed that Arg-92 was very important for the interaction. A synthetic peptide carrying the epitope for a monoclonal antibody to the PreS1 region competed with L-HBsAg for HBcAg, but peptides corresponding to a linear sequence at the tip of the nucleocapsid spike did not, showing that the competing peptide does not resemble the tip of the spike.

CHAPTER 1
INTRODUCTION

1.1 Hepatitis B virus

Human hepatitis B virus (HBV) is the prototype of the family of *hepadnaviridae*. Other family members have been found in Pekin ducks (DHBV), ground squirrels (GSHV), herons (HHBV), and woodchucks (WHV) (Seeger *et al.*, 1991). HBV has a narrow host range, infecting only human and other higher primates such as chimpanzees, but does not infect baboons, lower primates or other mammals. Until recently, this virus could not be propagated in cell culture systems.

HBV is transmitted in the same way as HIV, through direct contact with infected blood or body fluids, through perinatal transmission from an infected mother to her offspring. After an initial acute infection, a proportion of patients who fail to eliminate the virus become chronic carriers. The risk of becoming a carrier is strongly depends upon the age at which a person is infected. Infants who are infected at or shortly after birth have a greater than 85% chance of becoming carrier (Stevens *et al.*, 1985); however, this incidence declines to below 19% after the age of five. Currently, there are estimated to be more than 300 million carriers of HBV world-wide, particularly in the Southeast Asia and Africa. They have an increased risk of developing chronic liver disease such as cirrhosis and hepatocellular carcinoma (Dudley *et al.*, 1972; Redeker, 1975). The exact role of this virus in the development of hepatocellular carcinoma cells is not known, but these cells frequently contain integrated HBV DNA sequences. This suggests that integration of the HBV genome might alter the transcription of a cellular gene regulating cell division which culminates in the development of liver cancer (Dimmock and Primrose, 1994).

The infection can now be prevented through immunization with vaccines based on surface antigens (HBsAg) derived from human plasma (Szmuness *et al.*, 1984) or produced in yeast resulting from genetic engeneering approaches (McAleer *et al.*, 1984; Murray *et al.*, 1984). For those who had already acquired a chronic infection, α-interferon appears to hold the best hope to clear the virus (Thomas and Scully, 1985). This protein acts by enhancing the expression of major histocompatibility (MHC) class I by infected liver cells and not through an antiviral

activity. HBV-specific peptides are then presented in sufficient amount by MHC class I to allow $CD8^+$ cytotoxic T lymphocytes (CTLs) to lyse the infected cells (Dimmock and Primrose, 1994). The successful rate of this treatment is about 50% (Thomas and Scully, 1985).

1.2 Virion and genome structure

With the electron microscope, three morphological distinct forms of particle are observed in infectious sera: spheres and filaments of 22 nm diameter (Bayer *et al.*, 1968; Gerin *et al.*, 1969) and 42 nm double-shelled particles or known as Dane particles (Dane *et al.*, 1970).

The Dane particle (Fig. 1) is now known to be the infectious agent (Robinson and Luttwick, 1976). It consists of an outer envelope derived from the host cell membrane. Embedded in the envelope are three distinct but related forms of surface protein (HBsAg): S- (small), M- (middle) and L- (large) HBsAg. Internal to the envelope is the viral nucleocapsid or core (HBcAg). Within the core is the genome and the polymerase protein (P) which is covalently attached to the negative strand DNA (Ganem, 1991; Nassal and Schaller, 1993).

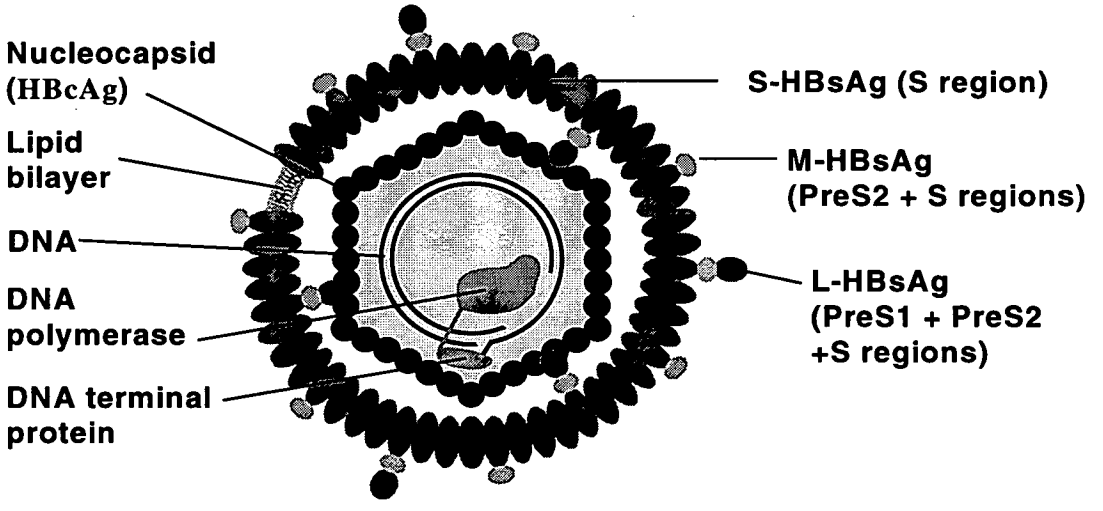


Fig. 1: A schematic representation of virion structure of HBV

The representation of the L, M, and S forms of HBsAg have no quantitative or positional significance. (modified from Nassal and Schaller, 1993).

The genome appears in the form of partially double-stranded circular DNA about 3.2 kb (Fig.2), the smallest of any animal DNA virus yet encountered. The DNA strands differ in size. The negative strand is about 3.2 kb long and serves as a template for transcription of all viral mRNAs. The positive DNA strand is of varying length, from 1.7 to 2.8 kb long. Neither of these DNA strands is covalently closed, but the circular configuration of the double-stranded DNA is maintained by hybridisation of complementary sequences at the ends of the long and short DNA strands which results in a molecule with a single-stranded gap. Near to the 5'-end of the two DNA strands is a pair of directly repeated 11-base-long sequences (DR1 and DR2) which play crucial roles in replication (Ganem and Varmus, 1987).

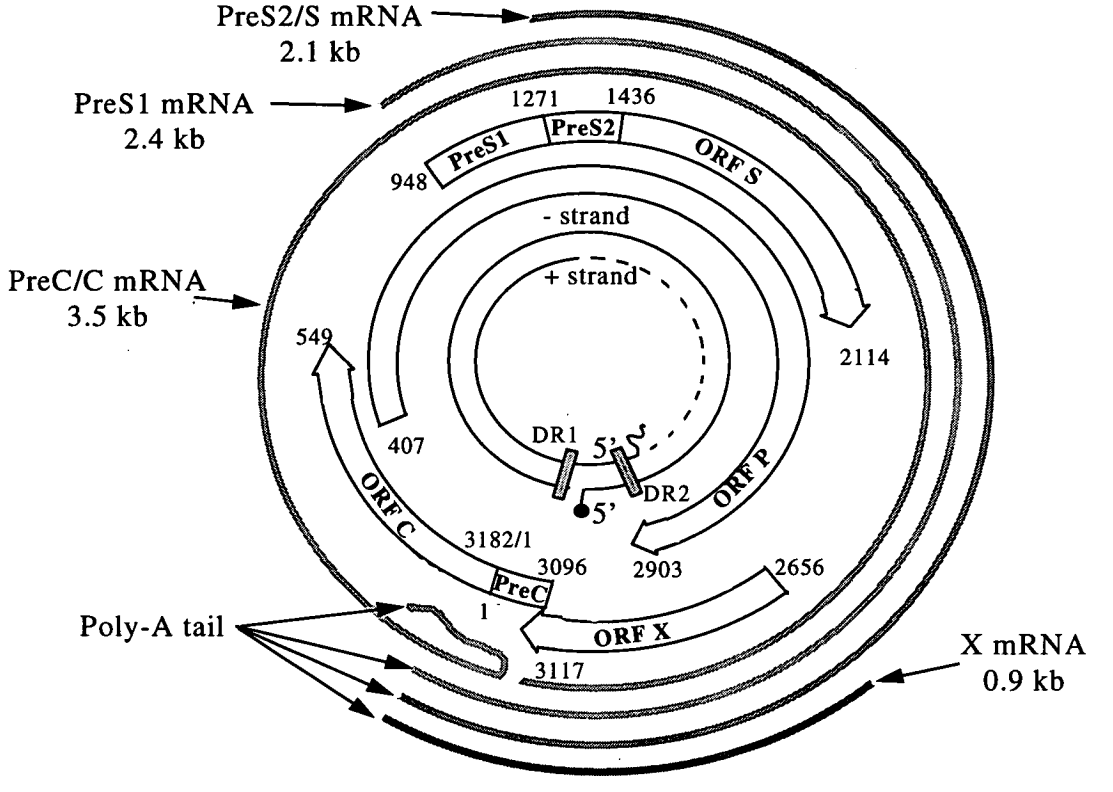


Fig. 2: The Stuctural and Functional Features of HBV DNA

The minus and plus strands of HBV DNA are shown in the innermost portion of the figure. The viral direct repeat (DR) sequences are boxed and the P protein is indicated by (•). Dashed line indicates the presence of the single-stranded gap. The viral open reading frames (ORFs) are indicated by arrow bars, with arrowheads indicating direction of transcription and translation. Major viral transcripts are indicated as bold lines in the outermost portion of the figure. (modified from Ganem and Varmus, 1987; Schaller and Fischer, 1991).

The 22 nm particles, which are enormously abundant in the blood stream of infected carriers, are composed of HBsAg and host-derived lipid which consists mainly of phospholipids, free and esterfied cholesterol and small quantities of triglycerides. No viral genetic information is found in these particles and therefore they are noninfectious. Also, these particles lack the nucleocapsid component and have no detectable cellular polypeptides (Ganem, 1991). The biological function of these noninfectious particles is unclear but they may weaken the efficacy of the immune response by acting as decoys that trap antivirus antibodies. They are highly immunogenic and therefore have been exploited as a subunit vaccine against HBV infection (Szmuness *et al.*, 1980).

1.3 Genome organisation and viral transcripts

The HBV genome carries four major open reading frames (ORFs) which overlap each other extensively and are encoded by minus-stranded DNA: ORFs S, C, P and X (Fig. 2). The nucleotide numbering system used in this study is that of Pasek *et al.* (1979) which select the first nucleotide of the coding sequence for HBcAg as number 1. Upstream of ORF S which encodes for HBsAg, is an in-phase reading frame (ORF PreS) with two conserved in-phase ATG codons which subdivide the PreS regions into two functional subregions, termed PreS1 and PreS2, whose products are known as L- and M-HBsAg. Similarly, the coding region for HBcAg (ORF C) is also preceded by a short upstream in-phase ORF, termed PreC, which produces an amino-terminally extended protein that is processed and secreted as HBeAg; this is believed to be involved in the virus-host interaction by influencing the immune response. ORF X encodes a product that is believed to be important for the establishment of infection by an unknown mechanism. Overlapping these coding regions, ORF P, encodes protein P which has polymerase, reverse transcriptase and RNAse activities (Ganem and Varmus, 1987; Ganem 1991; Nassal and Schaller, 1993).

Four species of mRNA transcripts (Fig. 2) are produced from a corresponding set of promoters on a covalently closed circular DNA (cccDNA) utilising the transcription machinery of a host cell. The RNA species can be divided

into two classes: 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 bifuntional, serving not only as the mRNA for the precore, core and P proteins, but also as the template for reverse transcription into the DNA genome. The subgenomic mRNAs consist of three species with 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 M- and S-HBsAg are encoded by the 2.1 kb mRNAs 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 are capped, unspliced and terminate at a common polyadenylation site, about 20 nucleotides downstream of the conserved hexanucleotide TATAAA (Ganem and Varmus, 1987; Schaller and Fischer, 1991; Nassal and Schaller, 1993).

1.4 Viral proteins

1.4.1 Core antigen (HBcAg)

The nucleocapsid or core particle of HBV is a protective container for the viral genome and polymerase. Its interior provides a suitable environment for the process of reverse transcription during replication. Furthermore, its exterior interacts with the viral envelope proteins in an unknown manner leading to maturation of the virus. The capsid is made up from many copies of core proteins; serologically known as core antigen (HBcAg) which is translated from the second in-frame initiation codon (AUG) of the C ORF and comprises 183 amino acids with a molecular mass of approximately 22 kDa. It is organised in two distinct domains: the first 144 amino acids are required for self-assembly; the C-terminal portion of the molecule (residues 145-183) is extremely basic, displaying non-specific nucleic acid binding and represent a motif required for packaging of the pregenomic RNA (Ganem, 1991).

Several groups have demonstrated that nucleocapsids isolated from infected liver (Petit and Pillot, 1995; Machida *et al.,* 1991), and produced in *E. coli* (Matsuda *et al.,* 1988; Gallina *et al.,* 1989; Hatton *et al.,* 1992) have non-specific nucleic acid binding activity. A detail dissection of the protamine-like region has revealed four repeated sequence motifs (Hatton *et al.,* 1992). The first motif, spanning amino acids 150-157 (RRRDRGRS) is the minimal region required for RNA binding and packaging, but does not bind DNA, while the three other repeats bind DNA much more efficiently than RNA. Nassal (1992) showed that only half of the C-terminal region spanning to amino acid 164 was sufficient for pregenomic RNA encapsidation with an efficiency similar to that of the full-length HBcAg. In addition, amino acids 165-173 of the C-terminus were required for proper genome replication. This suggest an extensive role of the protamine-like domain, not only in pregenomic RNA encapsidation, but also in genome replication.

HBcAg can be expressed in a variety of heterologous systems, including *E. coli* (Pasek *et al.,* 1979; Burrell *et al.,* 1979; Edman *et al.,* 1981; Stahl *et al.,* 1982), tissue culture cell lines (Gough and Murray, 1982; Rossinck *et al.,* 1986), yeast (Kniskern *et al.,* 1986), baculovirus (Hilditch *et al.,* 1990), *Acetobacter methanolicus* (Schroder *et al.,* 1991) and *Xenopus oocytes* (Zhou and Standring, 1991). In these systems, core proteins are able to self-assemble into nucleocapsids in the absence of other viral proteins. When observed in an electron microscope, such particles are morphologically similar to authentic nucleocapsids isolated from infected liver (Cohen and Richmond, 1982; Uy *et al.,* 1986; Miyanohara *et al.,* 1986; Ou *et al.,* 1986). Deletion mutagenesis of HBcAg indicates that the first 140 amino acids are sufficient for self assembly (Zlotnick *et al.,* 1996).

To date, the three dimensional structure of HBcAg particle or its subunits at atomic resolution has not been solved. Perhaps, purified HBcAg particles with heterogeneous populations, as shown by electron microscopy (Cohen and Richmond, 1982; Onodera *et al.,* 1982) do not allow the particles to crystallise easily. Usually, most viral capsid proteins assemble into particles of a fixed, predetermined size, but HBcAg particles are dimorphic. Using electron cryomicroscopy and image reconstruction Crowther *et al.* (1994) showed that HBcAg expressed in *E. coli* assemble into two sizes of particles: small and large, approximately 30 and 34 nm in

diameter, respectively (Fig 3). The large particle corresponds to a triangulation number T=4 containing 240 protein subunits. The minority species (less than 15% of the total population), comprising the smaller size of particle, corresponds to a T=3 shell of 180 monomers.

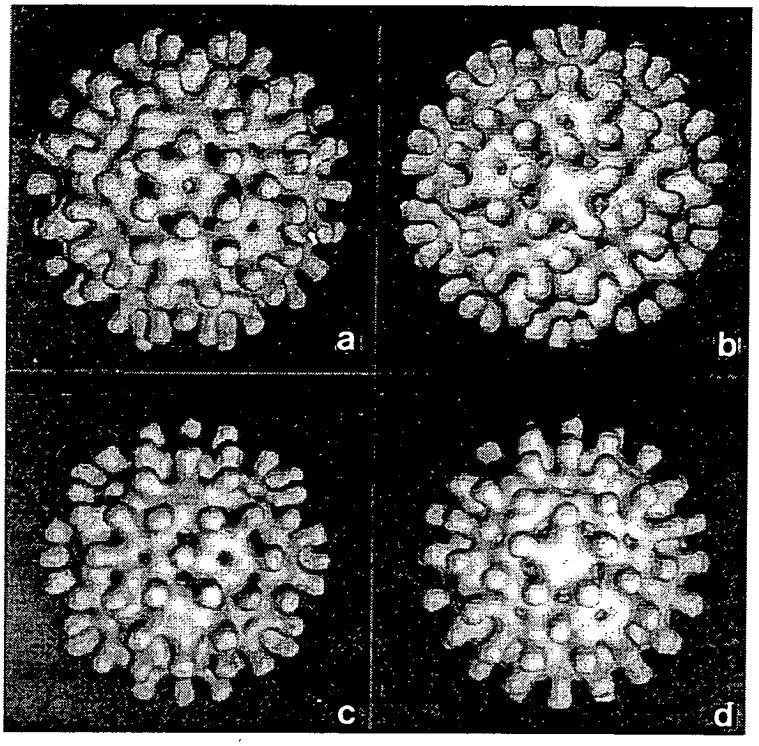


Fig. 3: Three-dimensional maps of hepatitis core particles

Large empty (a and b) and small empty (c and d) core particles determined by electron cryomicroscopy and image processing. a and c show the view along a two-fold axis of symmetry, b and d show the view along a five-fold axis (Source: Crowther et al., 1994).

Nucleocapsids isolated from human liver show a ratio of 13:1 of large to small particles (Kenney *et al.*, 1995). Recently, Zlotnick *et al.* (1996) showed that the proportions of small particles isolated from *E. coli* increased systematically with larger deletions from the protamine-like region; a mutant comprising 149 residues formed approximately 5% of small particles, while a shorter mutant containing 140 residues produced approximately 80% of small particles. However, a C-terminally truncated mutant at residue 144 with several additional amino acids extending its N- and C-termini also produced predominantly (approximately 85%) small particles. In

addition, when cysteine 61 of the 149-residue mutant was mutated to alanine, the amount of small particles increased to around 30% (Zlotnick *et al.*, 1996). This observation indicates that the C-terminus is not the only factor to affect the dimorphism. The factors which control the dimorphic distribution are not understood. Furthermore, the physiological implications of the two different species of core particles *in vivo* remain an enigma.

HBcAg particles produced in *E. coli* can be dissociated into dimers without unfolding the protein and reassembled under appropriate conditions (Wingfield *et al.*, 1995; Zlotnick *et al.*, 1996). This shows that the fundamental building block is a dimer, and the residues involved in dimer linkage are Cys-48 and Cys-61 (Zheng *et al.*, 1992). Each dimer consists of a protruding spike that stick out from the underlying shell domain. The large and small particles show 120 and 90 spikes, respectively (Crowther *et al.*, 1994). The spikes are similar in structure for the two types of particles which are shaped like an inverted capital 'T' (Zlotnick *et al.*, 1996). The stem of the T's form the external spikes and from the calculated volume, the authors estimated that the spikes received a contribution of about 35 amino acids from each monomer. Most recently, using electron cryomicroscopy, Böttcher *et al.* (1997) solved the structure of the large particle to 7.4 Å resolution, and at the same time Conway and colleagues (1997) published a 9 Å-resolution map. Both groups showed that each dimer consists of a 4-helix bundle, formed by the pairing of α-helical hairpins from both subunits. A diagram of the polypeptide fold of the monomer subunit as proposed by Böttcher *et al.* (1997) is shown in Fig. 4. It appears that the immunodominant loop of amino acids 74-89 in the core protein (Argos and Fuller, 1988; Salfeld *et al.*, 1989; Borisova *et al.*, 1993) is exposed near the tip of these protruding domains. Since this region is nearest to the viral envelope, it is very likely to be involved in the interaction with the envelope proteins. In principle, this could be established by biopaning the HBsAg with a bacteriophage display library (Scott and Smith, 1990) and the peptide sequences obtained should reflect the docking site on the core protein.

With three dimensional image reconstruction, it is clear that the assembled particles have holes penetrating the protein shell (Crowther *et al.*, 1994; Kenney *et al.*, 1995; Zlotnick *et al.*, 1996). At 17 Å resolution, the largest holes on T=4

particles are measured to be around 500 Å^2. Since the HBV capsid serves as a compartment for reverse transcription of pregenomic RNA to double-stranded DNA, it is important for the inside of the core to be accessible to the small molecules necessary for DNA synthesis (Crowther *et al.*, 1994).

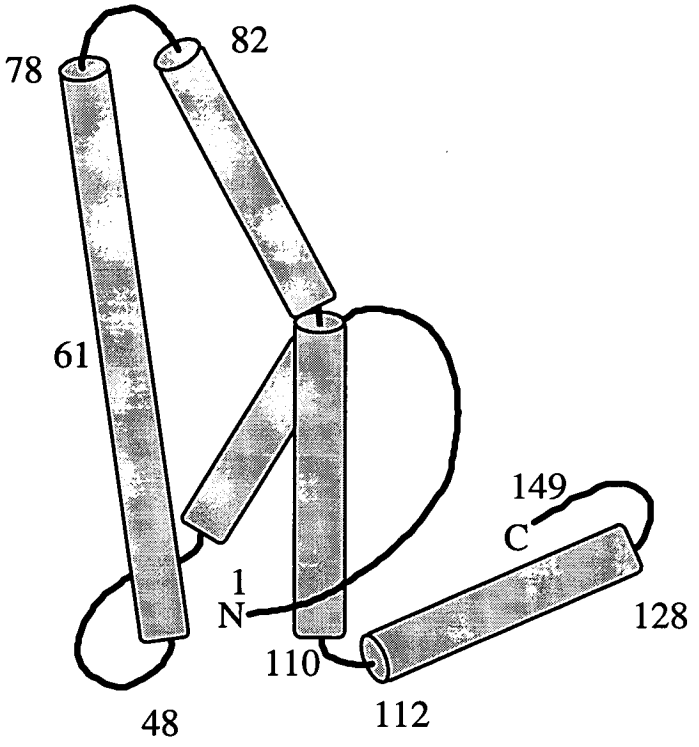


Fig. 4: Polypeptide fold of the HBcAg monomer

Each monomer consists of 4 α-helices (indicated as cylinders). The α-helix near to the N-terminus is ~15 Å long. The two longest α-helices are ~42 Å long and joined by a loop corresponding to amino acids 78-82. The first long helix is roughly straight, whereas the second long helix has a kink in the middle. The C-terminal helix corresponds to residues 112-128 is ~23 Å long. The numbers indicate approximate amino asid positions in HBcAg and putative amino and carboxy ends are indicated by N and C respectively. (modified from Böttcher *et al.*, 1997).

A striking feature of HBcAg is that the extremely basic C-terminal region can be replaced by foreign sequences of varying length, without disruption of the ability to form particles (Borisowa *et al*, 1989; Stahl and Murray, 1989; Schoedel *et al.*, 1992). Those fused foreign peptides, in some cases, can be detected by specific antibodies on the outside of the assembled particles. Since the arginine-rich C-terminus has been shown to bind nucleic acids and play a role in genome replication,

therefore, it is likely that the inserted foreign sequences emerge through or located near the edge of the holes in the assembled particles, and are thus accessible to antibodies (Crowther *et al.,* 1994; Kenney *et al.,* 1995). Another explaination is that these particles could be less stable, disrupted in the reaction conditions used and exposed to antibodies. Image reconstruction from cryoelectron micrographs of those assembled particles could be a straightforward technique to prove those suggestions.

The assembly of core-like particles is also unimpeded by the addition of heterologous sequence to the N-terminus of HBcAg (Clarke *et al.,* 1987; Stahl and Murray, 1989). However, certain amino acids in this region are important for the formation of correct determinants. Deletion of the first eleven amino acids of HBcAg resulted in a drastic loss of antigenicity (Salfeld *et al.,* 1989). Similarly Stahl *et al.* (1982) demonstrated that particles, produced in *E. coli,* which have eleven foreign amino acids fused to residues 3-183 of HBcAg reacted well with polyclonal anti-HBcAg while a protein which differed only in the deletion of amino acids 3 and 4 of HBcAg failed to react significantly with antibody. Further investigation by Stewart (1993) showed that protein lacking amino acid 3 reacted less well with polyclonal anti-HBcAg, relative to wild-type.

1.4.2 HBeAg

Upstream of the ORF C lies another in-frame initiation codon of the pre-C region. Translation of the complete pre-C/C gene produces a 25 kDa precore protein, p25, which has an additional of 29 amino acids at the N-terminus of HBcAg. The first 19 amino acids act as a signal sequence which targets p25 into membranous organelles, probably endoplasmic reticulum and Golgi (Ou *et al.,* 1986). The protein is then proteolytically processed at both ends and eventually secreted in the sera of infected individuals as a soluble 17 kDa protein, p17e, which known as hepatitis B e antigen (HBeAg). It consists of the first 149 amino acids of core polypeptide, preceded by 10 amino acids remaining from the pre-C sequence (Takahashi *et al.,* 1983; 1991).

Although HBeAg and HBcAg share a long fragment of identical primary sequence, these proteins show several distinct properties: HBeAg is nonparticulate, whereas HBcAg self-assembles into core particles; HBcAg particles induce a rapid, strong and long lasting humoral immune response in almost all infected individuals. In contrast, HBeAg induces a distinct, weaker, and delayed response giving anti-HBe antibodies, which usually correlate with viral clearance (Hoofnagle *et al.*, 1981). The biophysical and serological distinctions between HBeAg and HBcAg indicate that these antigens adopt different conformations. Nassal and Rieger (1993) demonstrated that Cys-7, located in the 10-amino acid N-terminal extension, forms an intramolecular disulphide bond with Cys-61. This intramolecular disulphide bridge keeps HBeAg in a specific conformation that prevents dimerisation and consequently particle formation, explaining the nonparticulate nature of HBeAg. Mutation of Cys-7 to Ala or Ser produces dimeric HBeAg exhibiting both HBe and HBc antigenicity.

1.4.3 Surface antigen (HBsAg)

HBV encodes three related surface antigens (HBsAg) within one single ORF by using three different in-frame start codons and a common stop codon (Fig. 5). Hence, the proteins differ at their N-termini, but have a large sequence in common. Initiation at the 5'-most AUG results in the production of the largest of the three proteins, the L-HBsAg (39 kDa), which contains 108 or 119 residues (depending on serotype) encoded in the PreS1 region as well as the 55 PreS2 amino acids added to the N-terminus of the S region which comprises 226 amino acids (Heermann *et al.*, 1984). The second largest protein, the M-HBsAg (31 kDa), arises from translation beginning at the middle initiation codon and gives N-terminal extension of 55 amino acids termed PreS2 (Machida *et al.*, 1984). The smallest of these polypeptides, S-HBsAg (24 kDa), begins at the innermost initiation codon and is the most abundantly expressed of the HBV surface glycoproteins. All these proteins exist in two isomeric forms, either glycosylated at Asn-146 of the S region or unglycosylated at this site. The M-HBsAg is additionally glycosylated at Asn-4 of the PreS2 region.

A myristic acid group is amide-linked to the N-terminal glycine residue of L-HBsAg (Persing *et al.*, 1987).

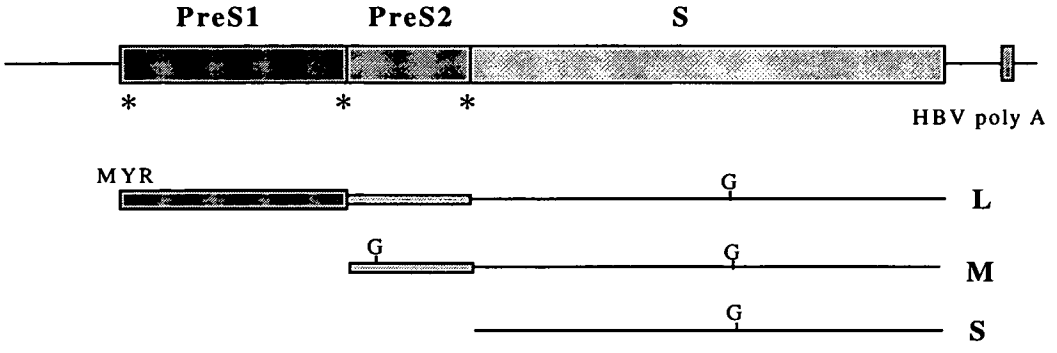


Fig. 5: Envelope Proteins of HBV

The translation products of the HBsAg gene are shown as lines of different tickness. S-, M- and L-HBsAg are translated from a common open reading frame of the HBsAg gene by the use of three in-frame initiation codons (*) at the N-termini of PreS1, PreS2 and S. Glycosylation is indicated by G. MYR represents the myristic acid.

The S-HBsAg is unique among viral envelope proteins in that it is capable of self-assembly with host-derived lipids to form empty spherical particles 22 nM in diameter (Liu *et al.*, 1982). This suggests that all the information necessary for this assembly process resides in the S region. Formation of the 22 nm particles is initiated by insertion of the S-HBsAg into the endoplasmic reticulum (ER) membrane. After aggregation of approximately 100 transmembrane monomers, assembled complexes are believed to mature by budding into the lumen of the pre-Golgi compartment, where their mannose asparagine-linked carbohydrates are processed to the complex, endoglycosidase H-resistant form (Patzer *et al.*, 1984; Huovila *et al.*, 1992).

S-HBsAg expression is not sufficient, but is required for the envelopment of nucleocapsids. L-HBsAg, however, has been shown to be absolutely necessary for virion formation (Bruss and Ganem, 1991a; Ueda *et al.*, 1991). M-HBsAg is dispensable for formation of both virion and 22 nm subviral particles (Fernholz *et al.*, 1993; Bruss and Ganem, 1991b). Assembly of HBV virions is presumed to proceed along a similar pathway to that for 22 nm particles. The kind of particle that

is formed seems to be determined by the ratio of S- and L-HBsAg proteins coassembling during morphogenesis. The 22 nm spherical particles from serum contain only traces of L-HBsAg. A higher proportion of the L-HBsAg coassembling with S-HBsAg is found in the filamentous form of HBsAg particles. A still higher proportion of L-HBsAg is apparently necessary for virion formation (Heermann *et al.*, 1984). However, overexpression of the L-HBsAg leads to the inhibition of virion formation (Cheng *et al.*, 1986; Chisari *et al.*, 1986).

The envelope proteins are synthesised in the ER where they gain a defined transmembrane topology. The S-HBsAg is believed to traverse the ER membrane at least twice. Four transmembrane α-helices have been predicted in this polypeptide, but only the two N-terminal domains have been shown to be directly inserted in the membrane (Fig. 6) (Eble *et al.*, 1987; Prange *et al.*, 1992). These two domains corresponding to amino acids 11-28 and 80-98, respectively, are separated by a hydrophilic loop of 55 amino acids which is exposed on the cytoplasmic side of the membrane (Bruss *et al.*, 1994). The second hydrophilic region of 70 amino acids which carries the major surface antigen and a glycosylation site, is predicted to be on the luminal side. Both the N and C-terminal sequences are luminally exposed. The topology of the M-HBsAg is believed to be very similar to that of the S-HBsAg. The PreS2 region does not contain a signal sequence, but the M-HBsAg is translocated into the ER lumen by the first signal in its S region (Eble *et al.*, 1990). Consequently, the PreS2 region and the major epitope regions of S and M are exposed on virions.

The transmembrane topology of the L-HBsAg is still unclear. PreS regions of the L-HBsAg are involved in the envelopment of viral nucleocapsids and binding to the host cell, so the PreS regions of the L-HBsAg must be both luminally and cytoplasmically exposed. Ostapchuk *et al.* (1994) showed that PreS regions are cytoplasmically disposed when synthesised in a cell free system. Therefore, PreS regions could be transported across the membrane to fulfil a dual function in virion assembly and attachment to the host cell. Two possibilities have been suggested: (i) the PreS regions are post-translationally translocated across the bilayer following interactions with core or; (ii) the membrane bilayer itself is attenuated or reorganised during budding to expose the PreS regions (Ostapchuk *et al*, 1994). Recently,

Prange and Streeck (1995) demonstrated that a substantial alteration in the transmembrane conformation of the L-HBsAg occurs post translationally, resulting in two topologically different populations of L-HBsAg which dispose their PreS regions both to the cytoplasmic and the luminal side of the membrane. Moreover, they found that the S region of the L-, M- and S-HBsAg is not uniformly oriented in the membrane of microsomes, as assayed in a cell free system. This is inconsistent with the models proposed in Fig. 6.

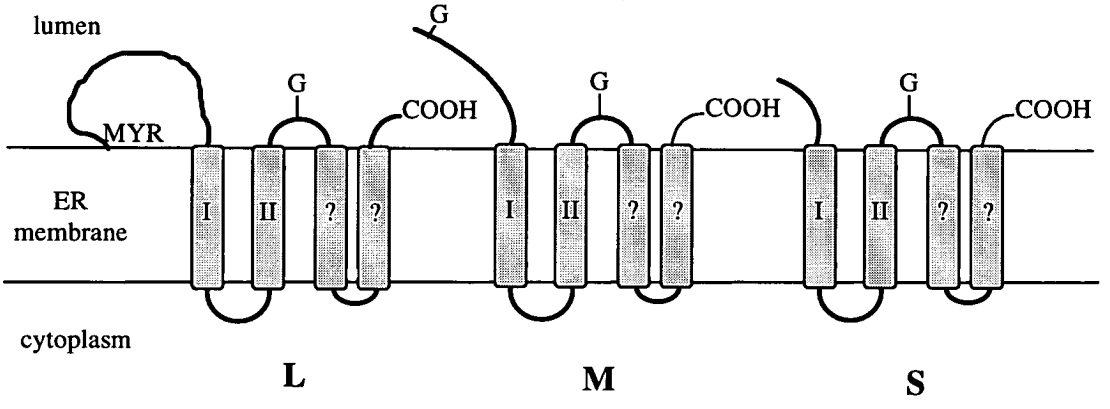


Fig. 6: Topologies of S-, M- and L-HBsAg in the ER Membrane

The N- and C-terminal ends of the proteins are lumenally disposed as well as the entire PreS regions of M- and L-HBsAg. I and II indicate the membrane spanning helices. The spanning helices marked with a '?' are based on model building but have not been examined experimentally. Sites where asparagine-linked glycosylation occurs are indicated by G. A myristic acid group (MYR) is amide-linked to glysine at the N-terminus of L-HBsAg. (modified from Ostapchuk et al., 1994).

Exposure of the PreS regions on the cytosolic side of ER is important for virion formation, perhaps forming a docking site for the nucleocapsids. Recent mutation studies by Bruss and Thomssen (1994) showed that the first 102 amino acids of the PreS1 region (119 amino acids; HBV subtype *adw*) are dispensable for virion formation. This suggests that the sequence between amino acid 103 in the C-terminal region of the PreS1 region and amino acid 254 upstream of Signal II in the S region might form a docking site for nucleocapsids. More recently, Dyson and Murray (1995) used a random hexapeptide library (Scott and Smith, 1990) to explore interactions between the HBcAg and L-HBsAg. Filamentous bacteriophages

displaying random peptides on their gpIII proteins were used to screen for binding to HBcAg. Bound phages were released and amplified via infection of *E. coli*. After several rounds of affinity selection and amplification, a subset of the selected phages were grown up individually and the identity of the peptides that bind to HBcAg were obtained by sequencing the gpIII gene carrying the insertion. A peptide sequence of LLGRMK and some related sequences were selected from the fusion phage library. However these specific sequences were not found as a contiguous sequence within the sequence of HBsAg, but two related sequences did occur (Fig. 7): residues 63-65 (LLG) in PreS1 region and residues 21-26 (LLTRIL) of the S region. These suggest that amino acids of the hexapeptide and other related sequences may be brought together in the folded protein of HBsAg to form a mimotope. Moreover, the interaction of L-HBsAg with HBcAg in a cell free system was inhibited by the peptide ALLGRMKG. This information provided the basis for the mutational analysis carried out in this project to define the exact docking sites on L-HBsAg. It is hoped that with the knowledge of these important sites, inhibiting molecules can be designed to block the assembly of HBV *in vivo*.

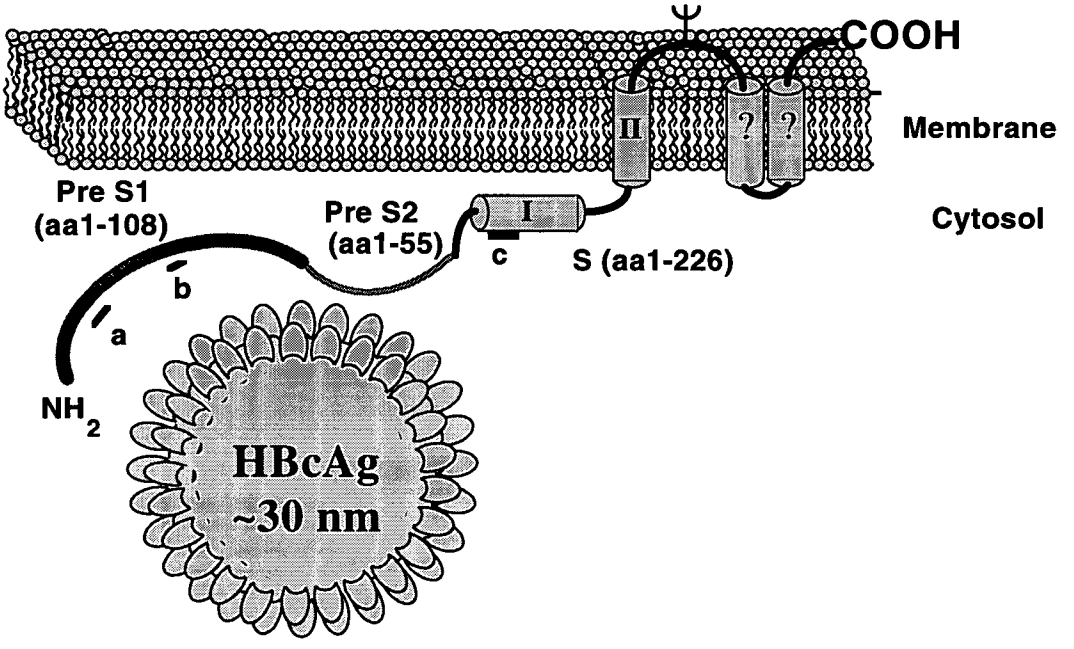


Fig. 7: Domain structure of L-HBsAg and association with HBcAg

a, b and c indicate the peptide sequence LDPAFR (19-24), LLG (63-65) of PreS1 region and LLTRIL (21-26) of S region respectively. (modified from Dyson and Murray, 1995).

1.4.4 P protein

The P protein is synthesised from the second cistron of the genomic RNA. The P ORF shares homology with the polymerase genes of retrovirus and other genetic elements that employ reverse transcription (Toh *et al.*, 1983; Khudyakov and Makhov, 1989). Mutational analysis and sequence comparison with other reverse transcriptases show that the P protein consists of three domains: terminal protein (TP), polymerase or reverse trancriptase (Pol) and RNAse H (RH). The TP is attached to the 5'-end of the RNA pregenome and primes DNA synthesis. The RH activity degrades the RNA template in the hybrid molecule of RNA pregenome and minus strand DNA (Barterschlager and Schaller, 1988; Radziwill *et al.*, 1990). All available data suggest that the P protein binds to the encapsidation signal ε on the RNA pregenome which consists of some 70 nucleotides close to the 5'-end of the genomic RNA (Pollack and Ganem, 1993a; Knaus and Nassal, 1993). Rieger and Nassal (1995) demonstrated that the signal can form a characteristic bipartite stem-loop structure which may represent the actual P protein binding site. A more detailed discussion of the interaction between the P protein and the genomic RNA is presented in section 1.5.

1.4.5 X protein

The X protein (HBxAg) comprising 154 amino acids and with a molecular weight of 16.5 kDa, is a product of the X ORF. HBxAg is a multifunctional protein that displays a variety of independent activities, many of which could be important for virus infection or the development of carcinoma.

Chronic HBV infection has been demonstrated as a risk factor in the development of human hepatocellular carcinoma, HCC (Buendia, 1992; Slagle *et al.*, 1992). The mechanism of HBV involvement in liver carcinogenesis remains obscure, but several studies have suggested a possible role for HBxAg which transactivates gene expression of viral and cellular regulatory elements in the development of HCC (Hu *et al.*, 1990; Zhou *et al.*, 1990; Balsano *et al.*, 1991; Twu *et al.*, 1993). Indeed insertion of the HBxAg gene under its own promoter into mice,

resulted in pathological changes in the liver, which culminated in the development of HCC (Kim *et al.*, 1991).

Several studies have clearly demonstrated that HBxAg can function as a transactivator of a wide variety of viral and cellular transcriptional control elements. These include Simian virus 40 (SV40), Rous Sarcoma virus, human immunodeficiency virus (HIV), MHC class I and class II, β inteferon , c-*myc*, interleukin 8, intercellular adhesion molecule 1, RPB5 subunit of RNA polymerase II, TATA-binding protein, a putative DNA repair protein, and UV-damaged DNA binding protein (UV-DBB) (Twu and Schloemer, 1987; Seto *et al.*, 1988; Siddiqui *et al.*, 1987; Twu and Robinson, 1989; Hu *et al.*, 1990; Zhou *et al.*, 1990; Balsano *et al.*, 1991; Mahe *et al.*, 1991; Hu *et al.*, 1992; Cheong *et al.*, 1995., Lee *et al.*, 1995; Qadri *et al.*, 1995). HBxAg has also been shown to transactivate HBV promoters; PreS1, PreS2, X and enhancer I or II (reviewed in Rossner, 1992). On the other hand, HBxAg does not interact directly with DNA (Siddiqui *et al.*, 1987; Avantaggiati *et al.*, 1993). This fact has led to the speculation that HBxAg acts indirectly, possibly by modifying the activities of cellular factors that modulate transcription, rather than direct binding of a specific DNA sequence by X protein (Rossner, 1992; Slagle *et al.*, 1992). In this respect, several reports have shown that HBxAg stimulates the Ras-Raf-Map kinase cascade, which lead to enhanced cell proliferation and the activation of transcription factors AP-1 and NF-κB (Cross *et al.*, 1993; Benn and Schneider, 1994; Natoli *et al.*, 1994). Recently Doria *et al.*, (1995) showed that cytoplasmic HBxAg is a dual specificity activator of transcription, which activates Ras-Raf-Map kinase signal transduction pathways in the cytoplasm and stimulates HBV enhancer I in the nucleus.

The broad gene-regulating function of HBxAg suggests that it may not only regulate the expression of HBV transcripts but may also modify the biochemical environment by transactivating cellular genes in infected cells to facilitate viral replication. The mechanism of transactivation by HBxAg is currently unclear, however, the understanding of this mechanism may provide information of value in the development of compounds to inhibit the replication of HBV.

1.5 Replication of HBV

Due to the narrow host range of HBV and the lack of a routine cell culture system to propagate the virus, study of the replication of HBV has been heavily dependent on the closely related animal hepatitis virus systems: DHBV and WHV, and most of the available data have been obtained by genetic approaches using transfection of cells with HBV DNA. A brief overview of HBV replication is shown in Fig. 8 (Nassal and Schaller, 1993).

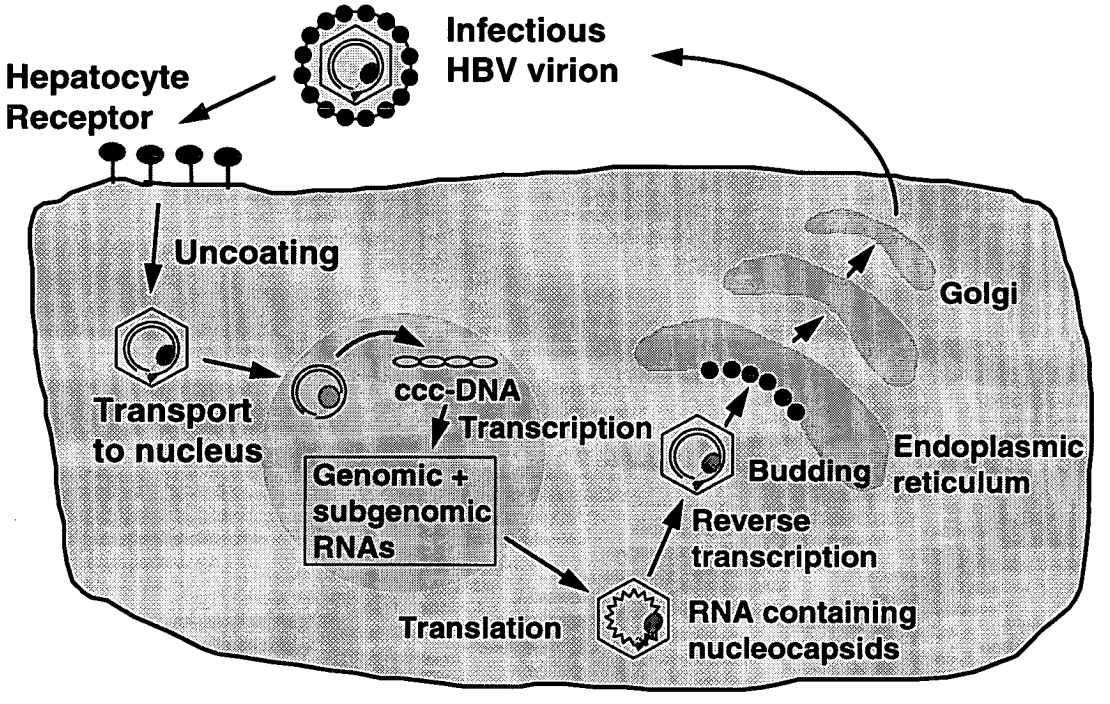


Fig. 8: The Replication of HBV

Steps involved are: (i) attachment and penetration of the host cell; (ii) uncoating and transport of the viral genome into the nucleus; (iii) conversion into ccc-DNA as the template for transcription; (iv) RNA synthesis and transport to the cytoplasm; (v) translation; (vi) assembly of RNA-containing nucleocapsids; (vii) reverse transcription of RNA and (viii) export from the cell as enveloped virions.

The receptor for binding of the virus to hepatocytes has not been identified. However, peptide sequences in the PreS1 region have been proposed by Neurath *et al.* (1986) and Pontisso *et al.* (1989) as binding sites of the HBsAg to the cellular receptor. Moreover, antibodies against these sequences neutralise the infectivity of

HBV in chimpanzees (Neurath *et al.*, 1989). From studies of the binding of radiolabelled DHBV particles in primary duck hepatocytes, Schaller (1993) predicted that there are about 5 x 10^3 binding sites per cell with an apparent K_d of 10^{-11} M. Hence, isolation of the cellular receptor(s) should eventually lead to the construction of infectable cell lines and test animals which could help to address the question of virus-host interaction. As a result, peptides or chemical compounds homologous to either the attachment sites on the HBsAg or on the receptor could then be designed for blocking virus attachment.

Following uptake of the virion, viral nucleocapsid is transported to the cell nucleus in which the genome is converted to a covalently closed circular (ccc) form that serves as the template for the transcription of genomic and subgenomic RNAs by host RNA polymerase. The genomic and subgenomic RNAs are then transported to the cytoplasm and translated into their respective proteins. When sufficient amounts of core protein, P protein and RNA pregenome have synthesised in the cytoplasm, these components coassemble into core particles (Fig. 9). The P protein binds to the ε-segment at the 5'-end of the RNA pregenome (Kaus and Nassal, 1993; Pollack and Ganem, 1993b; Fallows and Goff, 1995; Rieger and Nassal, 1995). Then, HBcAg dimers are attracted to the RNA via their non-specific nucleic acid binding domains and assemble into particles. A second specific recognition site on the P protein interacts with an additional site between DR2 and DR1* at the 3'-end of the pregenome, thus maintaining the circular structure of the RNA for replication. Reverse transcription of the RNA then begins from the 3' DR1* towards the 5'-end of the RNA, generating negative strand DNA by copying the RNA. Concomitantly, the RNAse H activity degrades the RNA template to the very 5'-end, sparing only 15-18 nucleotides, which serve as a primer for the synthesis of the plus strand DNA by copying the first DNA strand to produce partially duplex DNA intermediates (Lien *et al.*, 1986; Loeb *et al.*, 1991; Nasssl and Schaller, 1993).

The assembled capsid is either exported as enveloped virion or returned to the nucleus to amplify the intranuclear copies of cccDNA. L-HBsAg is particularly important for virion formation (Bruss and Thomssen, 1994) and hence is probably involved in specific interactions with the capsid that result in budding into the endoplasmic reticulum (ER). The recognition sites for envelopment of the

nucleocapsid have yet to be characterised, but would serve as an excellent target for antiviral attack.

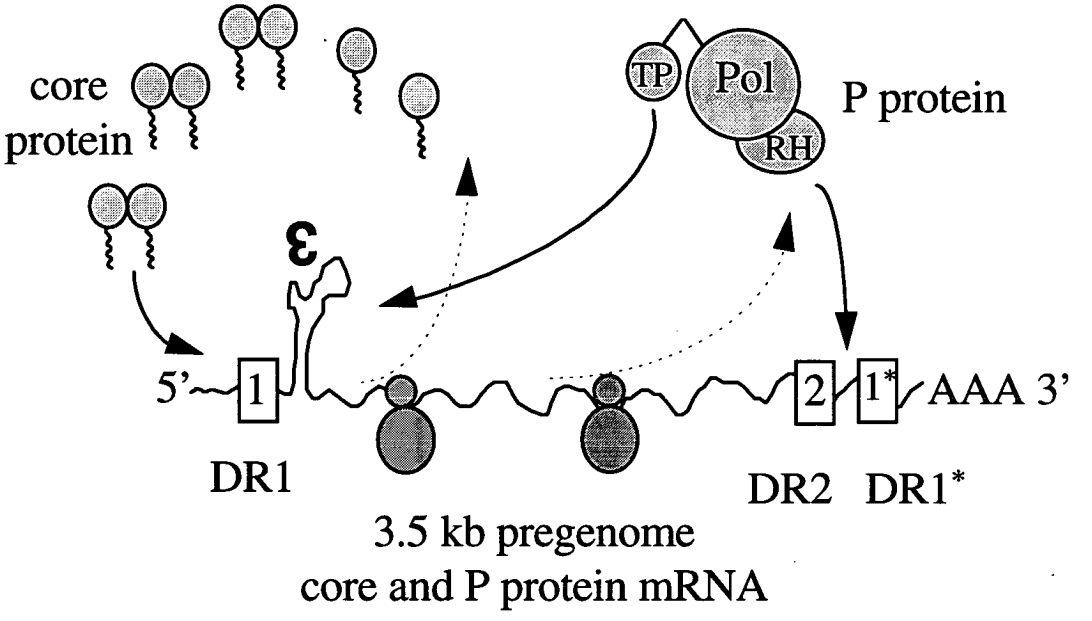


Fig.9: Nucleocapsid assembly

RNA pregenome (3.5 kb) serves as a template for the translation of core and P proteins. P interacts with the encapsidation signal, ε, on its own RNA and triggers the addition of core protein dimers via their non-specific nucleic acid binding domains and complete the capsid shell. The interaction of P protein between DR2 and the second copy of DR1 (DR1) at the 3'-end of the pregenome induces circularization and prepare the RNA for replication. (modified from Nassal and Schaller, 1993).*

1.6 Mutagenesis approach to study protein interaction

Interactions of proteins with other molecules are of fundamental importance in biology. An enzyme interacts with a substrate to accelerate a chemical reaction. Interaction between myosin and actin stimulate muscle contraction. An antibody combines with an antigen to trigger immunological responses. Envelopment of HBV requires the interaction between the surface proteins and nucleocapsids.

Protein-protein interaction depends on steric and electrostatic complimentarity of the combining surfaces. When proteins interact with each other

and form stable complexes, the surfaces buried between them cover large areas. A comparative study by Janin and Chothia (1990) showed the buried surface areas between antigen and antibodies, and proteinases and their inhibitors to be roughly the same size, 1600 ± 350 $Å^2$. They also found that 8 to 13 hydrogen bonds contribute to their binding energies, some of them formed by one ionisable group with another either neutral or oppositely charged group.

X-ray crystallography, which can be used to determine the three dimensional structure of a protein at atomic resolution, is by far the most powerful method to study directly the relationship between the structure and function of a protein, and in principle, the intermolecular interactions between two proteins. However, the crystallisation of protein-protein complexes is considerably more difficult than the crystallisation of the interacting molecule alone. Furthermore, some proteins, especially membrane proteins, are very difficult to crystallise. Probably for this reason, there is no three dimensional information at atomic resolution on the structure of HBsAg and HBcAg. Other methods for structure analysis include group specific modification of amino acids and NMR spectroscopy. Although these method have yielded much information about protein structures and interactions, they can only be used with proteins which are available in large quantity and high purity (Creighton, 1984). With the great advances in recombinant DNA technology, it is now possible to target those amino acids that potentially play functional or structural roles and then delete or replace them with alternate residues to test the role of steric constraints, hydrophobic and electrostatic forces, charge, and the placement of hydrogen bonds, salt bridges, disulphide bonds, water, or metals (Cleland, 1995). The mutant proteins can then be purified and analysed both biochemically and biophysically to determine the effect of the amino acid modifications.

1.6.1 Site-directed mutagenesis

The classical approach of inducing mutations in genes included chemical mutagenesis, ultra violet light, and ionising radiation. New phenotypic mutants were then isolated and studied. However, since the entire organism had been subjected to mutagenesis, multiple mutations often arose amongst which mutations occuring in

the gene of interest could be very low. With the advent of recombinant DNA technology, it is now possible to specifically change any given base in a cloned DNA sequence rather than crudely mutagenising thousands of organisms and then screening millions of offspring to isolate a desired mutant (Rossi and Zoller, 1987). This powerful technique is known as site-directed mutagenisis, and the method as developed by Smith and colleagues was successfully applied to introduce an amber mutation in bacteriophage ØX174 (Hutchison *et al.*, 1978). Currently, there are a variety of methods available to construct site-specific mutations *in vitro*. The most commonly used method is oligonucleotide-directed mutagenesis. Generally, this can be divided into two categories based on the form of the DNA template: single-stranded template vectors derived from filamentous phages such as M13, fd and f1; and double-stranded expression vectors (Rossi and Zoller, 1987).

In principle, these two methods are similar; a short mutagenic oligonucleotide is hybridised to its complementary sequence of the wild type DNA. The oligonucleotide serves as a primer for *in vitro* enzymatic DNA synthesis of the rest of the DNA template by DNA polymerase, forming a heteroduplex molecule containing a mismatch introduced at a specific nucleotide via the oligonucleotide primer. The newly synthesised DNA strand is covalently closed by DNA ligase and used to transform *E. coli*. The resulting products are two homoduplexes whose sequences are either wild type or mutant. The two molecules can then be distinguished readily by colony hybridisation with the labelled oligonucleotide as a probe or by DNA sequencing.

Theoretically, mutants arising from oligonucleotide-mediated mutagenesis should comprise 50% of the transformants. In practice, this efficiency is very seldom realised due to exonucleolytic degradation of the mismatched site and repair following transformation of the host (Rossi and Zoller, 1987). Kunkel (1985) exploited the properties of host cells which strongly select against the unaltered genotype to improve the efficiency of mutagenesis. In this method (Fig. 10), bacteriophages are grown on an *E. coli* host that is deficient in the enzymes dUTPase (*dut*) and uracil-N-glycosylase (*ung*). The intracellular pool of dUTP is thereby greatly increased due to the inability of the cells to convert dUTP to dUMP, as a result of which some of the dUTP molecules are incorporated into the DNA in

place of dTTP. The *ung* mutation makes the host unable to excise the incorporated dUTP. M13 bacteriophages grown in a *dut⁻ung⁻* F' strain of *E. coli* usually contain 20-30 uracil residues per genome. These uracil-containing templates are then used for oligonucleotide-mediated mutagenesis. When an *ung*$^+$ strain of bacteria is transfected with the heteroduplexs, the template strands are destroyed by uracil-N-glycosylase, but the *in vitro*-synthesised mutant strands are not, which results in a substantial enrichment of mutant bacteriophages that harbour the mutation.

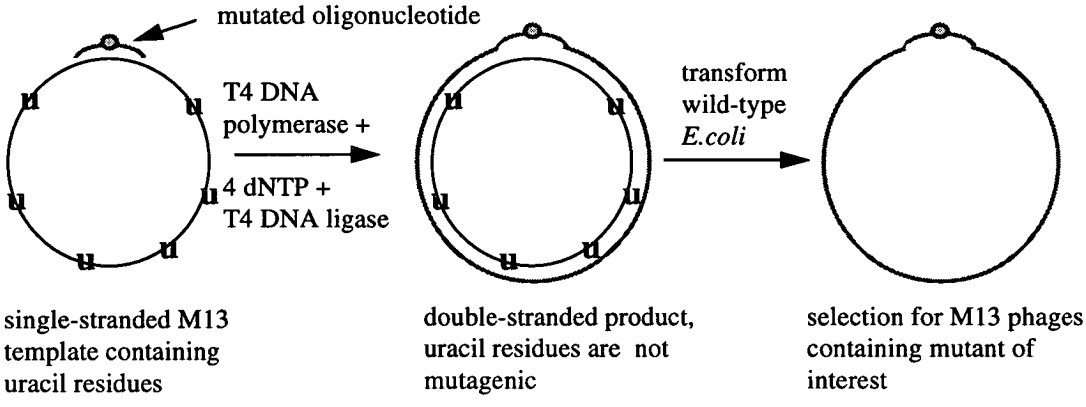


Fig. 10: Oligonucleotide-mediated Mutagenesis by Kunkel Method

ssDNA containing a small number of uracil (U) residues in place of thymine is prepared from phage propagated in a dut⁻ung⁻ strain. A synthetic oligonucleotide containing the mutation of interest is annealed to the template and treated with T4 DNA polymerase and T4 DNA ligase to produce a double-stranded circular molecule. Introduction of this heteroduplex molecule into a wild-type (dut$^+$ung$^+$) strain allows the efficient recovery of mutant DNA. (modified form Sambrook et al., 1989).

Alteration of one amino acid for another might sometimes affect the folding or stability of the entire protein. Therefore it is difficult to tell from a binding assay whether the effect of a mutation is really due to the substitution of a directly functional residue or to a non-specific disruption of protein conformation. There are several methods available to probe the structural integrity of mutated proteins. Biophysical techniques such as X-ray crystallography, NMR spectroscopy, circular dichroism spectroscopy and fluorescence spectroscopy can be used to analyse the structure of mutant proteins. Thermodynamic analysis can be used to determine the

free energy of thermal or chemical denaturation of mutant proteins as a measure of structural integrity (Sondek and Shortle, 1990). The conformation of mutant proteins can also be probed by analysing their ability to react with monoclonal antibodies that are specific for conformation epitopes (Bass *et al.*, 1991). To date, there is no way to predict accurately the perturbations in protein structure caused by such alteration. In the case of HBsAg, monoclonal or polyclonal antibodies that are specific for the native PreS regions could be used to alleviate this problem.

1.6.2 Deletion and insertion mutagenesis

These techniques are useful for mapping the boundaries of the functional domain of a protein and determining the overall domain structure of multifunctional proteins (Gibbs, 1995). Generally, deletion mutagenesis involves cleaving the target area of a DNA molecule with a specific restriction enzyme and then digesting random segments of DNA with an exonuclease such as *BAL 31* or exonuclease III for varying lengths of time. The termini generated by the exonucleases must be either staggered or blunt, and therefore, the termini are always repaired with either DNA polymerase I to fill in 3'-recessed termini, or by nuclease S1 treatment to cleave off single-stranded termini. The covalently closed forms of a nested set of deletion mutants are then generated by treatment with DNA ligases. However, theoretically, two out of three of the mutants will alter the translational reading frame and a certain percentage of these will create premature stop codons within the coding fragment (Rossi and Zoller, 1987). To generate scattered insertions, small fragments of synthetic DNA are ligated into the cleavage site. Prasad and Goff (1989) used linker insertion mutagenesis to map the p66 subunit of HIV reverse transcriptase, which is processed into two smaller subunits p51 and p15. The polymerase function was located to the N-terminal portion including the region that encodes the p51 subunit, and the RNase H activity was localised to the C-terminal portion including the region that codes for p15 subunit. Although these techniques may allow rapid identification of functionally important regions, they suffer from the disadvantage that protein structure is likely to be disrupted by insertions and deletions.

1.7 Aims of the project

A comprehensive understanding of the interaction between HBsAg and HBcAg is practically important for rational design of drugs to inhibit the envelopment of nucleocapsids. This process begins with the identification of contact residues at or near the interface. At the time when this project was initiated, there was no information available on the structure of HBsAg and very little was known about the structure of HBcAg. Furthermore, no experiments had been reported that directly determine or indicate the regions of HBsAg which interact with HBcAg. The only hint suggesting the region in HBsAg which might directly interact with HBcAg was obtained from affinity selection from fusion phage display libraries (Dyson and Murray, 1995), but the amino acid sequence, LLGRMK, implicated in the binding with HBcAg appears to be a discontinuous sequence in which the residues constituting this sequence are brought together from different positions in the L-HBsAg polypeptide. Chapter 4 therefore focuses on deletion and site-specific mutations of L-HBsAg in order to define the exact contact residues involved in the interaction.

Mutagenic analyses of the interaction require that the genes for the interacting molecules are cloned and the proteins expressed in substantial quantities for biochemical and biophysical assays. As HBsAg could not be expressed efficiently in *E. coli*, the wild-type and mutant proteins were synthesised *in vitro* with rabbit reticulocyte lysates. It is fortunate that HBcAg can be expressed in the prokaryotic cell where it assembles into two sizes of particle with morphology indistinguishable from nucleocapsids isolated from HBV-infected individuals. Chapter 3 addresses the questions concerning the biological and physical properties of the HBcAg monomer and the particles it forms: (i) yield improvement of HBcAg synthesised in *E. coli*; (ii) separation of the two kinds of particle and; (iii) heterogeneity of HBcAg monomer.

The goal of the work presented in the final chapter was to evaluate whether the deletions and specific changes in L-HBsAg have affected its binding with HBcAg. Initially, the evaluation was performed with immunoprecipitation of the interacting molecules with polyclonal antibodies. However, this method does not

allow precise discrimination of small changes in the binding. Therefore, steps were taken to set up an equilibrium binding assay in solution which allows determination of relative dissociation constants.

CHAPTER 2

MATERIALS AND METHODS

2A Materials

2A.1 Chemicals

The chemicals used in this study were mostly of analytical grade and were purchased from BDH Chemicals Ltd, Boehringer Mannheim, Biorad, GIBCO BRL, Scotlab, Sigma Chemicals and United States Biochemicals (USB). Solvents were supplied by Fisher Chemicals and radioactive reagents were supplied by Amersham International plc.

2A.2 Liquid and solid media

Media	Recipes, per litre
Luria broth (LB)	Bacto-tryptone (10 g), Bacto-yeast extract (5 g), Nacl (5 g), adjusted to pH 7.2
LB agar	As Luria broth plus agar (15 g)
Minimal agar	Agar (15 g), Glucose (20% w/v;10 ml),vitamin B1 (625 µg), Spitzizen Salts[a] (200 ml)
Top layer agar	Trypticase (10 g), Nacl (5 g), Bactor-agar (10 g)
2x Yeast tryptone (YT)	Bacto-tryptone (16 g), Bacto-yeast extract (10 g), Nacl (5 g)

[a]Spitzizen salts : 0.2% $(NH_4)_2SO_4$, 0.14% K_2HPO_4, 0.6% KH_2PO_4, 0.1% Trisodium citrate 0.2% $MgSO_4$. Media was prepared by the department media service.

2A.3 Antibiotic Solutions

antibiotic	stock concentration[b]	working concentration (µg/ml)
Ampicillin	100 mg/ml in H_2O	50
Chloramphenicol	30 mg/ml in ethanol	30
Kanamycin	20 mg/ml in H_2O	20
Tetracyclin	20 mg/ml in ethanol	20

[b]Stock solutions of antibiotics dissolved in H_2O were filtered through a 0.22-micron filter and all the solutions were kept at -20 °C.

2A.4 Standard solutions and buffers

Solution/Buffer	Components
10x TBE	Tris-HCI (0.9 M), boric acid (0.9 M), EDTA (20 mM; pH 8.0)
PBS	K_2HPO_4 (20 mM), KH_2PO_4 (5 mM), NaCl (150 mM)
TE pH 7.6	Tris-HC1 (10 mM; pH 7.6), EDTA (1 mM; pH 8.0)
TE pH 8.0	Tris-HCl (10 mM; pH 8.0), EDTA (1 mM; pH 8.0)
STE	Tris-HCl (10 mM; pH 8.0), EDTA (1 mM; pH 8.0), NaCl (0.15 M)
TBS	Tris-HCl (50 mM; pH 7.5), NaCl (150 mM)
NET-gel	Tris-HCl (50 mM; pH 7.6), NaCl (150 mM), NP-40 (0.1% v/v), EDTA (1 mM), Gelatin (0.25% w/v), NaN_3 (0.02% w/v)

2A.5 *Escherichia coli* strains

strain	Genotype	Reference
CJ236	*dut*1, *ung*1, *thi*-1, *rel* A1/ pCJ105 (Cam^r F')	Kunkel *et al.*, (1987)
K91Kan	HfrC (λ^-), *thi*	Smith and Scott, (1993)
SG935	F^- *lac* (*am*), *trp* (*am*), *pho*(*am*), *sup*C(*ts*), *rsp*L, *ma*1 (*am*), *htp*R (*am*), *tsx*: : TN10 lon Δ100	S.Goff (unpublished)
TG1	*sup*E, *hsd*Δ5, *thi*, Δ (*lac-pro*AB), F'[*tra* D36, *pro*AB^+, *lac*I^q, *lac*ZΔM15]	Gibson, (1984)
W3110IQ	F' *lac* with *lac*I^q	Bachmann (1972)
XL1-Blue	*sup*E44, *hsd*R17, *rec*A1, *end*A1, *gyr*A46, *thi*, *rel*A1, *lac*, F' [*pro*AB^+, *lac*I^q, *lac*ZΔM15, Tn10(*Tet*^{r})]	Bullock *et al.*, (1987)

2A.6 Plasmids and cloning vectors

Plasmid	Description and selectable Marker	Reference
pAGA9 (4.68 kb)	a derivative of pGP1-2 (Tabor and Richardson, 1985) carrying the T4 AGA tRNA gene (Mazzara *et al.*, 1981). It consists of gene 1 of phage T7 under the control of the inducible λ P_L promoter, and the gene for the heat sensitive λ repressor, cI857; Amp^R Kan^R	Zahn and Landy, (1996)
pCITE-2c (3.735 kb)	contains a Cap-Independent Translation Enhancer (CITE) sequence of the encephalomyocarditis virus (EMC) RNA 5' non-coding region, which functions as an internal entry point for initiation of translation by eukaryotic ribosomes. It also carries the T7 promoter for high efficiency *in vitro* transcription with T7 RNA polymerase; Amp^R	Novagen catalogue (597 Science Dr. Madison, WI 53711, USA)
pMDHBs3 (5.131 kb)	contains a 1.35 kb-insert of the L-HBsAg gene derived from pHBV130 (Gough and Murray, 1982). The insert is located in the polycloning site of the pCITE-2c vector, with the *Eco* R1 and *Sal* I sites at the 5'- and 3'-ends, respectively; Amp^R	Dyson and Murray, (1995)
pMDHBs4 (4.639 kb)	same as pMDHBs3, but it has a 858 bp-insert of the S-HBsAg gene derived from pHBV130; Amp^R	Dyson and Murray, (1995)
pR1-11E (4.823 kb)	contains the coding sequence for amino acids 1-8 of β-galactosidase followed by 3 linker amino acids (Glu,Phe,His) followed by amino acids 3-148 of HBcAg, under control of the *tac* promoter.	Stewart, (1993)
pTacpcore (5.378kb)	contains the coding sequence for amino acids 1-8 of β-galactosidase followed by 3 linker amino acids (Glu,Phe,His) followed by amino acids 3-183 of HBcAg, under control of the *tac* promoter.	Stewart, (1993)
M13mp18 (7.25 kb)	a bacteriophage M13 vector used for subcloning the L-HBsAg gene for preparation of ssDNA. IPTG and X-gal were used to identify recombinant phages; these form colourless plagues, while nonrecombinant phages form blue plagues.	Messing et al., (1977)
pM13-MDHBs3 (8.565 kb)	carries a 1.35 kb-insert of theL-HBsAg gene with M13mp18 vector as backbone.	

2B METHODS

2B.1 Preparation of DNA

i • Phenol extraction and ethanol precipitation of DNA

Contaminating proteins in DNA solutions were removed by extraction with phenol (equilibrated at pH 8.0), phenol:chloroform (1.1), and chloroform. An equal volume of these organic solvents was added to the aqueous DNA solution and the mixture was vortexed vigorously for 2 to 3 min. Organic and aqueous phases were separated by centrifugation at 15,000 x*g*, for 5 min at 4 °C. The upper aqueous phase containing DNA was recovered and the DNA was precipitated by the addition of 3 M sodium acetate (pH 5.2; 0.1 vol), and absolute ethanol (-20 °C; 2 vol). This was left at -20 °C for 20 min or overnight. Precipitated DNA was recovered by centrifugation at 15,000 x*g*, for 15 min or longer at 4 °C. The pellet was then washed with 70% (v/v) ethanol (-20 °C), dried under vacuum, and dissolved in TE (pH 7.6 or 8.0) or distilled water.

ii • Miniprep of plasmid DNA

The alkaline lysis method (Birnboim and Doly, 1979; Birnboim, 1983) was frequently used to isolate small quantities of plasmids and RF DNAs from bacterial cells for the identification of recombinant clones by restriction analysis.

Well isolated bacterial colonies were grown in LB medium (5 ml) supplemented with appropriate antibiotic(s) at 37 °C overnight with shaking (225-250 rpm). Cells (1.5 ml) were pelleted by centrifugation for 5 min and resuspended in ice-cold solution I (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8.0; 100 µl). The cells were then lysed by the addition of solution II [0.2 M NaOH, 1% (v/w) SDS; 200 µl], mixed well and incubated on ice for 5 min. Ice-cold solution III [5 M potassium acetate, 11.5% (v/v) glacial acetate acid; 150 µl] was then added, mixed and incubated on ice for 5 min. The white precipitate was sedimented by centrifugation at 15,000 x*g* for 5 min at 4 °C, and the supernatant transferred to a fresh tube. The solution was then extracted once with phenol:chloroform (1:1) and

the DNA was precipitated with ethanol (2 vol, 4 °C). The DNA was washed with 70% (v/v) ethanol, dried and resuspended in TE (25 µl, pH 7.6).

iii • Large-scale preparation and purification of plasmid and replicative forms (RF) of bacteriophage DNA

This method, which requires two steps, was used to prepare large quantities (~0.5 mg) of highly purified mutant plasmids derived from pMDHBs3 for *in vitro* transcription. It was also used to prepare RF DNA of M13mp18. Firstly, crude lysate was prepared by the alkaline lysis method (Sambrook *et al.*, 1989). Secondly, PEG precipitation (Lis and Schlrif, 1975; Lis, 1980) was applied to purify plasmid DNA.

Cells from overnight culture (250 ml; LB medium) were harvested by centrifugation at 4,000 x*g* for 15 min at 4 °C and resuspended in ice-cold STE (100 ml). The suspension was then spun under the same conditions and pellets were resuspended in solution I (18 ml; ice-cold). Cells were lysed by the addition of lysozyme [10 mg/ml in 10 mM Tris-HCl pH 8.0; 2 ml] and solution II (40 ml). The mixture was incubated at room temperature for 10 min. After the addition of solution III (20 ml; ice-cold), the mixture was incubated on ice for 10 min, and the white precipitate was separated by centrifugation at 5,000 x*g* for 15 min at 4 °C. The supernatant was then filtered through 4 layers of cheese cloth, isopropanol (0.6 vol) was added and the mixture was stored at room temperature for 10 min. This was then spun at 5,000 x*g* for 15 min at room temperature, and the pellet was washed with 70% (v/v) ethanol, dried and dissolved in TE (pH 8.0; 3 ml). An equal volume of LiCi (5 M; ice-cold) was added, mixed well and the solution centrifuged at 11,000 x*g*, for 10 min at 4 °C. The supernatant was then transferred to a fresh tube containing an equal volume of isopropanol and centrifuged at 11,000 x*g* for 10 min at room temperature. The pellet was rinsed with 70% (v/v) ethanol, dried and dissolved in TE (pH 8.0; 0.5 ml) containing DNAase-free pancreatic RNAase (20 µg/ml). After incubation at room temperature for 30 min, PEG 6000 [13% (w/v) in 1.6 M NaCl; 0.5 ml] was added and the solution was centrifuged at 13,000 x*g* for 10 min at 4 °C. The DNA pellet was dissolved in TE (pH 8.0; 0.4 ml), extracted once with phenol, phenol:chloroform (1:1) and chloroform. DNA was then precipitated

with ammonium acetate (10 M; 100 µl) and ethanol (2 vol), washed, dried and dissolved in TE (pH 8.0; 0.5 ml).

iv • Small-scale preparation of M13mp18 ssDNA

ssDNA produced by this method (Sambrook *et al.*, 1989) was used as template for dideoxy chain-termination DNA sequencing (Method 2B.6)

Well isolated colourless plaques of transformants were grown in LB medium (3 ml) containing log phase TG-1 cells (75 µl) as the host. After incubating at 37 °C for 5 h, the cultures (1.5 ml) were spun at 15,000 x*g* for 10 min at 4 °C. The supernatant was transferred to a fresh tube and PEG solution [20% (w/v) PEG 6000 in 2.5 M NaCl; 200 µl] was added, mixed and left at room temperature for 15 min. This was then centrifuged at 15,000 x*g*, 10 min at 4 ° C and the supernatant was removed by aspiration. The pellet was suspended in TE (pH 8.0; 100 µl) and extracted once with phenol, followed by chloroform. DNA was precipitated with ethanol, washed, dried, and dissolved in TE (pH 8.0; 25 µl).

v • Preparation of uracil-containing single-stranded DNA

Plasmid pM13-MDHBs3, a M13mp18 derivative containing the L-HBsAg gene, was used to transform *E. coli* strain TG-1. The transformant, in the form of single-stranded bacteriophage M13-MDHBs3, was grown in *E. coli* strain CJ236 (*dut⁻ ung⁻* F') to produce uracil-containing ssDNA which was used as template for oligonucleotide-mediated mutagenesis (Method 2B.5.*i*). This method was originally described by Kunkel (1985; Kunkel *et al.*, 1987).

A single plaque was transferred into 2x YT medium (1 ml) and incubated at 60 °C for 5 min to kill TG-1 cells. This was vortexed vigorously for 30 s to release bacteriophages trapped in the top agar, and then centrifuged at 12,000 x*g* for 2 min at 4 °C. Supernatant (50 µl) was transferred to a flask (500 ml) containing 2x YT medium (50 ml) supplemented with uridine (0.25 µg/ml). A mid-log-phase culture of *E. coli* strain CJ236 (5 ml) was added and the culture was incubated at 37 °C for 6 h with shaking at 250 rpm. The cells were pelleted by centrifugation at 5,000 x*g* for 30 min at 4 °C and the clear supernatant was transferred to a centrifuge bottle. An aliquot of the phage suspension(200 µl) was titered on CJ236 and TG-1 cells. It

was found that the titer on strain CJ236 was 5 to 6 orders of magnitude greater than on TG-1. To the remaining bacteriophage suspension, PEG solution (15% w/v PEG 6000, in 2.5 M NaCl; 0.25 vol) was added, mixed and then left on ice for 1 h. The precipitated bacteriophage particles were recovered by centrifugation at 5000 x*g* for 20 min at 4 °C and the pellet was resuspended in TE (pH 7.6, 4 ml). The suspension was vortexed vigorously for 30 s, and then left on ice for 1 h. The suspension was vortexed one more for 30 s and spun at 5,000 x*g* for 20 min at 4 °C. The supernatant was transferred to a polypropylene tube, extracted twice with phenol, once with phenol: chloroform (1:1), ethanol precipitated, washed, and dissolved in TE (pH 7.6; 200 µl).

vi • Quantitation of DNA

The concentration of DNA in solution was measured by its absorbance at 260 nm and 280 nm with a spectrophotometer (UNICAM model SP 500 Series 2). An OD_{260nm} of 1.0 represents a concentration of 50 µg/ml for dsDNA and 40 µg/ml for ssDNA. The ratio OD_{260nm}/OD_{280nm} provides an estimation of the purity of the nucleic acid. A value of 1.8 indicates a reasonably pure preparation of DNA (Sambrook *et al.*, 1989).

2B.2 Resolution and recovery of DNA fragments

Agarose gel electrophoresis was routinely used to separate, characterise, and purify DNA fragments.

i • Agarose gel electrophoresis and photography

Mini agarose gels [5 cm x 7.5 cm; 7 cm x 10 cm; 1% (w/v)] were used to check the integrity of DNA and identify the existence of the appropriate inserts in recombinant clones after digestion with restriction enzymes. Gels containing ethidium bromide (0.5 µg/ml) were prepared and electrophoresed in 1 x TBE. Prior to loading, DNA samples were mixed with loading buffer [0.42% (w/v) bromophenol blue, 0.42% (w/v) xylene cyanol FF, 50% (v/v) glycerol; 0.1 vol] and electrophoresis was carried out horizontally at constant voltage (60-100 V) for an

indicated period of time (usually, 30 min to 2 h). DNA size markers (GIBCO BRL, 1 kb Ladder; 5 µg) were loaded adjacent to the DNA samples in order to estimate the size of DNA fragments. DNA fragments were visualised on a transilluminator (UUP, INC), and photographs were taken with a UVP camera linked to a video copy processor (MITSUBISHI).

ii • Isolation and purification of DNA fragments from agarose gels

Prior to the ligation step in cloning processes (Method 2B.3), DNA inserts and vectors that had been digested with restriction enzymes, were separated, recovered and purified with the following protocol (provided by Epicentre Technologies).

DNA fragments were separated with low melting point agarose gels (1 to 1.5% w/v; GIBCO BRL), and visualised by UV illumination (UVP, model UVGL-58). Zones containing target fragments were cut out from the gels, weighed, and 50x GELase buffer [2 M bis-(2-hydroxyethyl)-imino-tris(hydroxymethyl)-methane (pH 6.0), 50 mM EDTA, 2 M NaCl; 2 µl per 100 mg of gel slice] was added. These were incubated at 70 °C until the gels were completely molten (around 20 min). The temperature of the molten gels was then equilibrated to 45 °C and GELase™ (Epicentre Technologies) was added (1.5 U per 80 mg of 1.5% w/v gel; 1 U per 80 mg of 1% w/v gel). After incubation at 45 °C for 1 h, DNA fragments were precipitated with ammonium acetate (5 M, pH 5.2; 1 vol) and ethanol (2 vol). The DNAs were dissolved in distilled water and were then ready for ligation.

2B.3 Cloning into plasmids and bacteriophage M13 vectors

The following procedures were used to insert DNA fragments encoding the L-HBsAg gene and mutants derived from it into plasmid pCITE-2c and M13mpl8 vector.

i • Digestion of insert and vector with restriction endonucleases

The vectors were cleaved at their multiple restriction sites with two different restriction enzymes to generate linear molecules with incompatible termini, sticky,

or blunt ends. The enzymes were also used to digest the fragment to be inserted so as to produce termini compatible with those of the vectors. In general, 1-5 U of restriction endonuclease were used per μg of DNA and incubated for 2-5 h with buffer and temperature conditions recommended by the supplier of the enzymes. Digestion reactions were terminated by heating at 65 °C for 10 min. The vectors were phenol extracted and ethanol precipitated prior to dephosphorylation.

Restriction enzymes used in cloning:

Enzyme	1x incubation buffer	Supplier
Eco RI	Tris-HCl (50 mM, pH 7.5), $MgCl_2$ (10 mM), NaCl (100 mM), DTT (1 mM).	Boehringer Mannheim
Sal I	Tris-HCl (50 mM, pH 7.5), $MgCl_2$ (10 mM), NaCl (100 mM), DTT (1 mM).	Boehringer Mannheim
Sma I	Tris acetate (33 mM), Mg acetate (10 mM), K acetate (66 mM), DTT (0.5 mM); pH 7.9.	Boehringer Mannheim

ii • Dephosphorylation

After linearisation of the vectors, calf intestinal alkaline phosphatase [CIP; Boehringer] was used to remove their 5'-phosphate in order to prevent recircularisation during ligation. The reaction (50 μl) consisted of DNA (20 μg in 44 μl H_20), 10 x CIP buffer [0.5 M Tris-HCl (pH 8.5), 1 mM EDTA] and CIP (24 U/μl; 1 μl). For 5'-recessed DNA, the mixture was incubated at 37 °C for 1 h, 50 °C for 30 min and 75 °C for 20 min, whereas for 5'-blunt-ended DNA the reaction was incubated at 37 °C for 1 h. CIP was inactivated with phenol:chloroform (1:1) extraction and DNA was ethanol precipitated and dissolved in TE (pH 7.6; 20 μl).

iii • Ligation

Before ligation, inserts and the dephosphorelated vectors were run on LMP agarose gels and the desired fragments isolated and purified (Method 2B.2.*ii*). Ligation was performed with a three-fold molar excess of inserts (100-200 ng) in a reaction mixture [50 mM Tris-HCl (pH 7.6), 10 mM DTT, 500 μg/ml BSA; 10μl] containing T4 DNA ligase (Boehringer; 1 U) and incubated at 16 °C overnight. For ligation of blunt-end DNA fragments, 5 U of ligase was used and incubated at

22 °C overnight. Ligation products (1 and 5 µl) were then used to transform competent *E. coli* (Method 2B.4).

2B.4 Introduction of DNA into *E. coli* cells

This method for transformation was described by Cohen *et al.* (1972) and Dargert and Ehrlich (1974).

i • Preparation of competent cells

To 50 ml of LB (supplemented with the appropriate antibiotic), an overnight culture of an *E. coli* strain (1/10 dilution) was added. The culture was grown at 37 °C with shaking (200-250 rpm) to an OD_{600nm} between 0.3 and 0.4. The cells (25 ml) were chilled on ice for 30 min, pelleted by centrifugation at 2,000 x*g* for 15 min at 4 °C, and resuspended in ice-cold $CaCl_2$ (0.1 M; 10 ml). After 30 min incubation, the cells were repelleted and resuspended in ice-cold $CaCl_2$ (0.1 M; 2 ml) and kept for 30 min on ice.

ii • Transformation of competent cells with plasmid DNA

Plasmid DNA (5-50 ng) or a ligation reaction mixture (Method 2B.3.*iii*) was mixed with competent cells (200 µl) and left on ice for 30 min. Cells were then heat-shocked for 2 min at 42 °C , LB (1 ml) was then added, and incubated at 37 °C for 1 h. Cells (50 µl and 200 µl) were then plated on appropriate selective LB agar and incubated overnight at 37 °C.

iii • Transfection of competent cells with bacteriophage M13 DNA

RF M13mp18 DNA (5 ng) or recombinant constructs derived from it (5 µl of a ligation reaction; Method 2B.3.*iii*) was mixed with competent cells (200 µl), left on ice for 30 min, and heat-shocked at 42 °C for 2 min. Log phase TG-1 cells (OD_{600nm} 0.6 to 0.8; 200 µl) were then added to the mixture, mixed well, and melted top layer agar (3 ml) containing X-gal (20 mg/ml in dimethylformamide; 40 µl) and IPTG (200 mg/ml; 4µl) was poured into the tube, gently mixed and poured on LB agar plates. The plates were incubated overnight at 37 °C, and white and blue plaques were counted the next day.

2B.5 Mutagenesis

i • Oligonucleotide-mediated mutagenesis (Kunkel, 1985)

- ***Oligonucleotides***

Negative sense oligonucleotides designed for site-directed mutagenesis of the L-HBsAg coding sequence were synthesised by the Oswel DNA service, University of Edinburgh (later moved to University of Southampton).

oligo[a]	position[b]	sequence[c]	mutation induced
K38A	1049-1069	CAGGTGTCC***GC***GTTGGGATTG	Lys-38 →Ala
K46A	1074-1093	GCTCCTACC***GC***GTTGGCGTC	Lys-46 → Ala
64Random	1129-1147	CTCCACCCC***NN***AAGGCCTC	Leu-64→ Random
L63/G65-Random	1124-1152	GAGGGCTCCAC***NN***CAAA***NN***GCCTC CGTGC	Leu-64 and Gly-65 → random
R88A	1200-1218	CTGACTGG***GC***ATTGGTAGA	Arg-88 → Ala
R92A	1213-1233	GGGTAGGCTGC***GC***TCCTGACT	Arg-92 → Ala
R102A	1240-1262	ATGAGTGGTT***GC***CAGAGGTGGAG	Arg-102 → Ala
R124A	1307-1327	CCTCTCACT***GC***GGGATCTTGC	Arg-124 → Ala
R126A	1314-1334	ATACAGGCCT***GC***CACTCTGGG	Arg-126 → Ala
R156A	1402-1422	CCCCAATC***GC***CGAGAAGATTG	Arg-156 → Ala
R187A	1498-1517	TGTGAGGATT***GC***TGTCAACA	Arg-187 → Ala
P192Amber	1513-1530	GACTCTGC***TA***TATTGTGA	Pro-192 → Amber
C211Amber	1571-1588	TGGCCAA***GT***CACACGGTA	Cys-211 → Amber
F243Amber	1667-1684	AAGATGAT***CT***AACGCCGC	Phe-243 → Amber
Q264Ochre	1729-1747	AACATACCT***TA***ATAGTCCA	Gln-264 → Ochre
K323Ochre	1905-1922	TAGGAATT***A***TCCGAAAGC	Lys-323 → Ochre
V347Ochre	1976-1998	GGGAAAGCCCT***TA***GAACCACTGA	Val-347 → Ochre
L372 Amber	2050-2069	GGGACTC***T***AGATGCTGTACA	Leu-372 → Amber

[a] oligonucleotide descriptions differ from those given by Oswel and have been adopted to describe the mutation induced. [b] in the numbering system of Pasek *et al.*, 1979. [c] nucleotides printed in italics (and underlined) identify the base changes used to induce mutations.

• ***Phosphorylation of oligonucleotides***

Prior to the mutagenesis reactions the oligonucleotides (150 pmole) were phosphorylated with T4 polynucleotide kinase (D-501, ICMB, University of Edinburgh; 1 U), in Kinase buffer [50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM; spermidine HCl, 0.1 mM EDTA (pH 8.0); 20 µl] in the presence of 0.5 mM ATP. The reactions were carried out at room temperature (24-26 °C) for 1 h and terminated by heating at 70 °C for 10 min.

• ***Primer extension***

The phosphorylated oligonucleotides (15 pmole) were annealed with M13-MDHBs3 (Method 2B.1.*v;* 0.5 pmole) in annealing buffer [20 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT; 10 µl] at 80 °C for 5 min and transferred to a beaker containing water at 80 °C and the mixtures were allowed to cool slowly to room temperature. The oligonucleotide-template mixtures were then added to mutagenesis reaction buffer [1 mM dNTPs, 1 mM ATP, 20 mM Tris-Cl (pH 7.5), 10 mM $MgCl_2$, 10 mM DTT; 10 µl] containing T4 DNA polymerase (Boehringer; 5 U) and T4 DNA ligase (Boehringer; 3 U). The reaction mixtures were incubated at 0 °C for 5 min, at room temperature (24-26 °C) for 5 min, and at 37 °C for 2 h and were used to transfect competent *E.coli* strain TG-1 cells (Method 2B.4 *iii*). ssDNAs were prepared from the transformants (Method 2B.1.*iv*) and DNA sequencing (Method 2B.6) was performed to determine the mutations that had been introduced. The L-HBsAg coding segment of the desired mutant was excised from the M13mp18 vector by digestion with *Eco* R I and *Sal* I and ligated into the pCITE-2c vector (Method 2B.3).

ii • Deletion mutagenesis with BAL 31 exonuclease

• ***Calibration of BAL 31 exonuclease activity***

A series of seven two-fold dilutions of (2 µl) *BAL 31* exonuclease (GIBCO BRL; 1.4 U) was carried out with *BAL 31* buffer [600 µM NaCl, 12 mM $CaCl_2$, 12 mM $MgCl_2$, 20 mM Tris-HCl (pH 8.0), 0.2 mM EDTA (pH 8.0); 2 µl]. Aliquots (1 µl) of these enzyme dilutions were then used to digest *Eco* RI linearised

pMDHBs3 DNA (0.55 μg) in *BAL 31* reaction mixture (10 μl) at 37 °C for 30 min. The reactions were terminated by adding of EGTA (200 mM, pH 8.0; 1μl) and heated at 65 °C for 5 min. The molecular weight of the digested DNAs was analysed by electrophoresis in 0.8% (w/v) agarose gel. The dilution of *BAL 31* which had totally digested the DNA was used in a large-scale digestion, as described below.

• ***Generation of nested deletions of linearised plasmid***

pMDHBs3 (20 μg) DNA that had been linearised with *Eco* RI was digested with *BAL 31* (3.13 U) in *BAL 31* buffer (361 μl) at 37 °C. At 1 min intervals, samples of the reaction mixture (45 μl) were removed and heated at 65 °C for 5 min in the presence of EGTA (200 mM, pH 8.0). After that, the DNAs were precipitated with ethanol and dissolved in H_2O (20 μl). The DNAs were then repaired with T4 DNA polymerase (Boehringer; 1 U) in polymerase buffer [20 mM Tris-HCl (pH 7.6), 1 mM $MgCl_2$, 0.1 mM DTT; 30 μl] in the presence of dNTPs (0.05 mM) as substrates. These reactions were incubated at room temperature for 15 min and Kleonow fragment of *E.coli* DNA polymerase I (Northumbria Biologicals; 5 U) was added and incubated at the same temperature for 15 min. The repaired DNAs were precipitated with ethanol and each of the DNA pellets was dissolved in H_2O (17 μl).

• ***Liberation of 5'-end truncated genes from the vector DNA***

The above DNAs were digested with *Sal* I (10 U) and electrophoresed in 1.5% (w/v) agarose gel to separate the 5'-end truncated L-HBsAg DNA fragments (target fragments) from their vector component. Target fragments were recovered (Method 2B.2 *ii*) and subcloned into M13mp18 vector that had been digested with *Sal* I and *Sma* I (Method 2B.3).

2B.6 DNA sequence determination

Sequencing reactions were performed by the dideoxynucleotide chain termination method of Sanger and Coulson (1975; Sanger *et al.*, 1977).

i • Single stranded DNA sequencing

The primer for the polymerase reaction was annealed to its template as follows. A microcentrifuge tube containing ssDNA (100 ng/µl; 8 µl), primer (1 pmole/µl; 1 µl) and 10 x TM (100 mM Tris-HCl pH 8.5, 50 mM $MgCl_2$; 1 µl) was heated at 80 °C for 5 min and then allowed to cool at room temperature (10-15 min), centrifuged briefly, and aliquoted (2 µl) to four wells (labelled A, C ,G and T) of a microtiter plate which contained nucleotide mixes (A, C, G and T respectively, as shown in the following table; 2 µl). Klenow reaction mix (9 mM Tris-HCl pH 8.0, 9 mM DTT, 0.2 U *E. coli* DNA polymerase I Klenow fragment, 0.56 µCi α-[^{35}S]-dATP; 2 µl) was then added to each well, mixed gently by tapping and incubated at room temperature for 15 min. This was followed by the addition of chase mix (0.25 mM dNTP; 2 µl) and incubation continued for a further 20 min. The reaction was stopped by adding gel loading dye [0.5% (w/v) xylene cyanol, 0.5% (w/v) bromophenol blue, 10 mM EDTA in formamide; 2 µl]. Prior to electrophoresis, the reaction mixtures were heated at 90 °C for 5 min and quickly chilled on ice.

Nucleotide mixes (volumes in µl)

	A	C	G	T
0.5 mM dCTP	500	25	500	500
0.5 mM dGTP	500	500	25	500
0.5 mM dTTP	500	500	500	25
10 mM ddATP	1	-	-	-
10 mM ddCTP	-	8	-	-
10 mM ddGTP	-	-	16	-
10 mM ddTTP	-	-	-	50
TE (pH 8.0)	500	1 000	1 000	1 000

• ***Denaturing polyacrylamide gel electrophoresis and autoradiography***

The sequencing reaction products were separated on a 420 mm x 180 mm x 0.3 mm, 6% polyacrylamide gel [60 ml of: 6 M urea, 5.7% (w/v) acrylamide, 0.3% bisacrylamide, 1x TBE, polymerised by adding 300 µl of 10% (w/v) ammonium persulphate and 75 µl TEMED]. Usually, the gel was pre-electrophoresed for 30 min in 1 x TBE, before the reaction mixtures (3 µl) were loaded in sharkstooth comb wells. Electrophoresis was performed at a constant power of 60 W for the indicated period of time. After electrophoresis, the gel was soaked with fixer [10% (v/v) acetic acid; 10% (v/v) ethanol] for 20 min, transferred to blotting paper, dried under vacuum at 80 °C for 40 min and autoradiographed against an x-ray film (DUPONT). After the appropriate length of time, the film was developed using an X-O GRAPH COMPACT x 2 automatic X-ray film processor.

ii • Automated DNA sequencing

The automation of manual sequencing steps was developed by research at the California Institute of Technology (Smith *et al.,* 1986). In this method, the steps were not fully automated. Sequencing and gel electrophoresis were still performed manually. However, the detection and analysis of DNA sequences by a machine have greatly increased the number of bases that can be determined per reaction.

Thermal cycle sequencing was carried out using the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer). To a microcentrifuge tube (0.5 ml), the following reagents were added: template (0.1 µg ssDNA or 0.4 µg dsDNA), terminator ready reaction mix [ddNTP-dye terminator, dNTP, Tris-HCl (pH 9.0), $MgCl_2$, thermal stable pyrophosphatase, AmpliTaq DNA polymerase (FS); 8 µl] and primer (3.2 pmole). The reaction mixture (20 µl) was overlaid with mineral oil (40 µl) and subjected to the following temperature profile: 96 °C for 30 s, 50 °C for 15 s and 60 °C for 4 min, 25 cycles. The sequencing products were precipitated by adding 3 M sodium acetate (pH 4.6; 0.1 vol) and 95% (v/v) ethanol (2.5 vol), washed with 70% (v/v) ethanol (250 µl) and vacuum dried. Electrophoresis of sequencing reaction products was performed by Nicola Preston and Sandra Bruce (ICMB) on an automated sequencer (ABI Prism 377), and the data

obtained were analysed using Sequence Navigator DNA and protein sequence comparison software (Applied Biosystems).

2.B.7 *In vitro* transcription and translation

Plasmid pMDHBs3 and mutated derivatives (30 μg) were first linearised with *Sal* I, then purified with phenol, phenol:chloroform (1:1) and chloroform extraction and precipitated with 3 M sodium acetate (pH 5.2; 0.1 vol) and ethanol (100%; 2 vol) overnight. The linearised DNAs (1μg) were used as templates in reaction mixtures [80 mM HEPES-KOH (pH 7.5), 24 mM $MgCl_2$, 2 mM spermidine, 40 mM DTT; 20μl] containing 7.5 mM rNTPs and T7 RNA polymerase (Promega, Madison, WI; 600 U). Transcriptions were performed at 37 °C for 2 h and the synthesised RNAs (2 μl aliquots) were kept at -70 °C. Prior to translation reactions, the synthetic RNAs were diluted 10-fold and samples of these dilutions (1 μl) were added to reaction mixtures (9 μl) containing rabbit reticulocyte lysate (Promega, Madison, WI; 5 μl), amino acid mixture minus methionine (20 μM), [^{35} S]-methionine (1 Ci/mol, Amersham; 0.35 μl), $Mg(OAC)_2$ (0.6 mM), KCI (120 mM) and DTT (2 mM). Reactions were carried out at 30 °C for 2 h in the presence or absence of HBcAg (0.6 μg).

i • Determination of percent incorporation of [^{35}S]-Met and the yield of translation products

Translation mixtures were diluted (1:50; 100 μl) with 1 M NaOH containing 2% (v/v) H_2O_2 and incubated at 37 °C for 10 min. These were then diluted to 1 ml with 25 % (v/v) ice-cold TCA containing 2% (w/v) casamino acids and left on ice for 30 min. These mixtures (250 μl) were placed onto glass microfibre filters (Whatman, 2.1 cm diameter), which were situated in scintillation vials. The filters were washed three times with 5% (v/v) TCA (3 ml) and once with acetone (3 ml), each at interval of 10 min. The washing solutions were removed by aspiration and the discs dried at 60 °C for 20 min. Scintillation liquid (Ecoscint A, National Diagnostics, Atlanta; 5 ml) was added and radioactivity was determined with a liquid scintillation spectrometer (Beckman, LS 7 000). To determine total counts

present in the reaction mixtures, the TCA-treated reaction mixtures (5 µl) were spotted directly onto filters, dried and counted.

The percentage of the [^{35}S]-methionine incorporated into translation products was calculated with the following formula:

Percent incorporation = (cpm of washed filter/cpm of unwashed filter x 50) x 100

From these values, the amount of [^{35}S]-methionine in each reaction mixture, and the number of methionine residues in each polypeptide, the total amount of peptide synthesised in each translation system was determined.

2B.8 Protein purification and analysis

i • Analysis of the yield of HBcAg synthesised in E. coli supplemented with T4 AGA tRNA

E. coli strain W3110IQ containing plasmids pTacpcore (encoding full-length HBcAg) and pAGA9 (encoding T4 AGA tRNA), or pTacpcore alone, or without the two plasmids were grown in LB medium (80 ml, supplemented with appropriate antibiotics) at 38 °C with shaking (~200 rpm) to an OD_{600nm} approximately 0.8, and IPTG (0.5 mM) was added. The culture was then split into two flasks, one continued shaking at 38 °C and the other was shaking at 42 °C (waterbath shaker) for 15 min and then continued shaking at 38 °C. The 42 °C-heat treatment for the second flask was a precautionary step to ensure all the heat sensitive λ repressors, cI857, driven by the P_{lac} promoter of the pAGA9 plasmid were completely denatured. After 24 h incubation, cells were pelleted by brief centrifugation, resuspended in loading buffer, boiled and analysed on SDS-PAGE (2B.8.*vi*).

ii • Purification of full-length and truncated HBcAg particles

The method for purification of HBcAg particles was initially described by Murray *et al.*, (1984) with some modification by Dyson and Murray (1995).

Several *E. coli* strains harbouring plasmids pTacpcore or Rl-11E, respectively, encoding full-length or truncated HBcAg were grown at 37 °C with shaking (250 rpm) in LB (500 ml) supplemented with ampicillin (100 µg/ml). The expression of HBcAg was induced with IPTG (0.5 mM) when the cultures reached OD_{600nm} around 0.8 to 1.0. After 16-20 h of shaking, the cells were harvested by centrifugation at 4,000 x*g* for 15 min at 4 °C and the pellet suspended in lysis buffer [50 mM Tris-HCl (pH 8.0), 0.1% (v/v) Triton; 12 ml]. $MgCl_2$ (4 mM), lysozyme (50 mg/ml; 48 µl) and DNase 1 (7.5 mg/ml; 30 µl) were then added to the suspension which was shaken gently (on orbital shaker) at room temperature for 2 h. The cell extract was recovered by centrifugation at 11,000 x*g*, for 20 min at 4 °C and precipitated by ammonium sulphate (35% saturation). The precipitate was centrifuged with the same conditions, the pellet resuspended in TBS (150 µl), dialysed against two changes of TBS (2 l) at 4 °C overnight. The dialysed solution was applied to 8-40% sucrose gradients (in TBS; 12 ml), and centrifuged at 100,000 x*g* (TH641 rotor, Sorvall) for 5 h at 4 °C. Fractions (0.5 ml) were collected by punching a hole at the bottom of the tube and those containing HBcAg were pooled and concentrated with a 300 kDa cut-off MICROSEP filter (Flowgen).

iii • Separation of large and small particles of HBcAg

After the first sucrose gradient centrifugation, the preceding and succeeding fractions of the peak containing HBcAg particles were pooled separately and dialysed against TBS (2 l) at 4 °C overnight. The dialysed HBcAg particles were then concentrated to 100 µl with a 300 kDa cut-off MICROSEP filter and layered on top of 8- 40% sucrose gradients, and centrifuged at 100,000 x*g* for 5 h at 4 °C. The above steps were repeated until two separate peaks containing large and small particles were obtained. The peaks were then pooled, dialysed and concentrated separately. The monodispersity of these particles was analysed by light scattering (Method 2B.8.*v*).

iv • Determination of HBcAg concentration with the Bradford assay.

Protein samples were diluted to 40 µl with TBS in a disposable tube. Then, 1 ml of 1x Bradford dye [0.01% (w/v) Coomassie blue G-250, 4.75% (v/v) ethanol, 8.5% (v/v) phosphoric acid] was added and vortexed immediately. The mixture was left at room temperature for 10 min and the OD_{595nm} measured with a spectrophotometre (UNICAM, model SP 500 series 2). Protein concentration was determined from a standard curve obtained from measurements with BSA (2.5, 5, 7.5, 10, 15 and 20 µg) in the same way. All samples were analysed in triplicate.

v • Light scattering analysis of HBcAg

The dispersity of HBcAg particles was analysed with a dynamic light scattering instrument (DynaPro-801™, PROTEIN SOLUTIONS Ltd). The operation of the instrument was assisted by Professor Malcolm Walkinshaw (Department of Biochemistry, University of Edinburgh) according to the user manual. Purified, and separated large and small particles of HBcAg (0.1 mg/ml; 150µl) were injected into a sample cell through Whatman Anatop™ plus syringe filters (0.1 µm porosity). The sample were illuminated by a miniature solid state laser of approximately 25 mW power and 780 nm wavelength. Scattered light at 90° angle was detected and analysed with the autocorelation function to deduce the translational diffusion coefficient (D_T) of the molecules in the sample cell. The hydrodynamic radius (R_h) of the molecules was calculated from D_T through the Stokes-Einstein equation:

$$R_h = k_b T / 6 \pi \eta D_T$$

where k_b is Boltzman's constant, T is the absolute temperature in degrees Kelvin, η is the solvent viscosity. The instrument assumes the viscosity of water, $\eta = 1.019 \times 10^{-3}$ Nsm^{-2}.

vi • SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were fractionated by SDS-PAGE (Laemmli, 1970) with a discontinuous buffer system. A Mini-Protein II gel apparatus (BioRad) was used and gels were prepared with the following recipes:

		Separating gel 15% (w/v)	Stacking gel 5% (w/v)
polyacrylamide mix [a]	:	15 ml	1.66 ml
4 x lower buffer [b]	:	7.5 ml	-
4 x upper buffer [c]	:	-	2.50 ml
H_2O	:	7.5 ml	5.84 ml
10% w/v ammonium persulphate	:	187.5 µl	66.6 µl
TEMED	:	30.0 µl	26.0 µl

[a] 30% (w/v) acrylamide, 0.8% (w/v) bisacrylamide

[b] containing 1.5 M Tris-HCl (pH 8.6), 0.4% (w/v) SDS

[c] containing 0.5 M Tris-HCl (pH 6.8), 0.4% (w/v) SDS

Protein samples (1 to 2 µl) were prepared by mixing with 1x loading buffer [62.5 mM Tris-HCl (pH 6.8), 2% (w/v) SDS, 10% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 50 mM DTT; 5 vol] and boiling for 5 min. Electrophoresis was performed in Tris-Glycine Buffer [3% (w/v) Tris, 14.4% (w/v) glycine, 0.1% (w/v) SDS; pH 8.4] at a constant current of 30-35 mA for 45-60 min. Gels were stained with coomassie brilliant blue R-250 [0.5% (w/v); in 25% (v/v) isopropanol and 10% (v/v) acetic acid] for 5-10 min at room temperature with gentle agitation and destained in solution containing methanol (10% v/v) and acetic acid (10% v/v) overnight.

2B.9 Isolation of protein for microsequence analysis

The following procedure was used to isolate protein fragments with lower molecular weight than full-length HBcAg monomer (~22kDa, 192 residues) which assemble into HBcAg particles. First of all, denatured HBcAg particles were fractionated by SDS-PAGE at neutral pH. The proteins were then transferred to a polyvinylidene difluoride (PVDF) membrane and stained with coomassie blue. Bands of interests were excised and analysed with an automated protein sequencer.

i • SDS-PAGE

The method used was that described by Moos *et al.,* (1988) in which a polyacrylamide gel was preelectrolised with glutathione to eliminate the by-products of acrylamide polymerisation, followed by electrophoresis at a lower pH to reduce the blockage of amino termini of the proteins.

Denaturing minigel (Method 2B.8.*vi*) was prepared as follows:

		Separating gel 15% (w/v)	Stacking gel 5.4% (w/v)
polyacrylamide mix[a]	:	20 ml	0.666 ml
4 x gel buffer[b]	:	10 ml	-
H_2O	:	9.7 ml	3.033 ml
10% (w/v) ammonium perssulphate	:	0.14 ml	0.025 ml
TEMED	:	0.026 ml	0.025 ml

[a] 30% (w/v) acrylamide, 0.8% (w/v) bisacrylamide
[b] containing 10.31% (w/v) bis-Tris, 0.66% (v/v) HCl; pH 6.61

The gel was preelectrophoresed in 1x gel buffer containing glutathione (1 mM final concentration) in the upper buffer reservoir, at 10 mA for 45 min. The gel was left overnight at room temperature, buffer was then decanted and wells blotted with filter paper. Prior to electrophoresis, protein samples were boiled

for 5 min in the presence of 1x loading buffer (Method 2B.8.*vi*) and loaded into the wells. The gel was then electrophoresed in 1x upper reservoir buffer [0.05% (w/v) mercaptoacetic acid, 1% (w/v) N-tris-(hydroxymethyl)-methyl-2-aminoethanesulfonic acid (TES), 2.37% (w/v) bis-Tris, 0.1% (w/v) SDS; pH 7.25] and 1 x lower reservoir buffer [1.31% (w/v) bis-Tris, 0.114% (v/v) HCl ; pH 5.9] at 10 mA until the bromophenol blue dye reached the bottom of the gel (~ 80 min).

ii • Electroblotting of protein onto PVDF membrane

The gel was removed from the electrophoresis cell and soaked in transfer buffer [10 mM 3-(Cyclohexylamino)-1-propanesulfonic acid (CAPS) pH 11 in 10% (v/v) methanol] for 5 min. Meanwhile, a PVDF (ProBlott™, Applied Biosystems) membrane was cut to approximately the same size as the gel, soaked with methanol for a few seconds, and transferred to a dish containing transfer buffer. A gel 'sandwich' was assembled in a blotting cassette submerged in transfer buffer with the following order: a Scotch-Brite pad, one sheet of blotting paper cut to the gel size, the gel, the PVDF membrane, another sheet of blotting paper and another Scotch-Brite pad. The blotting cassette was placed in the electroblotting apparatus (BioRad, Trans-Blot cell) containing transfer buffer by facing the membrane towards the anode side of the tank. Electroblotting was conducted at constant voltage of 50 V, at room temperature for 30 min. After electrophoresis, the membrane was removed from the sandwich and rinsed with H_2O.

iii • Coomassie blue staining and band excision

The membrane was soaked in absolute methanol for a few seconds and transferred to staining solution [0.1%(w/v) coomassie blue R-250 in 40% (v/v) methanol] and gently agitated for 5 min. The membrane was then destained with 50% (v/v) methanol by agitating until bands became clearly visible (~30 min). Protein bands of interest were excised with razor blades, placed in microcentrifuge tubes, dried at room temperature and stored at -20 °C. The samples were sequenced on an automated protein sequencer (Applied Biosystem) and analysed by Dr. Andy

Cronshaw (WELMET, Edinburgh protein characterisation facility, University of Edinburgh).

2B.10 Analysis of interactions between HBsAg and HBcAg

i • Immunoprecipitation with anti-HBcAg rabbit polyclonal serum

Immunoprecipitations were carried out at 4 °C by mixing 40x diluted translation mixtures [NET-gel buffer containing 2 mM DTT; 200 µl] with anti-HBcAg rabbit polyclonal serum (1 µg) and rolled gently for 1 h. Protein A sepharose (Pharmachia; 10 µl of a 50% slurry) was then added and rolling continued for a further 1 h. The immunoprecipitated complexes were recovered by centrifugation and washed four times with NET-gel buffer and once with 10 mM Tris-HCl (pH 7.6) containing 0.1% (v/v) NP-40. For every washing, the complexes were left rolling gently for 20 min. The last traces of the final wash from the sepharose pellets, and the walls and lids of the centrifuge tubes were removed by aspiration. These pellets were then heated together with SDS gel loading buffer [50 mM Tris-HCl (pH 6.8), 100 mM DTT, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 10% (v/v) glycerol; 10 µl] at 100 °C for 5 min. The protein A-sepharose was removed by centrifugation at 15,000 x*g* for 2 min at room temperature. The supernatants were analysed by 15% (w/v) SDS-PAGE. (Method 2B.8.*vi*). Gels were dried and [^{35}S]-labelled products were visualised by autoradiography and analysed on a phosphoimager (Molecular Dynamics).

ii • Development of an HBsAg-HBcAg binding assay in solution

The experimental establishment of a binding assay to determine the affinity constant of HBsAg-HBcAg complexes is described below. A detailed discussion of the purposes and theoretical background behind each step is presented in Chapter 5.

• *Time course experiment of the L-HBsAg binding to HBcAg coated wells*

Two sets of strip plates (Costar) with U-shaped polystyrene wells were coated with HBcAg (10 µg/ml in TBS buffer; 120 µl) overnight at room temperature with gentle agitation and washed three times with PBS. Then, one set of the plates

was blocked with BSA (10 mg/ml in TBS) at 4 °C (cold room) for 2 h, the other at room temperature (27 °C) with the same conditions. Afterwards, a constant concentration of [^{35}S]-L-HBsAg (100 µl of 200x dilution of *in vitro* translation reaction mixtures with NET-gel buffer containing 2 mM DTT) was pipetted into the HBcAg coated wells and incubated for different time periods (1, 2, 3, 4, 5, 6, 20, 24, 28.5 and 48 h) at 4 °C and at room temperature. After each incubation period, the relevant strip representing a point on a time course was detached from the plate and washed five times with NET-gel buffer containing 2 mM DTT. Wells were air dried and placed in scintillation vials containing scintillation liquid (Ecoscint A National Diagnostic, Atlanta; 5 ml) for quantitation of radioactivity with a scintillation counter (BECKMAN, LS 1701). A negative control was carried out with the same conditions but in the absence of [^{35}S]-L-HBsAg (an *in vitro* translation reaction that did not contain mRNA encoding for the L-HBsAg). Each time point was analysed in triplicate.

- ***Determination of the amount of [^{35}S]- L-HBsAg captured on HBcAg coated wells (input-output relationship)***

Two sets of plates were prepared with coating conditions as above and blocked with BSA at 4 °C. Various dilutions of *in vitro* translation mixtures (1/20, 1/40, 1/80, 1/160, 1/320, 1/640, 1/1280 and 1/2560) were pipetted into the coated wells of the first plate and incubated under the same conditions as above. After incubation, wells were washed five times with the same buffer as above and radioactivity counted. This experiment was repeated separately, with wells coated with 7.5, 5.0 and 2.5 µg/ml of HBcAg. Each dilution of the *in vitro* translation mixtures was assayed in triplicate.

- ***Determination of relative dissociation constants (K_d^{rel}) of HBsAg-HBcAg complexes***

HBcAg at various concentrations (0.01-80 µM, in NET-gel buffer containing 2 mM DTT) was incubated with a constant amount of HBsAg (1:50 or 1:100 dilution of *in vitro* translation mixtures with the same buffer) for 24 h at 4 °C , in a

final volume of 500 µl. An aliquot (100 µl in triplicate) of the mixtures was then transferred to a U-shaped well that had been coated with HBcAg (2.5 µg/ml; conditions as above) and blocked with BSA (10 mg/ml). After 1 h incubation at 4 °C, the wells were washed 5 x with NET-gel buffer containing 2 mM DTT, dried and radioactivity counted.

- ***Determination of K_d^{rel} values with the Scatchard plot***

The Scatchard equation derived for the above binding assay in solution is:

$$(cpm_o\text{-}cpm)/(cpm_o[C]) = 1/K_d^{rel} - [(cpm_o\text{-}cpm)/cpm_o]/K_d^{rel}$$

The mean value ± standard deviation of counts per minute (cpm) measured by the scintillation counter (which reflects the concentration of free [^{35}S]-HBsAg in solution) at various concentrations of HBcAg ([C]) was calculated and the value was subtracted from the counts per minute obtained for the well containing no HBcAg (cpm_o). The value obtained was then devided by cpm_o to generate fraction bound, Y, $[(cpm_o\text{-}cpm)/cpm_o]$. A plot of Y/[C] versus Y gives a straight line of slope equal to $-1/K_d^{rel}$. The theoretical background of the derivation of the Scatchard equation is described in Charpter 5.

- ***Determination of K_d^{rel} values with the curve fitting method***

A plot of Y versus [C] was prepared and a curve fitted to these points using equations for one binding site or two binding sites (GraFit™ version 3.0 Erithacus Software ltd).

Equation for one binding site:

$$Y = \{[C] \cdot B_{max}/(K_d^{rel} + [C])\} + background$$

where B_{max} is the maximal binding and background is the background value of Y.

Equation for two binding sites:

$$Y = \{[C] \cdot B_{max1}/(K_{d1}^{rel} + [C])\} + \{[C] \cdot B_{max2}/(K_{d2}^{rel} + [C])\} + background$$

where K_{d1}^{rel} and K_{d2}^{rel} represent the dissociation constants for the high- and low-affinity sites respectively. B_{max1} and B_{max2} refer to maximal binding for the high- and low-affinity sites respectively.

iii • Inhibition of HBsAg binding to HBcAg by synthetic peptides

[^{35}S]-HBsAg synthesised in rabbit reticulocytes was diluted 100-fold with NET-gel buffer (plus 2 mM DTT) containing various concentrations of synthetic peptide (100 nM to 1 mM). An aliquot of these mixtures (100 µl, in triplicate) was incubated in HBcAg coated wells (10 µg/ml; as in 2B.10.*ii*) for 24 h at 4 °C. Then, wells were washed five times with NET-gel buffer (plus 2 mM DTT) with 10 min intervals and placed in scintillation vials for quantitation of radioactivity. Peptides LDPAFR, LEDPASR and EDPASR were synthesised by the Department of Chemistry, University of Edinburgh. Peptide LLGRMKG was provided by Biogen Inc., Cambridge, MA, USA.

iv • Phage binding assay in solution

This method was established by Dyson *et al.*, (1995). Various concentrations of HBcAg (0.16 nM to 251 nM in TBS containing 20 mg/ml BSA) were incubated with a constant concentration of purified phage 2A-8 (1/200 dilution of 8 x 10^5 pfu/µl; in TBS containing 20 mg/ml BSA) for 18 h at 4 °C. In addition, a tube containing only phage was prepared. Then, an aliquot (100 µl) of the mixtures was transferred to a HBcAg-coated well (2.0 µg/ml; see 2B.10.*ii*) and incubated for 1 h at 4 °C. Wells were then washed six times with TBS containing BSA (20 mg/ml). Phages bound to the wells were eluted by incubation with elution buffer (0.1 M HCl, titrated to pH 2.2 by addition of solid glycine; 200 µl) for 10 min. Solutions from each well were transferred to microcentrifuge tubes containing Tris-HCl (1 M, pH 9.5; 37.5 µl) and vortexed for 20 s. The amount of phage recovered from each well was determined by titration as plaque forming units (pfu) as follows: An aliquot (1/100 dilution in TBS containing 1 g/l gelatin; 10 µl) of the mixtures was added to a suspension of K91Kan cells (300 µl) grown to OD_{600nm} 0.8. Melted top layer agar (equilibrated to 45 °C; 3 ml) was then poured into the tubes containing phages and cells, gently mixed and the mixtures poured on L-agar plates containing kanamycin (100 µg/ml). The plates were incubated overnight at 37 °C and plaques counted the next day. All assays were performed in triplicate.

- ***Determination of K_d^{rel} values with the curve fitting method***

The mean value ± standard deviation of plaque forming units (pfu) determined at various concentration of HBcAg ([C]) was divided by pfu obtained for the well in the absence of HBcAg (pfu_o). The ratios were then plotted against [C] and a curve fitted to these points using one binding site equations (GraFit™ version 3.0 Erithacus Software ltd).

CHAPTER 3

HBcAg PARTICLES: YIELD IMPROVEMENT, HETEROGENEITY AND FRACTIONATION

3.1 Introduction

HBcAg comprises 183 residues and can be divided into two domains; the first 140 amino acids are sufficient for self-assembly (Zlotnick *et al.,* 1996), and the C-terminal sequence is very rich in arginine residues which closely resembles protamines and is believed to interact with nucleic acids (Pasek *et al.,* 1979). HBcAg has been successfully expressed in *E. coli* for nearly two decades now (Pasek *et al.,* 1979) and *E. coli* is still the host of choice in our laboratory and elsewhere for routine large scale production of HBcAg particles. In the prokaryotic cell, HBcAg assembles into icosahedral particles morphologically resembling those found in infected liver (Cohen and Richmond, 1982). Two sizes of particle are produced: large and small particles containing 240 and 180 monomers, respectively (Crowther *et al.,* 1994). When the extremely basic protamine-like region (residues 149-183) is removed, the HBcAg still assembles into particles, but the packaging of RNA is abolished (Borisova *et al.,* 1989; Gallina *et al.,* 1989; Birnbaum and Nassal, 1990; Hatton *et al.,* 1992; Zheng *et al.,* 1992; Ulrich *et al.,* 1993). Furthermore, and of practical importance, the yield of truncated HBcAg lacking the protamine-like region in *E. coli* is much greater than that of full-length HBcAg (Zheng *et al.*, 1992; Stewart, 1993).

In this chapter, three questions concerning the HBcAg are addressed:

(i) : How to improve the yield of full-length HBcAg synthesised in *E. coli*?

(ii) : How to separate the large and small particles of HBcAg?

(iii): Are the purified HBcAg particles homogeneous?

3.2 Results

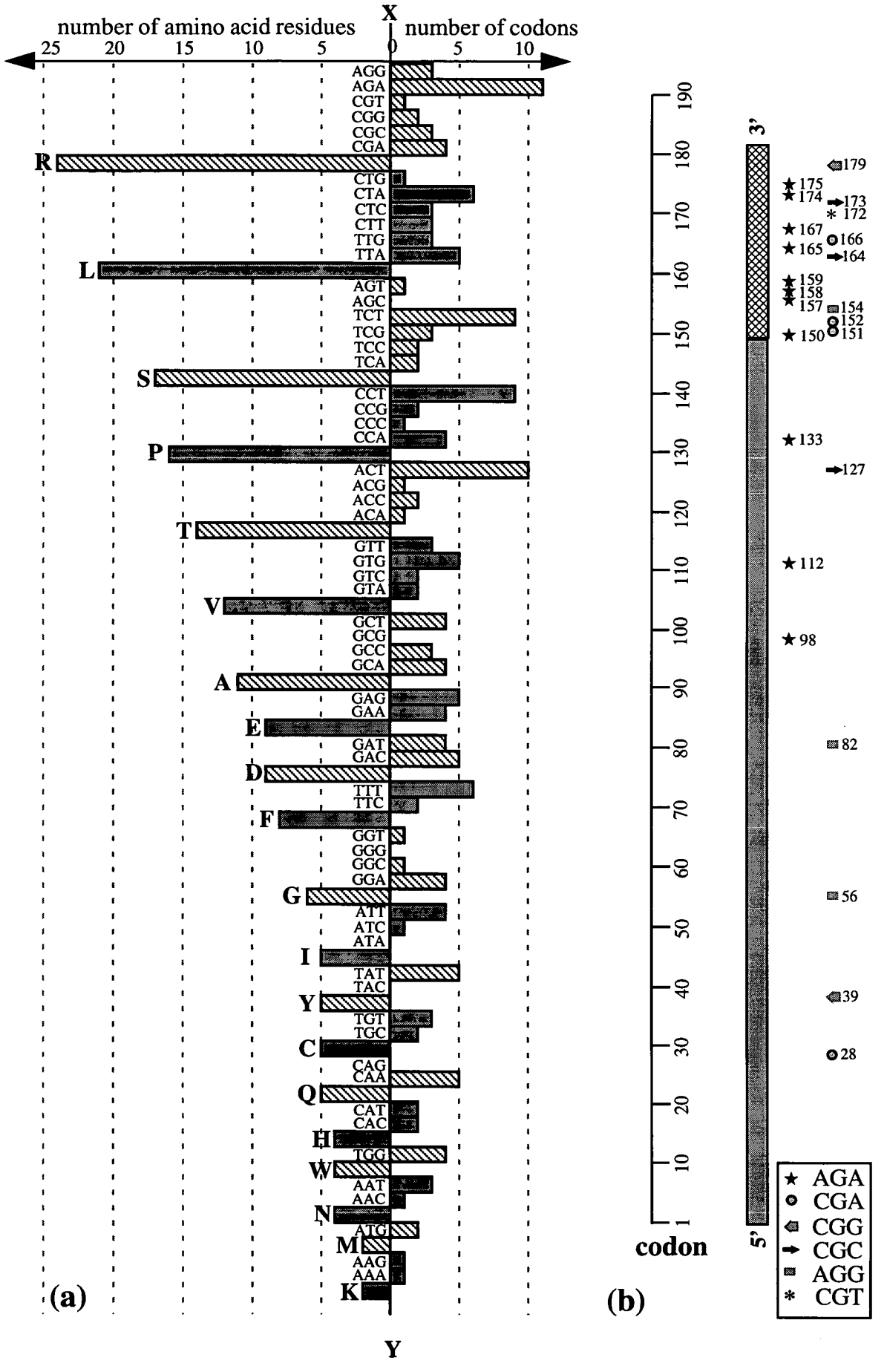
3.2.1 Codon usage in the HBcAg gene

Fig. 11a displays the codon usage of HBcAg gene (strain *adyw*), summarising the number of codons used and the total number of amino acids encoded. Although 55 of the 61 possible codon triplets are used to synthesise HBcAg, it is clear that codon usage is not random and there is a predominance of

Fig. 11:

(a) Codon usage in HBcAg gene. *The 61 possible triplets encoding the 20 different amino acids are listed on the cenral vertical line, xy. The total number of each triplet used in the coding region of the HBcAg gene is represented by bars on the right of the xy line. Bars on the left of the line show the number of amino acid residues encoded by each triplet.*

(b) Positions of arginine codons in the HBcAg gene. *The HBcAg gene is represented by the vertical bold line, (5'- to 3') from the bottom to the top, and the scale is shown on the left. The six different codons for Arg are represented by different symbols as indicated at the lower right of the diagram. Each codon's position is indicated on the right of the symbols.*


number of amino acid residues
X
number of codons
25
20
15
10
5
0
5
10
AGG
AGA
CGT
CGG
CGC
CGA
R
CTG
CTA
CTC
CTT
TTG
TTA
L
AGT
AGC
TCT
TCG
TCC
TCA
S
CCT
CCG
CCC
CCA
P
ACT
ACG
ACC
ACA
T
GTT
GTG
GTC
GTA
V
GCT
GCG
GCC
GCA
A
GAG
GAA
E
GAT
GAC
D
TTT
TTC
F
GGT
GGG
GGC
GGA
G
ATT
ATC
ATA
I
TAT
TAC
Y
TGT
TGC
C
CAG
CAA
Q
CAT
CAC
H
TGG
W
AAT
AAC
N
ATG
M
AAG
AAA
K
(a)
Y
codon
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
3'
5'
179
175
174
173
172
167
166
165
164
159
158
157
154
152
150
151
133
127
112
98
82
56
39
28
AGA
CGA
CGG
CGC
AGG
CGT
(b)

certain codons, especially AGA, the most frequently used codon, which comprises 6% of all codons in the HBcAg gene and 46% of the Arg codons. It is also interesting to note that the codons for Arg residues are the most abundant in the HBcAg gene, comprising 13% of all the codons, and 67% of these codons are located at the 3'-end of the HBcAg mRNA encoding the protamine-like region (Fig. 11b, codons 149-183). Furthermore, 50% of the Arg codons in this region are AGA. Thus, there is a strong bias towards the use of codon AGA, particularly at the 3'-end of the mRNA encoding full-length HBcAg.

3.2.2 Enhancement of HBcAg synthesis in *E. coli* by T4 AGA tRNA supplementation

To improve the yield of full-length HBcAg synthesised in *E. coli*, the plasmid pAGA9, carrying the T4 AGA tRNA gene under the control of the inducible P_L promoter, was introduced into a W3110IQ strain containing pTacpcore plasmids encoding full-length HBcAg, driven by the *tac* promoter. With cells harbouring pTacpcore plasmids, a 22 kDa protein band, corresponding to full-length HBcAg, was produced after induction with IPTG (Fig. 12a, lanes 6 and 7). Cells harbouring both pTacpcore and pAGA9 showed a significant increase in HBcAg synthesis as shown by a more intense 22 kDa band (Fig. 12a, lanes 9 and 10). In order to quantitate the yield of HBcAg produced, the dried gel was scanned (Hewlett Packard, Scanjet IIcx) and the image analysed with ImageQuant software (Molecular Dynamics). A rectangle was drawn along lanes 7 and 10, and the average pixel value along the length of the rectangle was integrated by the software that reported the results as a line graph and area under each peak was expressed as a percentage of the total protein in order to compensate for any differences in loading or staining between tracks. Fig. 12b shows the graphs obtained from lanes 7 and 10 which are superimposed on eachother. It is clear that the synthesis of HBcAg has increased as indicated by a higher and wider peak (asterisk) corresponding to the 22 kDa band in lane 10. The area under the peak that represents the amount of HBcAg synthesised relative to the total cell protein, increased from 13.7% to 20.8% (a 50% increase) when cell were supplemented with T4 AGA tRNA. In one litre cultures,

Fig. 12:

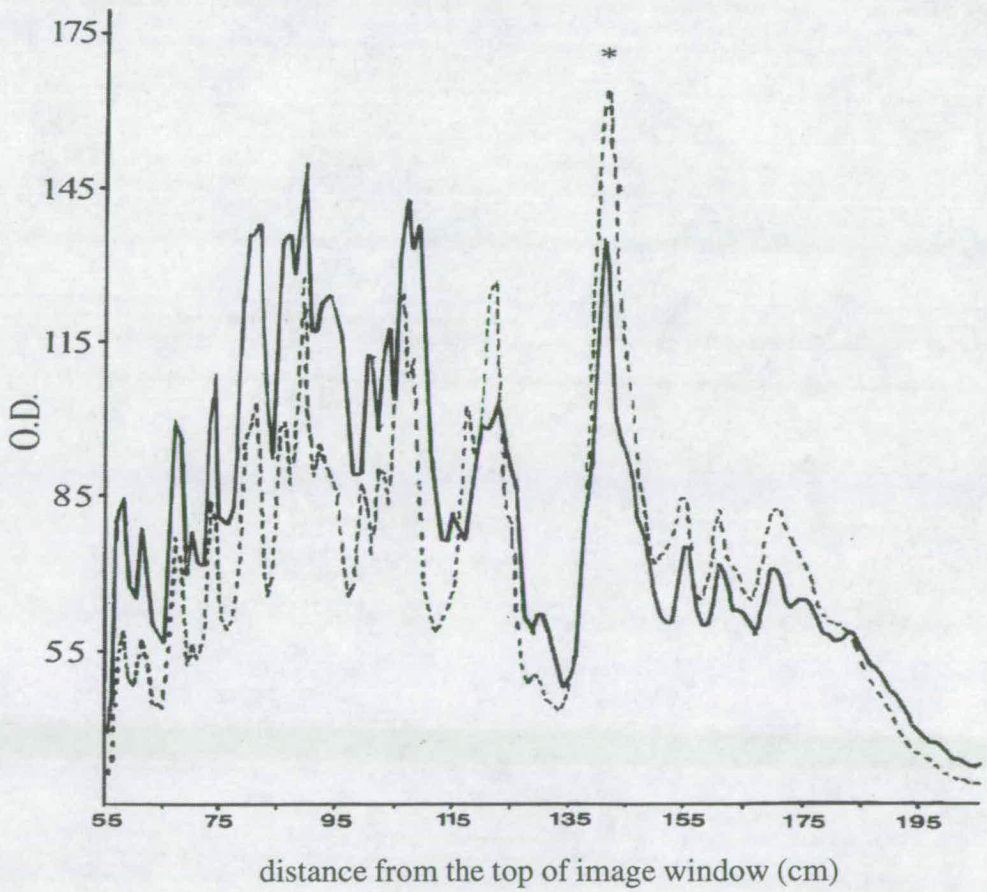
(a) Coomassie blue staining of crude cell extracts separated on SDS-Polyacrylamide gel (15%). *W3110IQ cells (lanes 2 to 4) , W3110IQ cells harbouring pTacpcore plasmids (lanes 5 to 7), and W3110IQ cells harbouring pTacpcore and pAGA9 plasmids (lanes 8 to 10) were induced with IPTG (lanes 3, 4, 6, 7, 9 and 10) or without IPTG (lanes 2, 5 and 8). After IPTG induction, cells on lanes 4, 7 and 10 were shaken at 42 °C for 15 min before incubate at 38 °C for overnight. Lanes 1 and 11 show the molecular weight marker (in kDa) and purified HBcAg, respectively.*

(b) Quantitation of HBcAg in crude cell extracts. *Graphs obtained from area integration of a rectangle drawn along lanes 7 and 10 of above gel are superimposed on another and peaks correspond to 22 kDa band are indicated by *. ImageQuant software (Molecular Dynamics) was used to integrate the area under each peak and results were reported as percent total (data not shown). Horizontal axis corresponds to the distance along the rectangle and the left end of the graph corresponds to the top of the rectangle. The units on the y axis are arbitrary.*

(a)

		W3110IQ			W3110IQ pTacpcore			W3110IQ pTacpcore pAGA9			
IPTG		-	+	+	-	+	+	-	+	+	
42 °C 15 min		-	-	+	-	-	+	-	-	+	
lanes	1	2	3	4	5	6	7	8	9	10	11

(b)



the yield of HBcAg particles approximately doubled, from about 5 mg to 10 mg.

3.2.3 Heterogeneity in purified HBcAg particles

When purified particles of the full-length HBcAg, either freshly prepared or having been kept for a few months, were subjected to SDS-PAGE in the presence of DTT (100 mM), the protein not only gave a sharp band with an apparent molecular weight of ~22 kDa , but also a smear of less intense bands below the sharp band (Fig. 13a). However, when the arginine-rich carboxy end was truncated, the protein migrated as a single band with the expected molecular mass of the monomer, ~17 kDa (Fig. 13b). To examine the components of the diffuse bands, the products from full-length HBcAg fractionated on a polyacrylamide gel were transferred to a PVDF membrane and stained with Coomassie blue. The sharp band of 22 kDa and diffuse bands were separately excised from the membrane and analysed with an automated protein sequencer (Method 2B.9). The first six amino acids of the N-terminus of the proteins in both the smear and intense bands were sequenced and found to be the same, that is TMITDS. This sequence corresponds to residues 2 to 7 of the β-galactosidase sequence fused to the N-terminus of HBcAg, resulting from the cloning vector. The first amino acid of the β-galactosidase sequence, methionine, was not detected in any of the samples.

3.2.4 Separation of the large and small particles of HBcAg with sucrose gradient centrifugation

As HBcAg assembles into particles, it can be easily separated from other host components with sucrose gradient centrifugation. Fig. 14a shows a typical sedimentation profile of full-length HBcAg particles on a sucrose gradient (8-40% w/v) after spinning at 100,000 x*g* for 5 h. The bell-shape peak of fractions 2 to 10 was found to contain a mixture of different sizes of HBcAg particles by light scattering. As the sedimentation rate is dependent upon particle mass, the early fractions of the peak should be enriched in large particles and later fractions in small particles. The first two fractions and the last fraction of the peak (from 3 sucrose

gradients) were therefore pooled and concentrated separately. These preparations were then overlaid on separate sucrose gradients (8-40%) and centrifugation was repeated as above. It is clear that the HBcAg particles sedimenting at different positions resulted in two separated peaks with some overlapping on the edge of the peaks (Fig. 14b). The first two and the last fractions of peak 1 and 2, respectively, were separately pooled and concentrated, and examined with light scattering.

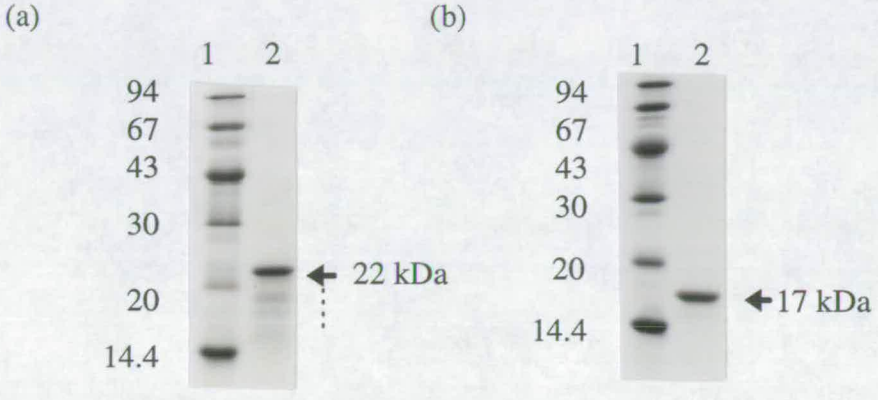


Fig. 13 SDS-PAGE of full-length and truncated HBcAg

Particles comprising purified full-length (a, lane 2) and truncated (b, lane 2) HBcAg (2 μg) were analysed on 15% polyacrylamide gels and stained with Coomassie blue. The specific band is indicated by an arrow head and the smear below the 22 kDa band is indicated by a vertical dashed line. Lane 1 is the molecular weight markers (in kDa).

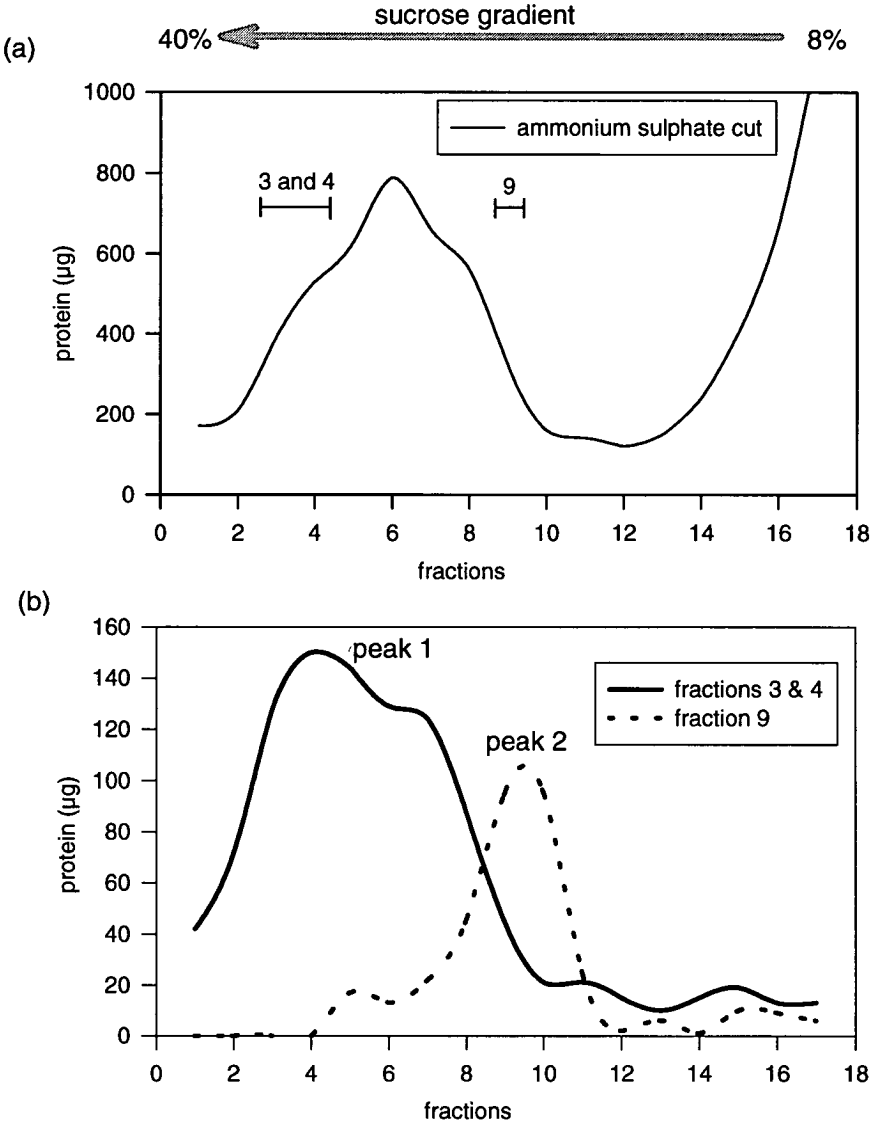


Fig. 14: Fractionation of large and small particles of HBcAg with sucrose gradient centrifugation

(a) Separation of HBcAg particles from host cell components. Ammonium sulphate precipitated fractions from cell extracts containing particles of full-length HBcAg were layered on top of a sucrose gradient and centrifuged at 100,000 xg for 5 h. Fractions (0.5 ml) were collected and protein concentration was determined with the Bradford assay. The arrow on top of the diagram indicates the direction of sedimentation.

(b) Separation of large and small particles of full-length HBcAg. Fractions 3 and 4, and fraction 9 of above sucrose gradient were separately pooled and concentrated, and layered on top of two sucrose gradients. Centrifugation and fractionation were carried out as before. Graphs obtained are superimposed on each other.

3.2.5 Light scattering analysis of unfractionated and fractionated core particles

The results of light scattering measurements on the purified but unfractionated HBcAg particles (corresponding to fractions 3-9 of Fig. 14a) are shown in Table 1. Guidelines used to interpret the polydispersity, base line and sum of squares (SOS) parameters in the monomodal analysis are summarised in the legend of Table 1. The unfractionated HBcAg particles were found to be polydisperse. Tables 2 and 3, respectively, represent the light scattering measurements on fractionated large (fractions 3 and 4 of peak 1, Fig. 14b) and small (fraction 9 of peak 2, Fig. 14b) HBcAg particles. Five to 10 measurements were taken at four different temperatures (4, 10, 20 and 30 or 37 °C) and their mean values are shown in the last row of each sub-table. For measurements indicating polydispersity with a low base line (1.002-1.005) and SOS noise (< 20), a bimodal regression (Autopro, Dynamics) was utilised to resolve the size distribution more accurately, and each component is printed in italics below the measurements.

The sample of large particles was generally monodisperse, for most of the measurements (24 out of 27) indicated a monodisperse size distribution which could be resolved with a single exponential autocorrelation decay function. Although three of the measurements at 4 and 20 °C showed polydispersity, bimodal regression analysis indicated two discrete populations in which a vast majority of the components were large HBcAg particles, as judged by the scattering amplitude recorded for the major component. The contaminants, which were present in a very small amount, are unlikely to be the small HBcAg particles because the molecular weight for the second component is less than one tenth that of the small HBcAg particles. The average D_T increased as the temperature of the sample cell was increased, but the hydrodynamic radius (R_h) decreased and was estimated to be 20.6, 19.2, 18.4 and 18.0 nm at 4, 10, 20 and 30 °C, respectively.

Around one third of the light scattering measurements on the small core particles indicated polydispersity (Table 3), but in general the size distribution of the particles was primarily monomodal as indicated by the low base line (<1.005).

#	Amp[a]	Diffn[b] Coeff	Radius (nm)	Polyd[c] (nm)	Estd MW (kDa)	Base[d] line	SOS[e] Error
1	0.693	67	19.8	11.0	4046	1.021	147.029
2	0.697	66	20.1	11.1	4182	1.025	131.795
3	0.709	63	21.0	11.7	4662	1.022	127.706
4	0.700	65	20.3	11.3	4318	1.022	120.734
5	0.705	64	20.8	11.5	4559	1.022	122.476
6	0.705	67	20.0	11.1	4158	1.016	105.937
7	0.689	67	19.8	11.0	4044	1.024	99.847
8	0.705	65	20.4	11.4	4369	1.015	104.235
χ	0.700	66	20.3	11.3	4292	1.021	119.970

Table 1: Light scattering measurements of unfractionated HBcAg particles at 4°C

[a]*Amplitude: the scattering amplitude recorded for this component. This is proportional to the intensity of the light scattered by particles of a given radius.*

[b]*Translational diffusion coefficient (D_T) for this component ($x10^{-13}m^2/s$)*

[c]*Polydispersity value indicates the standard deviation of the spread of particle sizes (in nanometer) based on the average radius measured. If the value is less than 15% of the average radius measured, it is reported as "--", which indicates a monodisperse solution, otherwise a value is generated for a polydisperse sample.*

[d]*Base line parameter represents the completeness of the regression applied (autocorrelation function) during the analysis. A baseline of 1.000 indicates that the actual size distribution can be fully resolved by the theoretical single exponential autocorrelation function or standard Gaussian distribution. A monomodal regression with a high baseline indicates that the size distribution does not fit within a standard Gaussian distribution. Interpretation of base line parameters is as follows:*

Base line parameter	*Interpretation*
0.997-1.001	*distribution is monomodal*
1.002-1.005	*distribution is primarily monomodal*
> 1.005	*size distribution is polydisperse*

[e]*Sum of squares (SOS) parameter represents the amount of residual noise or error associated with the autocorrelation function. The higher the SOS parameter, the wider the size distribution. Interpretation of SOS parameters is as follows:*

SOS parameters	*Interpretation*
1.000-5.000	*negligible error, low noise*
5.000-20.000	*somewhat significant background error*
> 20.000	*high noise/error due to high polydispersity in size distribution*

#: number or measurement; χ: mean values; Estd MW: estimated molecular weight

(a) at 4 ± 0.5 °C

#	Amp[a]	Diffn[b] Coeff	Radius (nm)	Polyd[c] (nm)	Estd MW (kDa)	Base[d] line	SOS[e] Error
1	0.665	65	20.3	----	4296	1.000	5.531
2	0.663	64	20.7	----	4484	1.004	3.250
3	0.662	64	20.8	----	4566	1.003	3.529
4	0.674	65	20.6	----	4436	1.002	4.422
5	0.680	65	20.5	----	4396	1.006	6.285
6	0.681	62	21.3	7.5	4826	0.999	4.452
	0.670	*60*	*22.0*	----	*5212*	*0.997*	*1.678*
	0.025	*936*	*1.4*	----	*6*		
7	0.672	64	20.6	6.0	4437	1.011	3.101
	0.629	*59*	*22.1*	----	*5252*	*1.008*	*1.950*
	0.050	*186*	*7.1*	----	*336*		
8	0.689	66	20.1	----	4176	1.002	5.483
9	0.694	64	20.7	----	4486	1.005	3.643
10	0.706	65	20.3	----	4314	1.002	2.900
χ	0.679	64	20.6	1.3	4442	1.003	4.260

(b) at 10 ± 0.3 °C

1	0.608	83	19.0	----	3660	1.003	2.708
2	0.550	82	19.5	----	3891	1.009	3.147
3	0.595	82	19.4	----	3831	1.005	2.031
4	0.621	83	19.1	----	3733	1.003	3.047
5	0.642	84	18.9	----	3599	1.000	1.870
χ	0.603	83	19.2	0.0	3743	1.004	2.561

(c) at 20 ± 0.3 °C

1	0.742	113	18.5	5.4	3453	1.001	3.766
	0.734	*110*	*18.9*	----	*3602*	*1.000*	*1.358*
	0.030	*2572*	*0.8*	----	*1*		
2	0.731	115	18.4	----	3377	1.000	2.856
3	0.723	115	18.4	----	3409	1.000	1.584
4	0.713	116	18.1	----	3256	1.000	3.162
5	0.722	114	18.5	----	3419	0.996	0.725
6	0.719	116	18.3	----	3330	1.002	1.627
χ	0.725	115	18.4	0.9	3374	1.000	2.287

(d) at 30 ± 0.3 °C

1	0.747	152	17.9	----	3167	1.001	1.140
2	0.724	152	18.1	----	3241	1.002	3.521
3	0.701	153	17.9	----	3158	1.002	2.083
4	0.675	150	18.1	----	3281	1.001	3.710
5	0.663	151	17.6	----	3265	0.998	14.529
6	0.645	155	17.4	----	3034	1.001	7.574
χ	0.692	152	18.0	0.0	3139	1.001	5.426

Table 2: Light scattering measurements of fractionated large HBcAg particles

Description of the data measured and calculated is shown in Table 1. The data for bimodal analysis are printed in italics. The temperature of the sample cell is shown on top of each table.

(a) at 4 ± 0.3 °C

#	Amp[a]	Diffn[b] Coeff	Radius (nm)	Polyd[c] (nm)	Estd MW (kDa)	Base[d] line	SOS[e] Error
1	0.695	73	18.0	----	3226	1.010	3.564
2	0.707	72	18.4	7.2	3411	1.004	6.748
	0.643	*80*	*16.5*	----	*2582*	*0.984*	*3.702*
	0.087	*18*	*71.8*	----	*92162*		
3	0.711	72	18.5	----	3427	1.003	4.473
4	0.714	73	18.1	5.3	3279	1.008	3.772
	0.708	*72*	*18.4*	----	*3393*	*1.007*	*1.220*
	0.029	*2844*	*0.5*	----	*0*		
5	0.718	72	18.6	----	3489	1.004	4.627
6	0.704	73	18.2	----	3305	1.004	3.383
χ	0.708	73	18.3	2	3356	1.005	4.428

(b) at 10 ± 0.3 °C

1	0.676	91	17.5	----	2984	1.005	3.372
2	0.664	90	17.8	----	3122	1.001	4.671
3	0.648	92	17.3	6.8	2904	1.004	3.907
	0.444	*117*	*13.5*	----	*1610*	*0.997*	*1.461*
	0.214	*54*	*29.1*	----	*10325*		
4	0.651	91	17.5	----	2984	1.001	4.228
5	0.650	91	17.6	----	3032	1.002	3.581
6	0.633	90	17.7	6.9	3080	1.001	4.689
	0.612	*85*	*18.7*	----	*3525*	*0.999*	*1.867*
	0.035	*627*	2.5	----	*27*		
χ	0.654	91	17.6	2.3	3018	1.002	4.075

(c) at 20 ± 0.3 °C

1	0.734	128	16.4	5.8	2557	1.004	4.653
	0.716	*122*	*17.0*	----	*2805*	*1.003*	*1.831*
	0.034	*1101*	*1.9*	----	*13*		
2	0.735	129	16.4	----	2556	1.003	2.535
3	0.706	128	16.5	----	2596	1.002	4.325
4	0.716	130	16.2	5.6	2490	1.001	4.571
	0.705	*126*	*16.7*	----	*2659*	*1.000*	*1.264*
	0.035	*2221*	0.9	----	*2*		
5	0.706	127	16.6	----	2630	1.000	2.958
χ	0.719	128	16.4	2.3	2566	1.002	3.808

(d) at 37 ± 0.4 °C

1	0.643	204	15.9	----	2382	1.003	9.843
2	0.617	199	16.2	6.3	2498	1.002	9.668
	0.363	*277*	*11.6*	----	*1111*	*0.998*	*6.357*
	0.265	*128*	*25.0*	----	*7149*		
3	0.591	210	15.6	----	2267	1.005	17.328
4	0.553	197	16.4	----	2567	1.004	10.592
5	0.525	200	16.3	----	2520	0.999	18.088
6	0.490	209	16.5	----	2240	1.005	29.791
χ	0.570	203	16.0	1.1	2412	1.003	15.885

Table 3: Light scattering measurements of fractionated small HBcAg particles

Description of the data measured and calculated is shown in Table 1. The data for bimodal analysis are printed in italics. The temperature of the sample cell is shown on top of each table.

Bimodal analysis of the polydisperse measurements showed that the solution contains a small amount of larger size molecules with a radius two- to five-times bigger than the average radius of small HBcAg particle (# 2 of a, # 3 of b, and # 2 of d). Small molecules, less than 30 kDa, were also detected (# 4 of a, # 6 of b, # 2 and # 4 of c), but the solution was free of the large HBcAg particle. The average radius of small HBcAg particle was estimated to be 18.3, 17.6, 16.4 and 16.0 nm at 4, 10, 20 and 37 °C, respectively.

3.2.6 Reducing and non-reducing SDS-PAGE of HBcAg particles

Fig. 15a shows the results of SDS-PAGE of unfractionated truncated, unfractionated full-length, and fractionated large and small full-length HBcAg particles, under reducing or non-reducing conditions. The densities of bands in lanes 2 to 9 of the gel were quantitated with ImageQuant software (Molecular Dynamics) which reported the results as line graphs and the area under each peak was expressed as a percentage of the total (Fig. 15b).

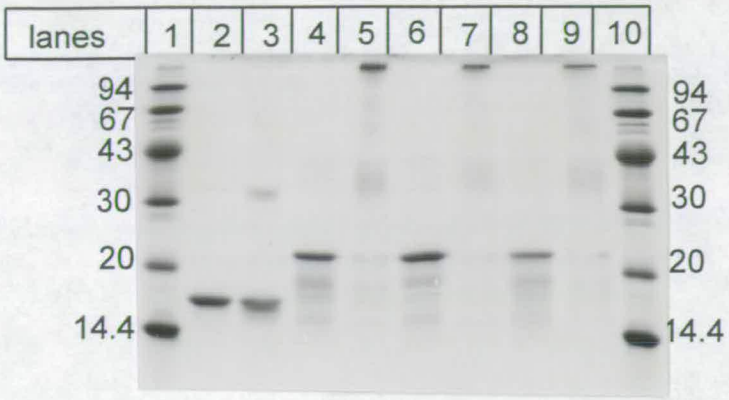
The unfractionated truncated HBcAg protein migrated as a single band at the expected apparent molecular mass of a monomer, ~17 kDa, under reducing conditions (Fig. 15a and b, lane 2). Under non-reducing conditions, ~83% of the protein migrated as monomers of ~17 kDa, and ~17% as dimers of ~34 kDa (Fig. 15a and b, lane 3). When unfractionted full-length, and fractionated large and small full-length HBcAg particles were analysed on reducing SDS-PAGE, ~50% of the protein migrated as a sharp band with an apparent molecular mass of ~22 kDa, and the other half of the protein migrated with a greater mobility forming a diffuse banding pattern (Fig. 15a and b, lanes 4, 6 and 8). Under non-reducing conditions, ~40% of the protein did not enter the gel, ~ 20 to 30% migrated as a diffuse band in the region of the dimer with apparent molecular mass between 34 and 42 kDa, and ~30% migrated as monomer and smear bands as observed under the reducing conditions (Fig. 15a and b, lanes 5, 7, and 9).

Fig. 15:

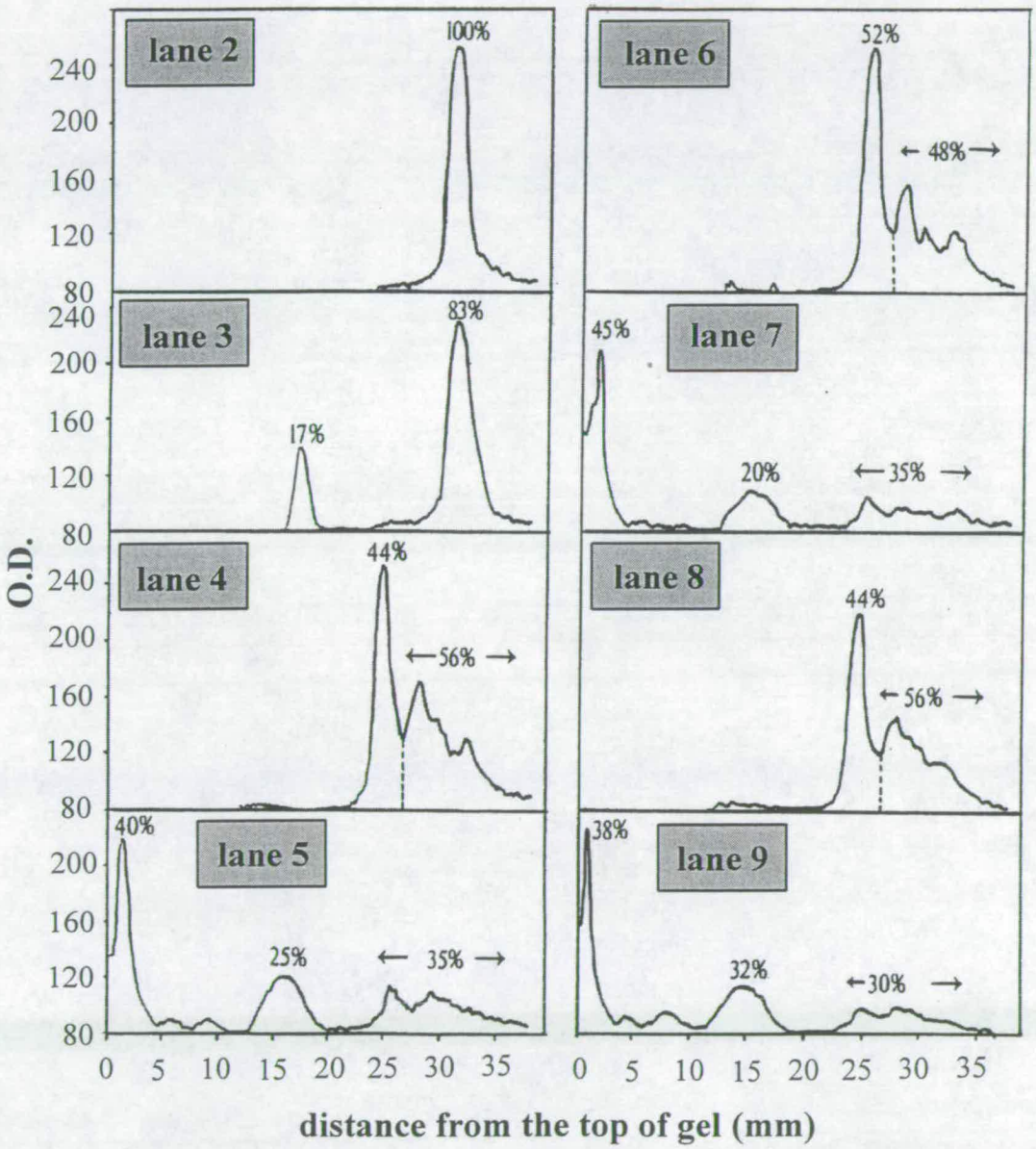
(a) SDS-PAGE of HBcAg particles under reducing and non-reducing conditions.
An aliquot (2 μg) of unfractionated particles of truncated (lanes 2 and 3), and full-length HBcAg (lanes 4 and 5), fractionated large particles of full-length HBcAg (lanes 6 and 7), and fractionated small particles of full-length HBcAg (lanes 8 and 9) was added to 2x SDS loading buffer (1 vol) in the presence of DTT (200 mM, reducing conditions, lanes 2, 4, 6, and 8) or absense of DTT (non-reducing conditions, lanes 3, 5, 7, and 9). All samples, reduced or non-reduced, were boiled for 5 min prior to electrophoresis. Cooled samples were loaded into the wells of a 15% SDS-polyacrylamide gel for electrophoresis. The gel was stained with Coomassie blue and dried between two sheets of cellulose membrane backing. Lanes 1 and 10 are molecular weight markers under reducing condition.

(b) Scanning and quantitation of bands along gel tracks of the above gel.
With ImageQuant software a rectangle was drawn along lanes 2 to 9 and the average pixel value along the rectangle was integrated and the results plotted as a line graph. Each graph corresponds to a lane of the above gel which is shown in the upper left corner of each graph. The area under the peaks was expressed as a percentage of the total and is shown on top of the peaks. The units on the y axis are arbitrary.

(a)


lanes
1
2
3
4
5
6
7
8
9
10
94
67
43
30
20
14.4

(b)


lane 2
100%
lane 6
52%
← 48% →
lane 3
83%
17%
lane 7
45%
20%
← 35% →
lane 4
44%
← 56% →
lane 8
44%
← 56% →
lane 5
40%
25%
← 35% →
lane 9
38%
32%
← 30% →
O.D.
240
200
160
120
80
0 5 10 15 20 25 30 35
distance from the top of gel (mm)

3.3 Discussion

HBcAg is highly enriched in Arg residues particularly in the C-terminal region. 46% of these residues are encoded by the triplet AGA and more importantly, 67% of these codons in the reading frame are clustered at the 3'-end of the HBcAg mRNA. Codons AGA and AGG, are amongst the rarest codons in genes of *E. coli*, AGA occurring at the level of 2.4% of total Arg codons and 0.14% of all codons (Aota *et al.*, 1988). Direct measurements of the level of AGA tRNA have shown that it is present in extremely low concentration during all phases of cell growth (Ikemura, 1982; Saxena and Walker, 1992; Emilsson *et al.*, 1993). As far as the HBcAg gene is concerned, there is overall bias towards the use of this extremely rare codon to synthesise the product. Furthermore, it has been known through the works of Zheng *et al* (1992) and Stewart (1993) that the expression of truncated HBcAg, which lacks the Arg-rich C-terminal region, is much greater than that of full-length HBcAg.

The result presented in section 3.2.2 showed that supplementation of *E. coli* with T4 AGA tRNA resulted in a 50% increment in HBcAg synthesis as measured by SDS-PAGE of crude cell extracts. A two-fold increment was observed when the sucrose gradient purified HBcAg particles was measured by the Bradford assay. The former quantitation method showed a lower increment in the protein synthesis because only the 22 kDa intense band was taken into consideration and it is clear from the SDS-PAGE of purified HBcAg particles that the HBcAg monomers, not only migrated as an intense band with an apparent molecular weight of 22 kDa, but nearly half of the proteins migrated as a smear below the sharp band. If the less intense bands were taken into account the increment is expected to be higher than 50%.

Further analysis of the smear bands with sequencing showed that the proteins have the same N-terminal sequence with the 22 kDa monomer. This suggests that around 50% of the HBcAg monomers have probably been proteolysed near the C-terminus, or prematurely terminated during the translation process due to 'queueing' of ribosomes at the 3'-end of HBcAg mRNA which is extremely rich in Arg codons. Attempts were made to eliminate the smaller species and hence improve the yield of

HBcAg by expressing the protein in a protease deficient strain (SG935), but no improvement was achieved. An unsuccessful effort was also reported by Zheng *et al.*, (1992) who tried to eliminate the low molecular weight species by inclusion of protease inhibitors. Besides, it has been shown by several groups that rare codons can induce premature termination (Robinson *et al.*, 1984; Misra and Reeves, 1985); frameshifting (Spanjaard and Van Duin, 1988), and decoupling of transcription from translation (Bonekamp and Jensen, 1988; Sorensen *et al.*, 1989), which in turn can lead to strong negative overall effects on the gene expression. Therefore it is likely that the rare Arg codons at the 3'-end of the HBcAg mRNA causes heterogeneity of the monomer which migrated with different mobility in a SDS-polyacrylamide gel.

The yield of full-length HBcAg is significantly improved by T4 AGA tRNA supplementation but the level of expression is still not as high as that of truncated HBcAg, which was estimated to be five-fold higher than that of full-length HBcAg. This suggests that other factors might contribute towards the low level. These include: (i) the protamine-like region might be, to some extent, toxic to the host cell because several groups have demonstrated that this region has a non-specific nucleic acid binding motif which incorporates the host cell's nucleic acids into HBcAg particles (Borissova *et al.*, 1989; Gallina *et al.*, 1989; Birnbaum and Nassal, 1992; Ulrich *et al.*, 1993), and indirectly, this might affect the overall level of gene expression; (ii) the mRNA of full-length HBcAg might adopt different secondary structure from that of truncated HBcAg, because the former is longer and its 3'-end contains a high frequency of Arg codons, which could make the transcript less stable and more susceptible to degradation by RNase.

Using cryoelectron microscopy and image reconstruction, Crowther *et al.*, (1994) showed the large and small particles of HBcAg comprise 240 and 180 subunits, respectively. Assuming the molecular mass of a subunit as 22 kDa and no nucleic acid is incorporated in HBcAg particles, the large HBcAg particle would have a mass of ~5,280 kDa, and the small particle ~3,960 kDa. By exploiting the difference in mass between the two different sizes of HBcAg particle, it is possible, in principle, to separate these particles with sucrose gradient centrifugation. Indeed, as expected, the large and small HBcAg particles sediment at different positions on a 8-40% sucrose gradient at 100,000 *xg*. Nearly five months after this observation,

Zlotnick *et al.* (1996) published a similar finding. Using an analytical ultracentrifugation they showed that large and small particles sediment at 45*s* and 40*s*, respectively.

Light scattering analysis of the two species resolved on a sucrose gradient showed that their size distribution is primarily monodisperse and no appreciable cross contamination was observed between the two species. In some measurements, particularly for small HBcAg particles, small amounts of lower and higher molecular weight particles were detected. This may be due to broken or aggregated HBcAg particles during preparation of the sample for analysis. It appears that the hydrodynamic radius (R_h) of the large and small HBcAg particles, derived from D_T, using the Stokes-Einstein equation (Method 2B.8.*v*), is reduced when the temperature of measurements was increased (from 4 °C to 30 °C or 37 °C). A possible explanation for this observation is that the formula in the software linked directly to the light scattering instrument assumes that the viscosity of water ($\eta=1.019\text{x}10^{-3}$ Nsm^{-2}) remains constant at different temperatures. In fact the viscosity of the solution used (50 mM Tris-HCl, 100 mM NaCl) is higher than this and decreases as the temperature increases. The average R_h calculated for the large and small HBcAg particles at different temperatures (4, 10, 20 and 30 or 37 °C) is estimated to be 19 and 17 nm, respectively, which is 2 nm longer than that determined by Crowther and colleagues (1994) with cryoelectron microscopy and image reconstruction. However, the difference in size (diameter) between the large and small HBcAg particles determined by light scattering is in good agreement with their measurements.

Both the large and small HBcAg particles gave identical results when analysed on SDS-PAGE, under reducing or non-reducing conditions, suggesting that the particles are made up from the same building blocks. Under reducing conditions, these particles showed heterogeneity in monomer, even though the particles were generally found to be monodisperse in size distribution. Zheng *et al.*, (1992) reported that the last amino acid of the monomer, Cys 183, is always involved in a disulphide bond with the Cys 183 of a different monomer to form a dimer, and Cys 61 always forms an additional disulphide bond with the identical residue between

two dimers. If this is, indeed, the case and all the full-length HBcAg monomers are homogenous in length, the protein would have not entered the gel under non-reducing conditions due to the formation of a large lattice of protein between the monomers. However, it appears that only a part of the protein (40%) did not enter the gel and the rest migrated as a dimer of variable size and corresponding monomer bands. The N-terminal sequence analysis showed that some of the monomers most probably lack the C-terminus, either caused by premature termination of translation or protease degradation, thus eliminating the possibility of disulphide bond formation between two Cys 183 residues. Another possibility is that Cys 61 is partly, but not always, involved in interchain disulphide bonds because some of the full-length HBcAg did, in fact, migrate as a monomer and, moreover, most of the truncated HBcAg (~80%) migrated as monomer under non-reducing conditions.

In summary, the three questions addressed earlier in this chapter have been answered: (i) the yield of full-length HBcAg expressed in *E.coli* can be improved by T4 AGA tRNA supplementation. This approach substantially reduced the time and cost spent to produce and purify the protein in bulk for further characterisation; (ii) the large and small HBcAg particles can be separated by sucrose gradient centrifugation. This method, although time consuming, is capable of consistently producing homogeneous HBcAg particles, and allows studies to be done on the interaction of the two different sizes of HBcAg particle with surface proteins and fusion-phages; (iii) clearly, full-length HBcAg subunits that assemble into HBcAg particles are heterogeneous at their C-terminus. However, more work would have to be carried out to clarify whether the heterogeneity is caused by incomplete synthesis of the protein or protease degradation of the carboxy end.

CHAPTER 4

MUTAGENESIS AND *IN VITRO* EXPRESSION OF SURFACE ANTIGEN

4.1 Introduction

The mutagenesis strategies applied in this study were based upon the information obtained from affinity selection from bacteriophage display libraries against HBcAg (Dyson and Murray, 1995). A peptide sequence of LLGRMK and some related sequences that should mimic the docking site on the HBcAg particle were selected; moreover, the interaction between HBsAg and HBcAg was inhibited by the synthetic peptide ALLGRMKG. However, this specific sequence was not found in HBsAg, but two related sequences did occur: residues 63-65 (LLG) in the PreS1 region and residues 21-26 (LLTRIL) of the S region. This suggests that at least some of the amino acids of the hexapeptide sequence may be brought together in the folded protein of HBsAg to form an essential contact area for HBcAg. Therefore, progressive deletion of the N- and C-termini of L-HBsAg was performed to define the minimum stretch of amino acids that contains the exact contact residues.

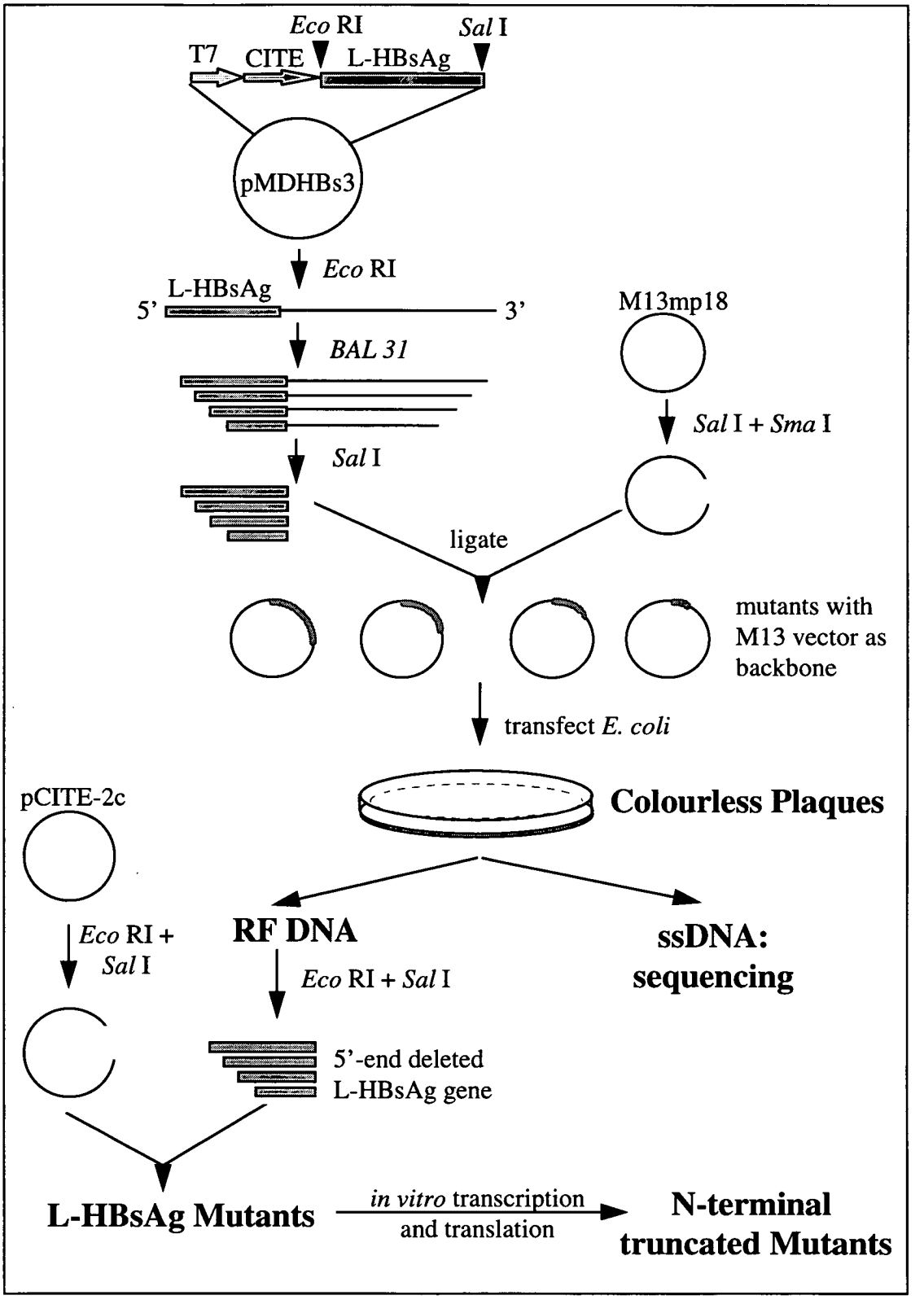
4.2 Results

4.2.1 N-terminal deletion mutagenesis of L-HBsAg with *BAL 31* exonuclease

The steps involved in generating a series of N-terminal truncated L-HBsAg are schematically represented in Fig. 16: (i) plasmid pMDHBs3, which contains an insert of 1.35 kb L-HBsAg coding region, was linearized with *Eco* RI, (ii) a set of deletions from both ends of the plasmid was generated by digestion with *BAL 31* exonuclease, (iii) the 5'-end deleted L-HBsAg coding sequence was separated from the vector by digestion with restriction endonuclease *Sal* I and agarose gel electrophoresis, (iv) the deleted fragments were then eluted from the gel and ligated to M13mp18 vector that had been digested with *Sal* I and *Sma* I, (v) ligation products were used to transform *E. coli* strain TG-1, (vi) well isolated colourless plaques of transformants were propagated in TG-1 cells, ssDNA was extracted from the bacteriophage and sequenced, (vii) RF DNA harbouring the desired deletions were extracted from the cells, (viii) the deleted L-HBsAg DNAs were released from

Fig. 16: Schematic representation of the steps involved in generation of a series of N-terminal deletion mutants of L-HBsAg

Plasmid pMDHBs3 contains an insert of 1.35 kb of the L-HBsAg gene. The vector is located in the polycloning site of the pCITE-2c vector, with the Eco RI and Sal I sites at the 5'- and 3'-ends, respectively. Immediately upstream of this insert is a 586 bp copy of the encephalomyocarditis virus (EMC) RNA 5' non-coding region which initiates translation by eukaryotic ribosomes. This region is known as Cap-Independent Translation Enhancer or CITE. The CITE sequence is located downstream of the T7 promoter to provide efficient transcription with T7 RNA polymerase. Refer to text for the explanations of the steps involved in generation of N-terminal deletion mutants(section 4.2.1, pg 75).


Eco RI
Sal I
T7
CITE
L-HBsAg
pMDHBs3
Eco RI
L-HBsAg
5'
3'
BAL 31
Sal I
M13mp18
Sal I + Sma I
ligate
mutants with M13 vector as backbone
transfect E. coli
pCITE-2c
Colourless Plaques
Eco RI + Sal I
RF DNA
ssDNA: sequencing
Eco RI + Sal I
5'-end deleted L-HBsAg gene
L-HBsAg Mutants
in vitro transcription and translation
N-terminal truncated Mutants

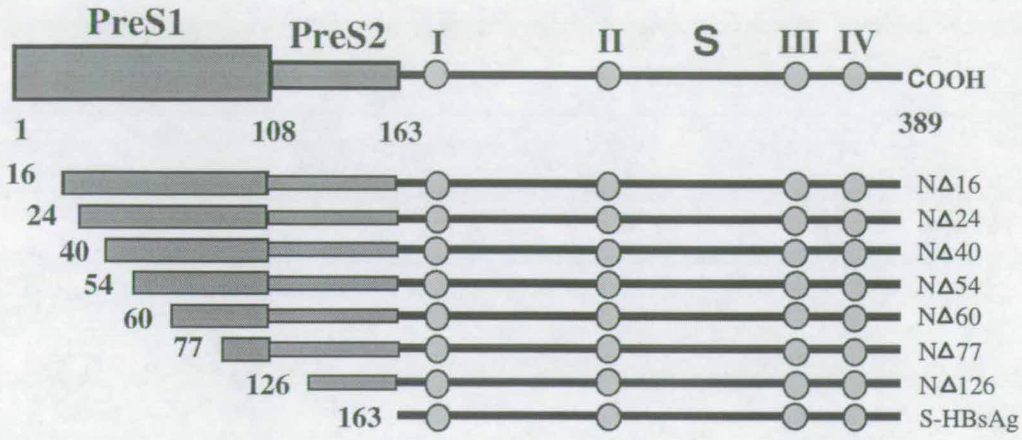
the M13mp18 vector and inserted into pCITE-2c vectors, (ix) *in vitro* transcription and translation were carried out to produce the N-terminal truncated L-HBsAg.

A total of 69 transformants were analysed with DNA sequencing, from which 30 mutants ranging from 1 to 150 amino acids deleted from the N-terminus of L-HBsAg were obtained. Around two thirds of the deletions altered the ORF of the L-HBsAg gene. Mutants NΔ16, NΔ24, NΔ40, NΔ54, NΔ60, NΔ77 and NΔ126 gave rise to a correct reading frame in the pCITE-2c vector were transcribed with T7 RNA polemerase, and their mRNAs were translated with rabbit reticulocyte lysates supplemented with [^{35}S]-methionine. To ensure that the proteins expressed were of the correct size, aliquots of the translation reactions were subjected to SDS-PAGE (Fig. 17). Translation of the L- and S-HBsAg mRNAs produced bands of 39 kDa and 24 kDa, respectively, which correspond to the unglycosylated forms of the L- and S-HBsAg. Translation of the deletion mutants gave rise to protein bands that decrease gradually in molecular mass depending on the number of amino acids that have been deleted. In addition, the intensity of the deletion protein bands decreased, especially NΔ126 and S-HBsAg, due to the deletion of the Met residues in the PreS regions.

4.2.2 C-terminal deletion mutagenesis of L-HBsAg

The coding segment of L-HBsAg was excised from pMDHBs3 and cloned into the RF DNA of M13mp18 that had been digested with *Eco* RI and *Sal* I. Uracil-containing ssDNA was prepared from *E. coli* strain CJ236 and stop codons were introduced at various positions in the coding sequence by oligonucleotide-mediated mutagenesis. The C-terminal deletion mutants are schematically represented in Fig. 18a. and the stop codon introduced is indicated at the end of each construct. The five shortest mutants (CΔ108, CΔ136, CΔ163, CΔ191and CΔ210) were constructed by Dr. M.R. Dyson, and the nucleotide changes for the other mutants (CΔ242, CΔ263, CΔ322, CΔ346 and CΔ371) as determined by DNA sequencing are shown in Fig. 18b. The coding sequences of these mutants were then excised from M13mp18 and inserted into pCITE-2c vectors to serve as template for transcription.

(a)



(b)

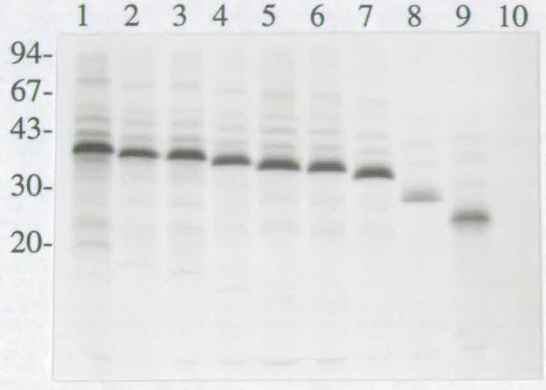


Fig. 17(a): Schematic representation of L-HBsAg and the N-terminal deletion mutants.

The diagram at the top represents the postulated domain organisation of L-HBsAg. The L-HBsAg of HBV subtype adyw is 389 amino acids long with three major regions: PreS1 (residues 1-108); PreS2 (residues 109-163); and S (residues 164-389). Four predicted transmembrane helices (Stirk et al., 1992) in the S region are indicated by circles: I (residues 171-191); II (residues 240-261); III (residues 323-347); and IV (residues 352-372). The lines beneath the L-HBsAg represent the N-terminal deletion mutants. The number at the beginning of each mutant indicates the number of residues deleted from the N-terminus of L-HBsAg. The abbreviated descriptions of the mutants are shown at the right of each mutant

(b): In vitro translation products of the N terminal deletion mutants

Lanes: 1 (L-HBsAg); 2 (NΔ16); 3 (NΔ24); 4 (NΔ40); 5 (NΔ54); 6 (NΔ60); 7 (NΔ77); 8 (NΔ126); 9 (S-HBsAg); and 10 (negative control; minus mRNA). Molecular weight standards in kDa are shown on the left of the gels. Aliquots (0.5 μl) of in vitro translation reaction mixtures were separated by 15% SDS-PAGE. The gel was dried and the products visualised by autoradiography after overnight exposure.

Fig. 18:

(a) Schematic representation of L-HBsAg and C-terminal deletion mutants

The diagram at the top represents domain organisation of L-HBsAg. C-terminal deletion mutants are shown beneath the L-HBsAg. The number at the end of each mutant indicates the final amino acid in the mutant. Each mutant was named as shown at the left.

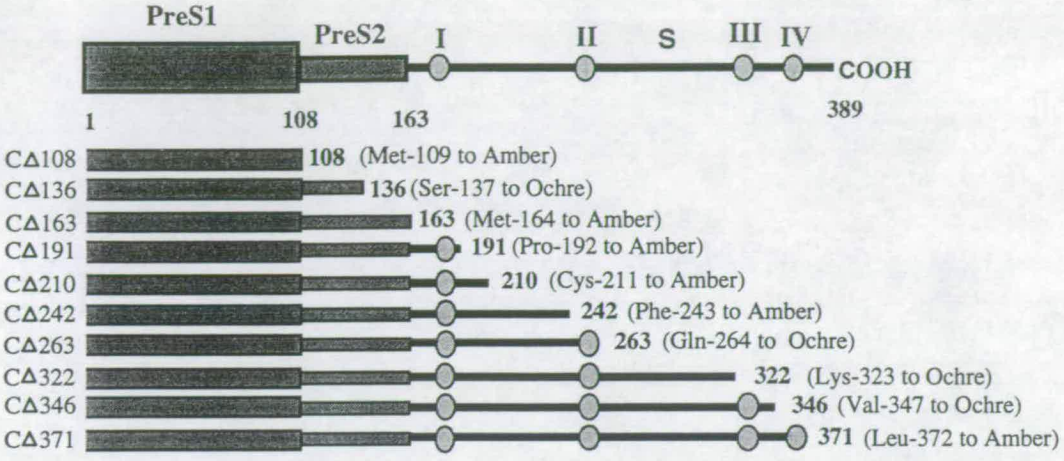
(b) Nucleotide sequence of C-terminal deletion mutants

*Each panel represents an electropherogram obtained from automated sequencing of a mutant. The nucleotide sequence of the non-coding strand is shown, read from left to right, 3' to 5'. The nucleotide changes are indicated by *. Stop codons introduced in the mutants are as follows (expressed as coding sequence, i.e. complementary to the sequence read from the traces): CΔ242 (TTT-243→TAG); CΔ263 (CAA-264→TAA); CΔ322 (AAA-323→TAA); CΔ346 (GTA-347→TAA) and; CΔ371 (TTG-372→TAG).*

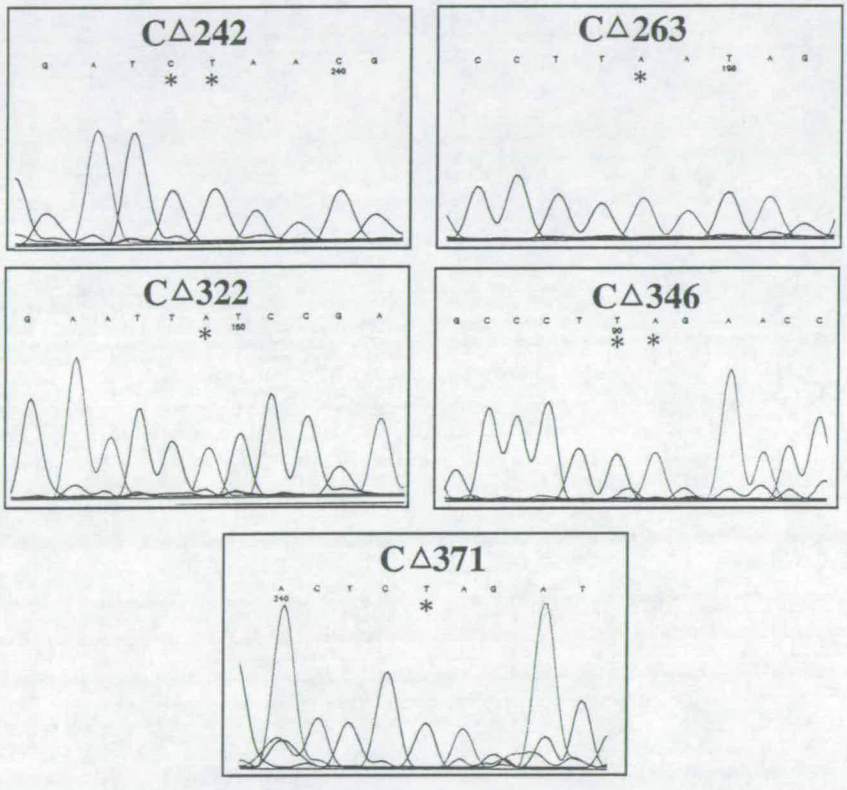
(c) In vitro translation products of the C-terminal deletion mutants

Lanes: 1 (CΔ108); 2 (CΔ136); 3 (CΔ163); 4 (CΔ191); 5 (CΔ210); 6 (CΔ242); 7 (CΔ263); 8 (CΔ322); 9 (CΔ346); 10 (CΔ371); and 11 (L-HBsAg). Molecular weight standards in kDa are shown on the left of the gels. Aliquots (0.5 μl) of in vitro translation reaction mixtures were separated by 15 % SDS-PAGE. The gel was dried and the products visualised by autoradiography.

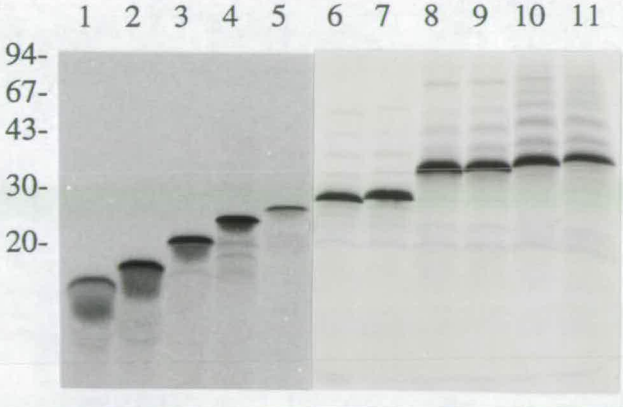
(a)


PreS1
PreS2
I
II
S
III
IV
COOH
1
108
163
389
CΔ108
108 (Met-109 to Amber)
CΔ136
136 (Ser-137 to Ochre)
CΔ163
163 (Met-164 to Amber)
CΔ191
191 (Pro-192 to Amber)
CΔ210
210 (Cys-211 to Amber)
CΔ242
242 (Phe-243 to Amber)
CΔ263
263 (Gln-264 to Ochre)
CΔ322
322 (Lys-323 to Ochre)
CΔ346
346 (Val-347 to Ochre)
CΔ371
371 (Leu-372 to Amber)

(b)


CΔ242
CΔ263
CΔ322
CΔ346
CΔ371

(c)


1 2 3 4 5 6 7 8 9 10 11
94-
67-
43-
30-
20-

In vitro translation was subsequently performed in rabbit reticulocyte lysates supplemented with [^{35}S]-methionine and the products analysed on SDS-PAGE (Fig. 18c).

4.2.3 Random mutagenesis of LLG (63-65) of the PreS1 region

Since the peptide sequence LLG (63-65) was implicated in the mimotope formation, oligonucleotide-mediated mutagenesis was carried out to change these amino acids. Two degenerate oligonucleotides were designed: L64Random and L63/G65Random. L64Random introduces changes at codon 64 (Leu) to 14 other possibilities, whereas L63/G65Random mutates Leu (63) and Gly (65) together to give rise to 196 possible mutations.

For the first attempt of this experiment, a 30:1 ratio of oligonucleotide to template was used. Twenty of the transformants derived from the mutagenesis reaction with L64Random were sequenced but no mutant was obtained. From the mutagenesis reaction with L63/G65Random, 22 out of the 36 sequenced transformants proved to be mutants (Table 4a), but most of them were single base changes. This shows that the conditions used favoured single base mutation.

In order to introduce mutations at codon 64 (Leu) using oligonucleotide L64Random and to increase the double mutation at codon 63 (Leu) and 65 (Gly) with oligonucleotide L63/G65Random, the oligonucleotide:template ratio was reduced to 10:1. In addition, the oligonucleotide-template annealing conditions were altered: for L64Random, the mixture was heated at 80 °C for 5 min and then plunged immediately into ice water; for L63/G65Random, the mixture was heated at 80 °C for 5 min and then allowed to cool down at room temperature. Ten transformants from each mutagenesis reaction were sequenced. The former reaction produced three mutants (Table 4b) and nine out of the ten transformants of the latter reaction were mutants (Table 4c), most of them being double base changes. Mutants L63H, L64E, G65R, L63H/G65R and L63A/G65E were selected for further analysis and their nucleotide changes are shown in Fig. 19.

(a)

L(63)	L(64)	G(65)	no of transformants
CTT	TTG	***GG***G	14
A (H)			2
G (R)			1
C (P)			1
GC (A)			1
A (H)		C (R)	1
		C (R)	4
		A (R)	2
		T (W)	5
		C (A)	1
		A (E)	1
		T (V)	3

(b)

L(63)	L(64)	G(65)	no of transformants
CTT	***TT***G	GGG	7
	GA (E)		1
	G (V)		1
	C (S)		1

(c)

L(63)	L(64)	G(65)	no of transformants
CTT	TTG	***GG***G	1
C (P)			1
GC (A)		A (E)	1
C (P)		AC (R)	1
GC (A)		C (R)	1
AC (T)		A (R)	1
		AT (M)	1
		CC (P)	1
		TA (amber)	1
		C (R)	1

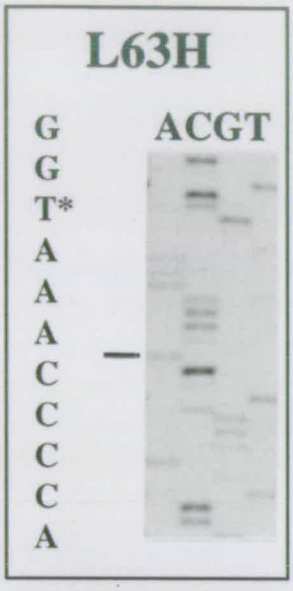
Table 4: Random Mutagenesis of LLG(63-65) of PreS1

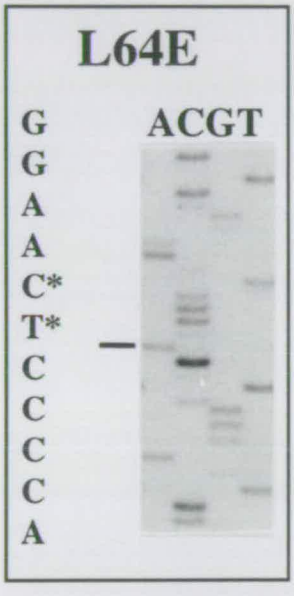
The codons encoding LLG(63-65) are shown on the second row of each table, underline indicates the mutated nucleotides. The nucleotide changes for every mutant are shown beneath the codon and the corresponding amino acid is shown in parentheses. (a) The mutagenesis reaction was performed with oligonucleotide L63/G65Random using a 30 to 1 primer:template ratio, heated at 80 °C for 5 min and allowed to cool slowly to room temperature by putting in water at 80 °C. (b) The mutagenesis reaction was performed with oligonucleotide L64Random using a 10 to 1 primer:template ratio, heated at 80 °C for 5 min and plunged into ice water. (c) The mutagenesis reaction was performed with primer L63/G65Random using the conditions in (b) but the primer-template mixture was allowed to cool to room temperature.

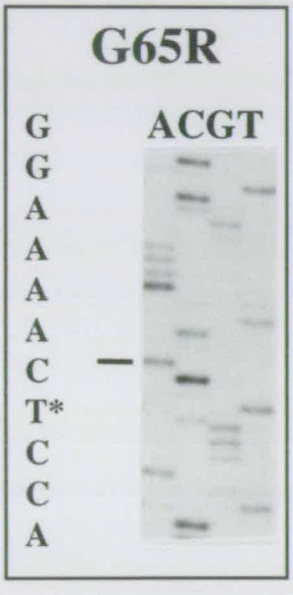
Fig. 19: Nucleotide sequence of mutants at codons 63-65 (LLG) of PreS1 region.

*The nucleotide sequences of the non-coding strands are shown on the left, starting from the horizontal line to the left of the photographs of the autoradiographs and reading upwards, 5' to 3'. Lanes from left to right represent A, C, G and T. Mutants L63H, G65R, L63H/G65R and L63A/G65E were generated using mutated oligonucleotide L63/G65Random. Mutant G64E was generated using oligonucleotide L64Random. Mutated nucleotides are indicated by *. The codons encoding LLG (63-65) and the changes introduced are shown in Table 4.*

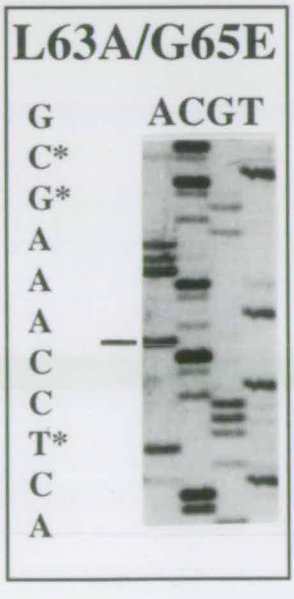

wild-type
G
G
A
A
A
A
C
C
C
C
A
ACGT


L63H
G
G
T*
A
A
A
C
C
C
C
A
ACGT


L64E
G
G
A
A
C*
T*
C
C
C
C
A
ACGT


G65R
G
G
A
A
A
A
C
T*
C
C
A
ACGT


L63H/G65R
G
G
T*
A
A
A
C
G*
C
C
A
ACGT


L63A/G65E
G
C*
G*
A
A
A
C
C
T*
C
A
ACGT

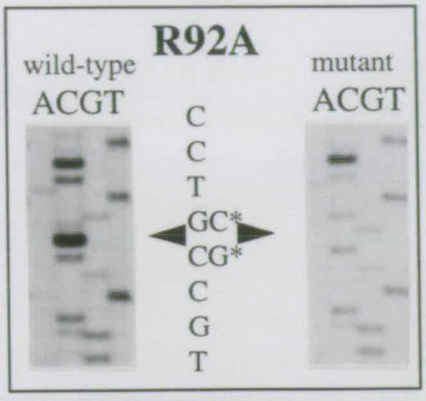
4.2.4 Substitution of positively charged amino acids with alanine

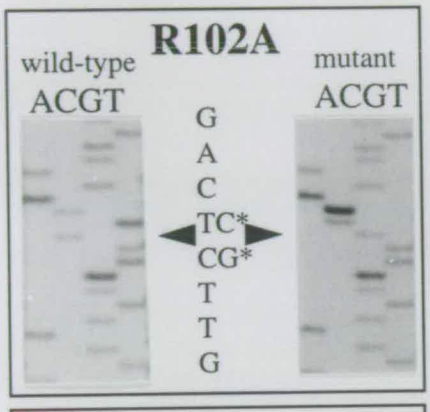
Since two out of six residues of the selected sequence, LLG**R**M**K**, from the fusion phage library are positively charged, it is likely that some of the positively charged amino acids in the PreS regions and helix I, might be brought together to form a discontinuous docking site for HBcAg. Furthermore, charged amino acids are more likely to be located on the surface of a protein (Chothia, 1976; Janin 1979), and substitution of these residues is less likely to cause structural perturbations because residues exposed to solvent are more tolerant to variation (Bowie *et al.*, 1990). The charged residues are also capable of forming hydrogen bonds and ion pairs and are therefore likely to be involved in the recognition of ligands. To investigate this further, Lys-38 and Arg at positions 88, 92, 102, 124, 126, 156 and 187 were individually substituted with Alanine. Alanine was chosen as a replacement residue because it is commonly located in both buried and exposed positions, and most importantly, it does not contribute new hydrogen bonding, sterically bulky, or unusual hydrophobic side chains. As shown in Fig. 20 the first two nucleotides of the codons code for these residues were mutated to G and C by site directed mutagenesis. It is quite likely that more than one residue from different segments of L-HBsAg may be involved in the contact with HBcAg, therefore two to five amino acids were substituted simultaneously to evaluate the contribution of certain amino acids towards the interaction (Fig. 21). An additional quintuple mutant (L63H/R88A/R92A/R102A/R187A+4a.a), containing a 4-amino acid (QPTP) insertion at position 97 of PreS1 which occurred erroneously during mutagenesis, was also analysed. The mutated coding regions were inserted into the pCITE-2c vector to serve as template for *in vitro* transcription and translation. SDS-PAGE analysis showed that all the mutated proteins were expressed at levels comparable to that of the wild-type (data not shown).

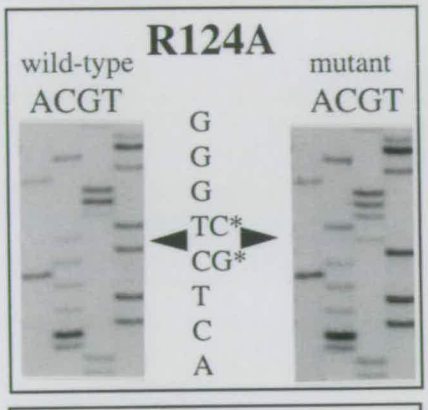
Fig. 20: Substitution of positively charged amino acids with alanine

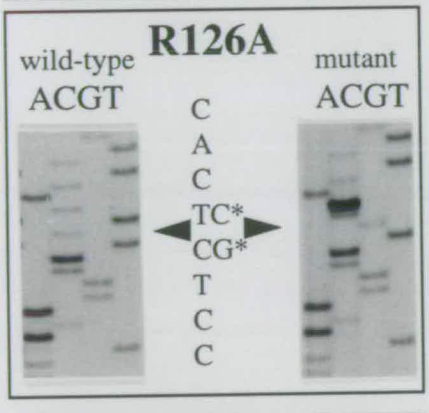
*The non-coding sequences of wild-type and mutants are shown at the left and right, respectively, of each panel. All sequences are read from left to right, ACGT, and from bottom to top, 5' to 3'. Mutation of Arg and Lys codons required substitutions of the first and second nucleotides with G and C; these are indicated by *. The nucleotide sequence of mutant K38A was analysed with an automated sequencer and the sequence from left to right represents the non-coding strand from 5' to 3'.*

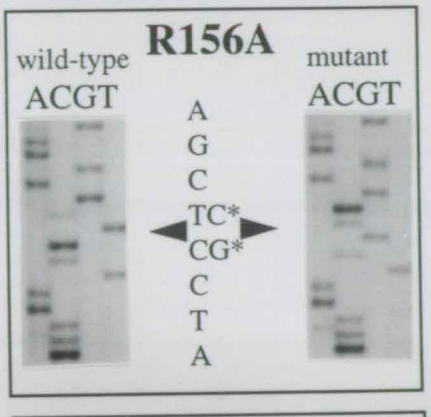

R88A
wild-type
ACGT
mutant
ACGT
T
T
A
GC*
CG*
G
G
T

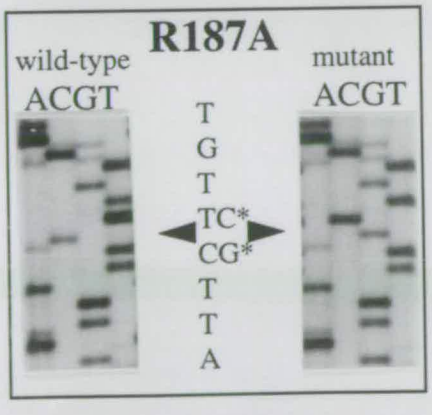

R92A
wild-type
ACGT
mutant
ACGT
C
C
T
GC*
CG*
C
G
T

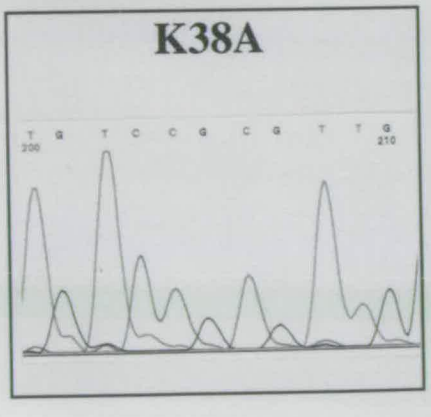

R102A
wild-type
ACGT
mutant
ACGT
G
A
C
TC*
CG*
T
T
G


R124A
wild-type
ACGT
mutant
ACGT
G
G
G
TC*
CG*
T
C
A


R126A
wild-type
ACGT
mutant
ACGT
C
A
C
TC*
CG*
T
C
C


R156A
wild-type
ACGT
mutant
ACGT
A
G
C
TC*
CG*
C
T
A


R187A
wild-type
ACGT
mutant
ACGT
T
G
T
TC*
CG*
T
T
A


K38A
200
210

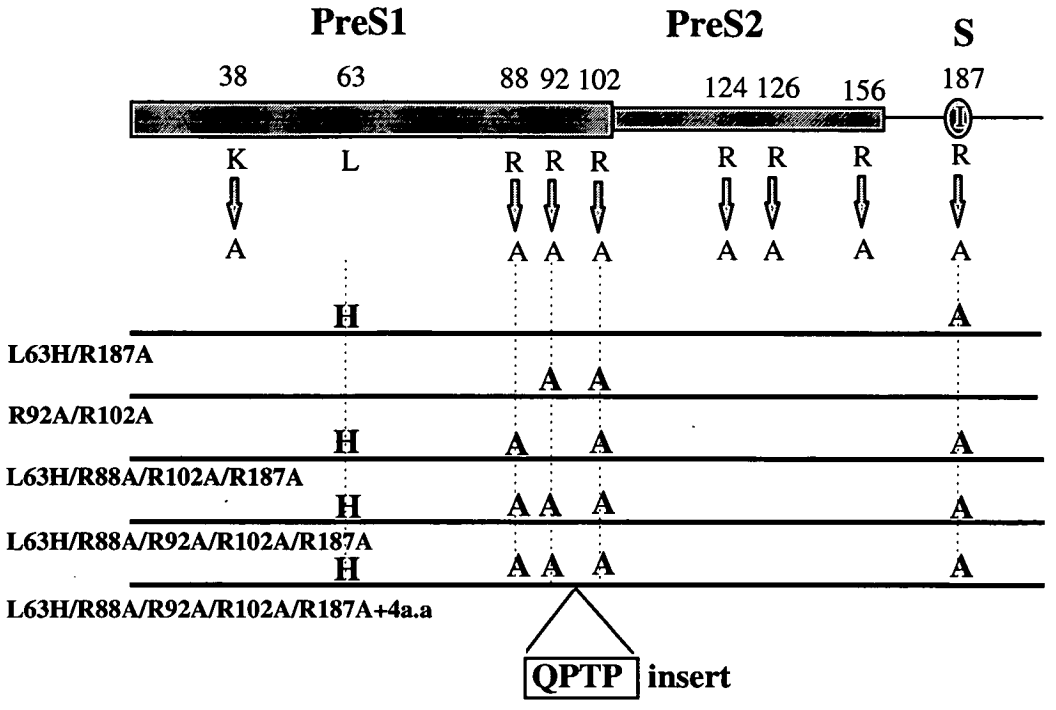


Fig. 21: Positively charged amino acid substitutions and multiple mutations in the PreS regions and helix I of L-HBsAg.

The top diagram represents the PreS regions and a part of the S region of L-HBsAg. The positions of amino acid substitutions are indicated on the top of the diagram. Single amino acid mutations are indicated by arrows. Multiple mutations are shown on the lines beneath the arrows. The abbreviations of the multiple mutations are shown at the left below each mutant.

4.3 Discussion

The structure of HBsAg is not known at present, although models for the topology of the S-, M-, and L-HBsAg in the ER membrane have been proposed based on computer modelling (Stirk *et al.*, 1992; Berting *et al.*, 1995), binding of monoclonal antibodies (Heermann *et al.*, 1984), antigenic determinants (Gavilanes *et al.*, 1982), post-translational modification and protease accessibility (Bruss and Ganem, 1991b; Bruss *et al.*, 1994; Ostapchuk *et al.*, 1994; Bruss and Vieluf, 1995; Prange and Streeck, 1995). However, there has been no chemical or structural analysis of the lipid component of the virion, which also might play a role in the structural organisation of the HBsAg. The only hint suggesting the region in HBsAg

which might directly interact with HBcAg was obtained from affinity selection from fusion phage display libraries with HBcAg (Dyson and Murray, 1995), but the region, LLGRMK, implicated in binding with HBcAg appears to be a discontinuous sequence in which the residues constituting this region are brought together from different positions in the L-HBsAg. Therefore large number of deletion and amino acid substitution mutants were generated and the effect of these mutations on the interaction with HBcAg is discussed in detail in chapter 5.

Rabbit reticulocyte lysates, an eukaryotic cell free system, were chosen to express the mutated HBsAg because attempts in our laboratory to overexpress the full-length L-HBsAg in *E. Coli* have so far failed (M.R. Dyson and S.A. Petukhov personal communication). Although HBsAg has been successfully overproduced in yeast and animal cells (Valenzuela *et al.*, 1982; Murray *et al.*, 1984), the proteins synthesised form particles having similar properties to the 22 nm particles released by virally infected human cells, which would not be suitable for the study of the interaction between HBsAg and HBcAg, particularly in determination of the regions and amino acids forming the contact area with HBcAg. For this study, rabbit reticulocytes offered several advantages: (i) the ability to incorporate amino acids labelled with [^{35}S]-methionine into HBsAg which allows detection of the labelled protein by autoradiography, phosphoimager, or scintillation counting; (ii) it can be used conveniently to verify the size of a mutant synthesised from a modified gene; (iii) reducing agents such as DTT can be added to avoid the formation of disulphide bonds, which is particularly important for HBsAg because it has been shown that S-HBsAg forms dimers and oligomers in the secreted subviral particle (Huovila *et al.*, 1992; Wunderlich *et al.*, 1996). The major limitation of this system is that only a very small amount of product is synthesised in a reaction, estimated to be around 10^{-15} mole per µl of reticulocyte.

CHAPTER 5

CHARACTERISATION OF THE INTERACTION BETWEEN SURFACE AND CORE ANTIGENS

5.1 Introduction

The PreS regions of L-HBsAg contain the hepatocyte receptor binding site (Neurath *et al.*, 1986) and have been found on the external surface of HBV(Heermann *et al.*, 1984; Kuroki *et al.*, 1990; Mimms *et al.*, 1990). They have also been found in the cytosolic side of cellular membrane vesicles (Bruss *et al.*, 1994; Ostapchuck *et al.*, 1994; Bruss and Vieluf, 1995; Prange and Streeck, 1995). These findings suggest two crucial roles for these regions: (i) attachment of the virion to host cells and; (ii) virion assembly. The second function of the PreS regions was further supported by the observation that in a cell free system, nucleocapsids interact directly with L-HBsAg, but not with S-HBsAg (Dyson and Murray, 1995). Furthermore, L-HBsAg produced in the absence of microsomal membrane was still able to interact with nucleocapsid (Dyson and Murray, 1995), indicating that the interaction does not require post-translational processing of the L-HBsAg such as membrane translocation (Eble *et al.*, 1986), acylation (Persing *et al.*, 1987) or glycosylation (Patzer *et al.*, 1986), which occur naturally *in vivo*. However, the exact docking site for HBcAg and specific amino acids involved in the interaction have not yet been identified. The goal of the work described in this chapter was to evaluate the effects of deletion and specific amino acid substitutions of the L-HBsAg (Chapter 4) on the interaction with HBcAg. To further understand the strength of the interaction that ultimately provides insight into the basic principles governing the virion assembly in host cells, an equilibrium binding assay in solution was established which allows the determination of relative dissociation constants.

5.2 Results

5.2.1 Immunoprecipitation with anti-HBcAg rabbit serum

Immunoprecipitation of the N- and C-terminal deletion mutants (Chapter 4) was performed following their interaction with HBcAg in order to define the

minimum contiguous sequence of amino acids required in L-HBsAg. At first translation reactions with the mutants were carried out in the presence and absence of HBcAg, then anti-HBcAg rabbit serum was added to form immune complexes with HBcAg. Staphylococus protein A covalently linked to sepharose beads (Goding, 1978) was used to precipitate the antibody-antigen complexes. The complexes were dissociated by heating in SDS-loading buffer, fractionated by SDS-polyacrylamide gels, and the surface antigen mutants were detected by autoradiography and phosphoimager analysis.

Fig. 22 shows the immunoprecipitation results with the N-terminal deletion mutants. In the presence of HBcAg, a major band of immunoprecipitated product was obtained with full-length L-HBsAg, NΔ16 and NΔ24, but the intensity of the precipitated band significantly decreased with NΔ40, NΔ54 and NΔ60, and only a very small amount of precipitated product was detected with NΔ77, NΔ126 and S-HBsAg after prolonged exposure (second lane of each mutant). These bands were absent from the precipitation of translation reactions that did not contain HBcAg (last lane of each mutant). Phosphoimager analysis further supported the observation that deletion of 40 amino acids from the N-terminus dramatically reduced the interaction, but deletion of 126 residues from the same end did not completely abolish it (Fig. 23).

On the other hand, immunoprecipitation of the C-terminal deletion mutants showed that protein bands of CΔ191, CΔ210, CΔ242, CΔ263, CΔ322, CΔ346 and CΔ371 were brought down by anti-HBcAg rabbit serum in the presence of HBcAg (Fig. 24 second lane of each mutant). However, no immunoprecipitation products were observed for CΔ108, CΔ136 and CΔ163. The negative controls performed in the absence of HBcAg gave small quantities of precipitate with mutants CΔ191 and longer (last lane of each mutant), which was probably caused by non-specific binding of the polyclonal serum to the translation products. However, these background bands were very faint compared to those in the presence of HBcAg.

Fig. 22: Immunoprecipitation of N-terminal deletion mutants with rabbit anti-HBcAg polyclonal serum. →

The mutants are indicated in the box on top of every four lanes. For each mutant, in vitro translation was carried out in the presence (+) or absence (-) of HBcAg (lanes 1 and 3 for each mutant). Immunoprecipitation (+) was then performed with rabbit anti-HBcAg serum (lanes 2 and 4 for each mutant). Numbers to the left of each gel show the positions of molecular weight standards expressed in kDa. Translation and immunoprecipitation products were examined by 15% SDS-PAGE and autoradiography. The period of exposure for (a),(b) and (c) was 12 h, and for (d) was 2 weeks.

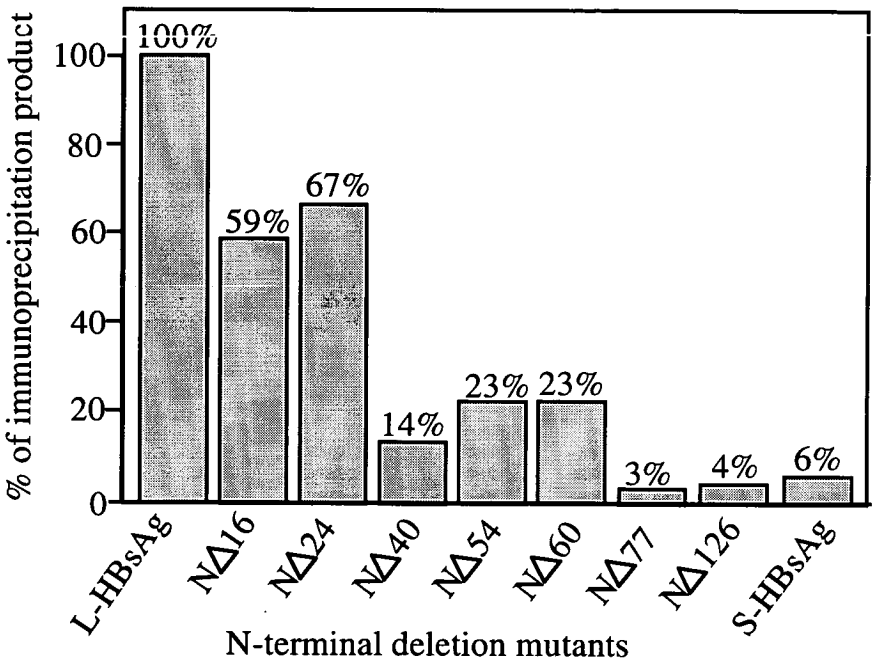
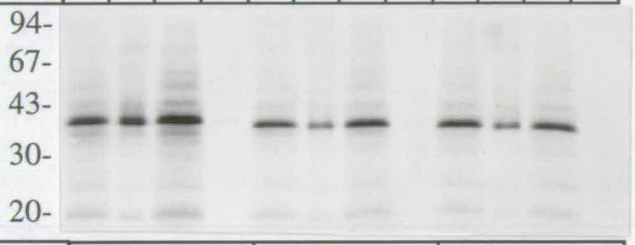


Fig. 23: Phosphoimager quantitation of immunoprecipitation products of N-terminal deletion mutants.

The ratio of the immunoprecipitation product for each construct was calculated with the following formula: (counts of the major immunoprecipitation band in lane 2/counts of the major translation product in lane 1) - (counts of the background band in lane 4/counts of the major translation product in lane 3), and the percentage (%) of immunoprecipitation product was calculated by considering the precipitated product of L-HBsAg as 100%.

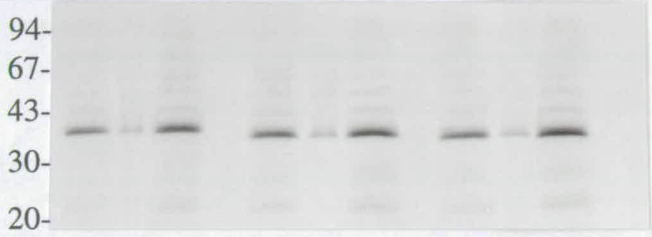
(a)

mutants	wild type				NΔ16				NΔ24			
HBcAg	+	+	-	-	+	+	-	-	+	+	-	-
Immunoprecipitation	-	+	-	+	-	+	-	+	-	+	-	+



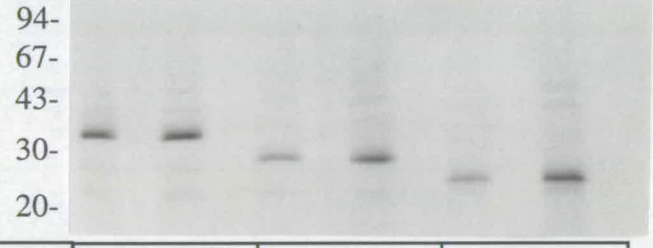
(b)

mutants	NΔ40				NΔ54				NΔ60			
HBcAg	+	+	-	-	+	+	-	-	+	+	-	-
Immunoprecipitation	-	+	-	+	-	+	-	+	-	+	-	+



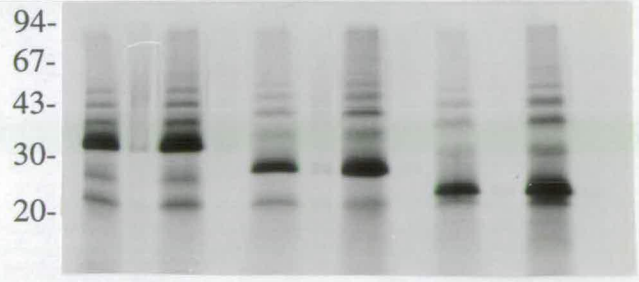
(c)

mutants	NΔ77				NΔ126				S-HBsAg			
HBcAg	+	+	-	-	+	+	-	-	+	+	-	-
Immunoprecipitation	-	+	-	+	-	+	-	+	-	+	-	+



(d)

mutants	NΔ77				NΔ126				S-HBsAg			
HBcAg	+	+	-	-	+	+	-	-	+	+	-	-
Immunoprecipitation	-	+	-	+	-	+	-	+	-	+	-	+



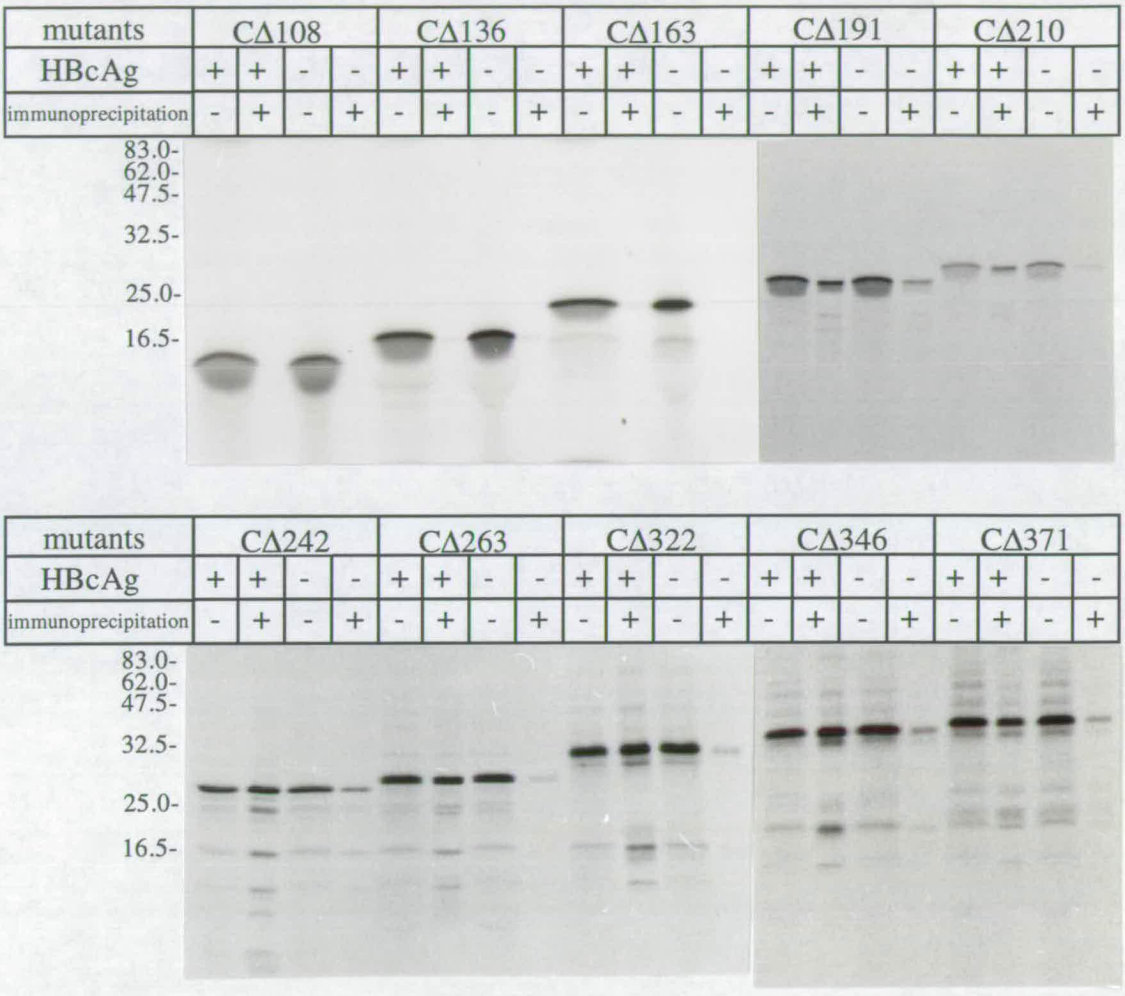


Fig. 24: Immunoprecipitation of C-terminal deletion mutants with rabbit anti-HBcAg polyclonal serum.

The mutants are indicated in the box on top of every four lanes. For each mutant, in vitro translation was carried out in the presence (+) or absence (-) of HBcAg (lanes 1 and 3 of each mutant). Immunoprecipitation (+) was then performed with rabbit anti-HBcAg serum (lanes 2 and 4 of each mutant). Numbers to the left of each gel show the positions of molecular weight standards expressed in kDa. Translation and immunoprecipitation products were examined by 15% SDS-PAGE and autoradiography.

5.2.2 Development of an equilibrium binding assay in solution

The method employed here for measuring relative dissociation constants (K_d^{rel}) between HBsAg and HBcAg in solution was first described for antibody-antigen reactions (Friguet *et al.*, 1985) and later adapted and modified for general use to determine the affinity constants of C4BP-HPS (C4b-binding protein with human protein S; Nelson and Long, 1991) and fusion phage-substrate complexes (Dyson *et al*, 1995) in solution. In the established assay, various concentrations of HBcAg were incubated with a constant concentration of [^{35}S]-HBsAg, that had been synthesised in a cell free system, until equilibrium was reached. The concentration of free [^{35}S]-HBsAg at equilibrium was quantitated by incubating an aliquot of the reaction mixture in HBcAg-coated wells for a short time period; wells were then washed and bound [^{35}S]-HBsAg was measured by scintillation counting of radioactivity.

The method is only valid if the following four criteria are fullfilled: (i) the temperature is kept constant throughout the experiment; (ii) the binding equilibrium is reached; (iii) the equilibrium in the solution is undisturbed during the subsequent incubation of the mixture in the HBcAg-coated well and; (iv) the concentration of the free [^{35}S]-HBsAg in solution is proportional to the radioactivity measured by a scintillation counter.

A time course study of the [^{35}S] L-HBsAg binding to the HBcAg-coated wells showed that equilibrium was reached after 6 h at room temperature or after 24 h at 4 °C (Fig. 25). This also shows that the binding rate of these molecules at room temperature is greater than that at 4 °C, which can be determined from the initial slope of the curves. After 24 h incubation, the counts at 4 °C were two times higher than those obtained at room temperature, therefore the temperature was kept constant at 4 °C throughout the assay.

Fig. 26 displays a linear relationship between the concentration of [^{35}S] L-HBsAg and the cpm measured in scintillation counting. Various dilutions of translation reaction mixtures containing the [^{35}S] L-HBsAg were incubated in wells coated with HBcAg for 1 h. After washing, the wells were placed into scintillation vials for counting. The cpm was directly proportional to the concentration of the

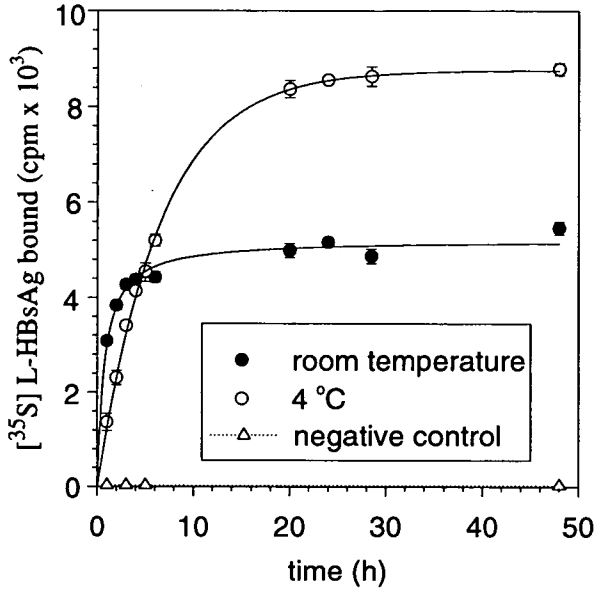


Fig. 25: Time course experiment of [^{35}S] L-HBsAg binding to HBcAg-coated wells.

A constant concentration of [^{35}S] L-HBsAg (200x dilution of in vitro translation reaction) was incubated with HBcAg-coated wells (10 µg/ml) for different time periods at room temperature and at 4 °C. cpm was obtained from scintillation counting and represents the arithmetic mean ± standard deviation of triplicate. Negative control was carried out with the same conditions but in the absence of transcription products.

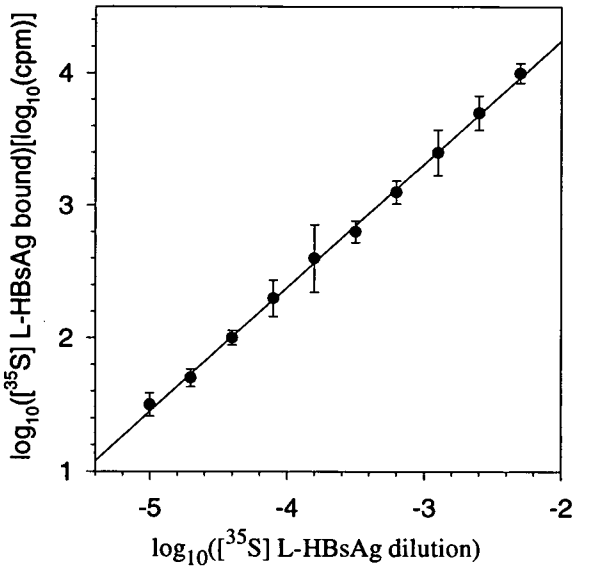


Fig. 26: Linear relationship between cpm and concentration of [^{35}S] HBsAg.

Various dilutions of in vitro translation mixtures were incubated for 1 h at 4 °C with HBcAg-coated wells (10 µg/ml). The wells were washed and radioactivity (cpm) was counted with a scintillation counter. Assays were performed in triplicate.

[^{35}S] L-HBsAg over a four log range. This linear relationship allows the determination of the free [^{35}S]-HBsAg concentration at equilibrium by scintillation counting.

The amount of the free HBsAg molecules bound to the HBcAg-coated on the solid phase must be very low in order to prevent any perturbation of the equilibrium in solution. To show that this requirement was fulfilled, varying dilutions of *in vitro* translation mixtures were incubated for 1 h in wells previously coated with HBcAg and the contents of each well were then transferred to a second coated well and incubated under identical conditions. The bound [^{35}S] L-HBsAg in the first and second sets of wells were then determined. Initially, 10 µg/ml of HBcAg was used to coat the wells, however it was found that the amount of the free [^{35}S] L-HBsAg bound to the wells was greater than 10%, which was deduced from the slopes of the straight lines obtained with the two sets of wells (Data not shown). Attempts were then taken to reduce the HBcAg concentration immobilised on the solid phase to 7.5, 5.0, and 2.5 µg/ml. Fig. 27 shows that with 2.5 µg/ml of HBcAg used for coating the wells, the slopes differ by 9.6%. At this level, it is unlikely that the equilibrium in solution will be significantly disturbed.

5.2.3 Mathematical and graphical analysis of binding data

Table 5 representatively shows the experimental data collected for the interaction between L-HBsAg and one of its derivatives, NΔ77, with HBcAg. The data were then analysed with the Scatchard plot and with non-linear hyperbolic curve fitting to obtain the K_d^{rel} values.

- ***i The Scatchard plot***

At equilibrium, the reaction between HBsAg and HBcAg in solution can be represented as follows:

$$S + C \leftrightarrow SC \qquad (1)$$

Here S is the free HBsAg, C represents the free HBcAg, and SC refers to the HBsAg-HBcAg complex. The interaction is presumed to be bimolecular and

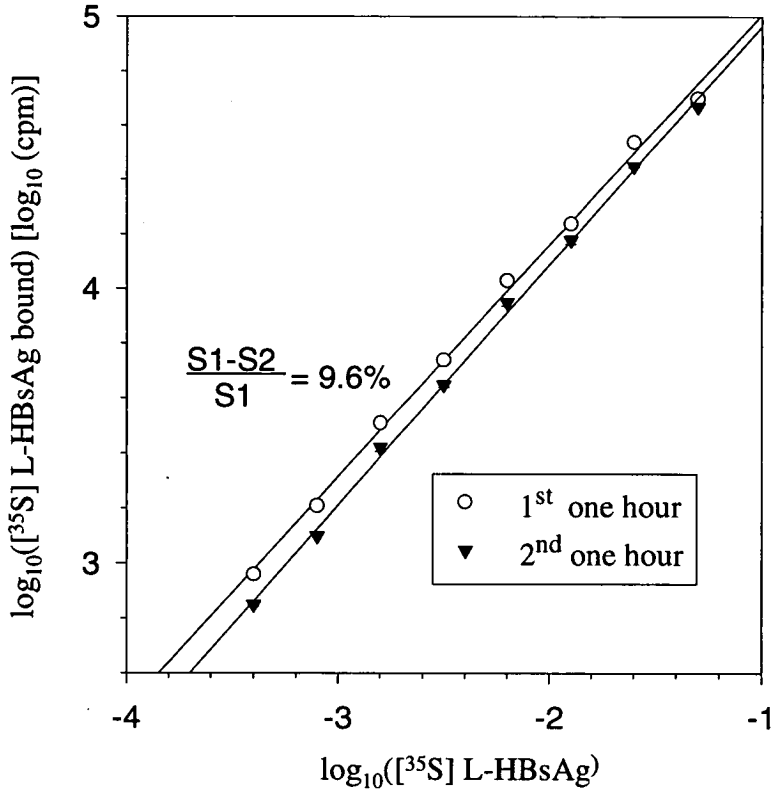


Fig. 27: Determination of the amount of [^{35}S] L-HBsAg captured on HBcAg coated wells with scintillation counting

Various dilution of in vitro translation mixtures were incubated for 1 hr at 4 °C with HBcAg coated wells (2.5 μg/ml) and the solution in each well were then transferred to another coated well and incubated under the same conditions. The wells were washed and radioactivity was counted with scintillation counter. The counts of the first and second sets of the wells are the arithmetic mean ± standard deviation of tripicate. The difference between the slopes of the straight lines was deduced from the ratio $(S_1-S_2)/S_1$, in which S_1 and S_2 are the slopes obtained with the first and second sets of the wells respectively.

(a)

[C] (μM)	cpm	Y=cpm_0-cpm/cpm_0	Y/[C] (μM^{-1})
80	323 ± 21	0.896 ± 0.007	0.011 ± 0
40	398 ± 16	0.872 ± 0.005	0.022 ± 0
20	547 ± 7	0.824 ± 0.002	0.041 ± 0
10	808 ± 27	0.742 ± 0.006	0.074 ± 0.001
5	1230 ± 3	0.604 ± 0.001	0.121 ± 0
2.5	1655 ± 30	0.467 ± 0.009	0.187 ± 0.004
1.25	2290 ± 79	0.263 ± 0.025	0.207 ± 0.023
0	3105 ± 122 (cpm_0)	---	---

(b)

[C] (μM)	cpm	Y=cpm_0-cpm/cpm_0	Y/[C] (μM^{-1})
5	345 ± 34	0.948 ± 0.005	0.190 ± 0.001
2.5	573 ± 30	0.914 ± 0.005	0.366 ± 0.002
1.25	714 ± 68	0.893 ± 0.010	0.715 ± 0.008
0.625	1086 ± 38	0.837 ± 0.006	1.339 ± 0.010
0.313	1190 ± 110	0.821 ± 0.017	2.624 ± 0.054
0.156	1272 ± 56	0.809 ± 0.008	5.188 ± 0.052
0.078	1544 ± 108	0.768 ± 0.016	9.850 ± 0.210
0.039	1718 ± 225	0.742 ± 0.033	19.026 ± 0.857
0.02	2077 ± 56	0.689 ± 0.008	34.433 ± 0.417
0	6670 ± 298 (cpm_0)	---	---

Table 5: Data obtained from binding assays of HBcAg with mutant NΔ77 (a) and wild-type L-HBsAg (b).

The experiments were carried out at 4 °C with various concentration of HBcAg, [C]. Counts per minute (cpm) was obtained from scintillation counting for 5 min and represents the arithmetic mean ± standard deviation of triplicate. cpm obtained for the sample containing no HBcAg was designated cpm_0. Fraction bound, Y, was calculated from cpm_0-cpm/cpm_0.

reversible, and the system is defined to have one identical and non-interacting set of binding sites, in which the probability of binding to any one site is the same as the probability of binding to any other site and occupancy of one site does not affect the probability of binding to any other. Phenomenologically, the dissociation constant, K_d, of this equilibrium can be defined as

$$K_d = [S][C]/[SC] \quad (2)$$

The fraction, Y, of the HBsAg bound to the HBcAg is given by

$$Y = [SC]/([SC] + [S]) \quad (3)$$

By combining equations (2) and (3), the Scatchard equation can be derived to give

$$Y/[C] = (1- Y)/K_d \quad (4)$$

As shown previously (Fig. 26), the free [^{35}S]-HBsAg was proportional to the quantity determined by scintillation counting, therefore the ratio between the free (S) and total (S_t) HBsAg concentrations can be written as follows:

$$[S] = [S_t] (cpm/cpm_o) \quad (5)$$

where cpm_o refers to counts per minute obtained for the well containing no HBcAg, and cpm is the counts per minute read for wells with different concentrations of HBcAg. According to mass conservation equation, S_t, S and SC are related to each other

$$[SC] = [S_t] - [S] \quad (6)$$

Combination of this relationship with equation (5) gives

$$[SC] = [S_t] (cpm_o-cpm)/cpm_o \quad (7)$$

, and substitution into equation (4) and appropriate rearrangement gives

$$(cpm_o-cpm)/cpm_o[C] = K_d^{-1} - K_d^{-1} (cpm_o-cpm)/cpm_o \quad (8)$$

When the starting concentration of HBcAg, $[C_t]$, is in large excess over the $[S_t]$ (the lowest concentration of HBcAg used in this study was 500-fold in excess compared to that of HBsAg). Thus, [SC] becomes negligible compared to $[C_t]$ and [C] can be approximated well by $[C_t]$. If the assumptions underlying development of equation 8 are true, then, a plot of $(cpm_o-cpm)/cpm_o[C_t]$ against $(cpm_o-cpm)/cpm_o$ (or $Y/[C_t]$ against Y) gives a straight line of slope equal to $-1/K_d$ or $-K_a$.

Fig. 28 shows these plots for L-HBsAg and NΔ77 with the data from Table 5. The plot for NΔ77 was linear with a K_d^{rel} value of 3.0 ± 0.3 μM (Fig. 28a), but the

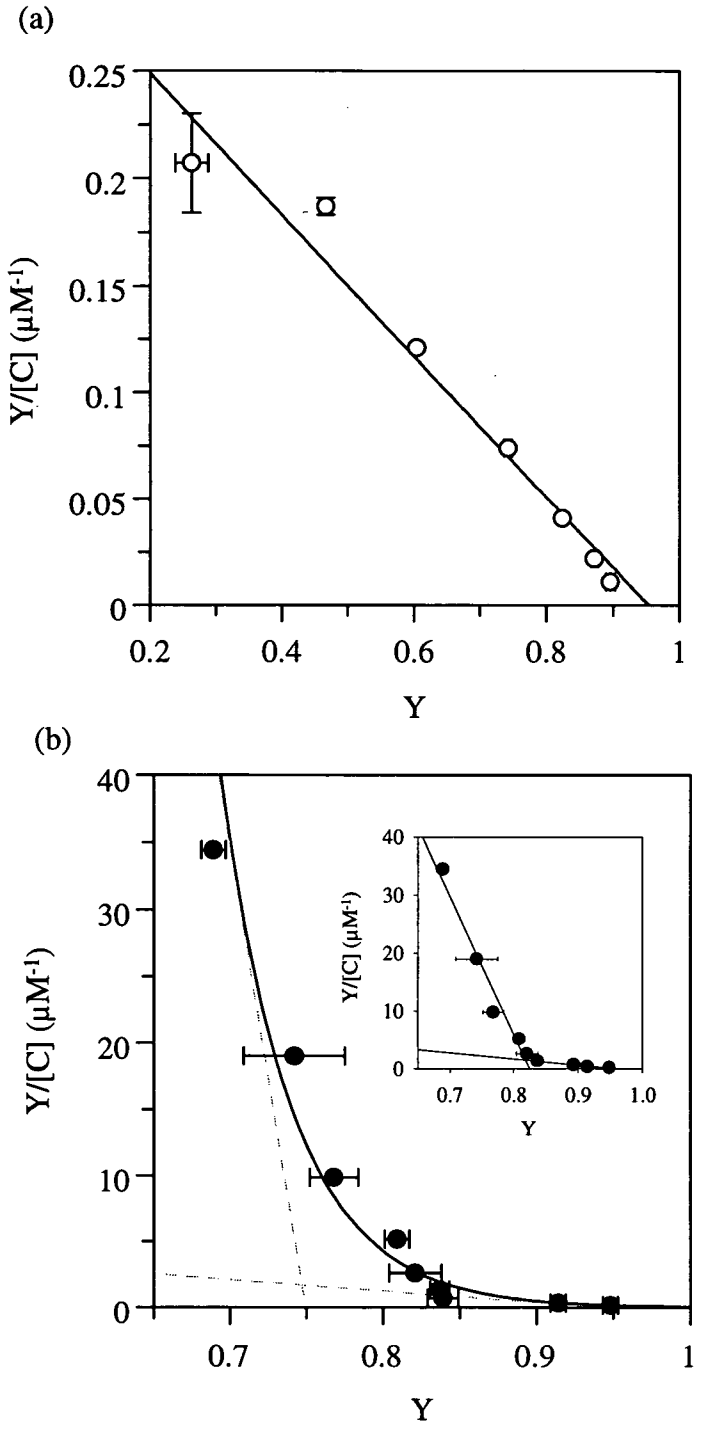


Fig. 28: Linear and concave Scatchard plots

(a) Scatchard graph of mutant NΔ77 binding to HBcAg. The data points were fitted with a straight line using the GraFit™ software. The equation y = mx + c was used to calculate the slope (m) of the line. (b) Scatchard plot of L-HBsAg binding to HBcAg. The single exponential decay function provided in the software was used to fit the points. The straight lines were fitted by eye to the apparently linear extreme parts of the curve (broken lines) and through all the data points (inset) to illustrate the incorrect procedures to obtain K_d values.

plot for L-HBsAg was with different limiting slopes at low and high concentrations of HBcAg (Fig. 28b). Initially, when this plot was encountered, there was a strong, naive temptation to put two lines through the points (Fig. 28b, inset) or the apparently linear end parts (limbs) of the curve (Fig. 28b, broken lines) by the least square method under some subjective influences, and then ascribe one of the lines, that of deeper slope, to a high affinity site and the other to a low affinity site. In fact, the slopes of these lines are complicated functions of site parameters (Klotz, 1971) and generally unsuitable for evaluation of binding constants. To avoid misinterpretation of non-linear Scatchard plots, many authors have suggested the use of non-linear fitting of appropriate binding equations as a more reliable method for data analysis (Reich *et al.*, 1972; Hancock *et al*, 1979; Munson and Rodbard, 1980; Hulme and Birdsall, 1992).

- ***ii non-linear fitting of hyperbolic equations***

For a simple interaction involving only one class of identical and non-interacting binding sites, equation 4 can be rearranged to give a hyperbola-like equation:

$$Y=[C]/[C]+K_d \qquad (9)$$

The hyperbolic function starts at 0 when [C]= 0 (all HBsAg remain free in solution; $cpm=cpm_o$) and rises to maximal binding, $B_{max}=1$ (cpm=0), as [C] increases. When half of the total HBsAg molecules are engaged in a HBsAg-HBcAg complex (i.e. Y=1/2 or $cpm=1/2\ cpm_o$), $K_d = [C]$.

The hyperbolic relationship between Y and [C] for the interaction between NΔ77 with HBcAg is shown in Fig. 29a. The hyperbolic curve was fitted to the data points using the equation for one binding site of the GraFit™ software. The K_d^{rel} calculated by the software for the interaction was 2.7 ± 0.2 μM, which is in good agreement with that determined by the Scatchard plot (Fig. 28a; K_d^{rel} = 3.0 ± 0.3 μM), and the difference is within the range of standard error.

For a system with two classes of independent binding sites, with respective dissociation constants K_{d1} and K_{d2} for high- and low-affinity sites, the equivalent of equation 9 is as follows:

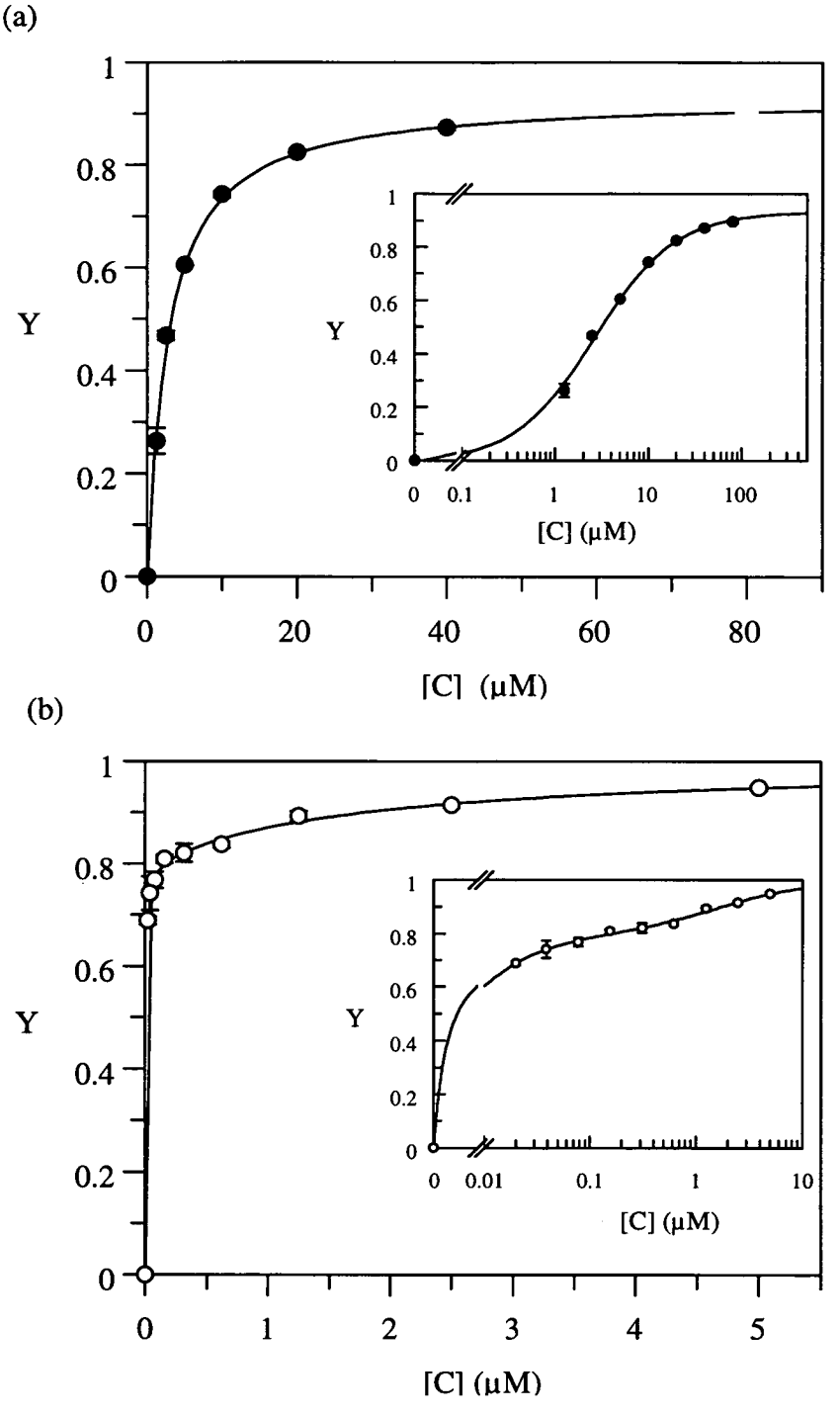


Fig. 29: Curve-fitting binding data with hyperbolic functions

(a) Graph of Y versus [C] for the binding of NΔ77 to HBcAg using the same experimental data as in Fig. 28a. The curve was fitted to the points using a hyperbolic function for one binding site (GraFit™). Inset: [C] was plotted as a function of log_{10} [C].
(b) The binding data for the interaction of L-HBsAg with HBcAg (Table 5b) was plotted with linear and semi-log scales (inset) for the [C] axis. The points were fitted with the equation for two binding sites with the same software.

$$Y=\{[C]/[C]+K_{d1}\} + \{[C]/[C]+K_{d2}\} \text{ (Cornish-Bowder, 1995)} \quad (10)$$

The data of L-HBsAg binding to HBcAg (Table 5a) which yield a biphasic Scatchard plot (Fig. 28b) were fitted with the equation for two binding sites (GraFit™) and the plot is shown in Fig. 29b. The relative dissociation constants for the high- and low-affinity binding sites were 3.2 ± 0.5 nM (K_{d1}) and 1.6 ± 0.7 μM (K_{d2}), respectively.

5.2.4. Interaction of large, small and unfractionated HBcAg particles with L-HBsAg

The binding of L-HBsAg to unfractionated HBcAg particles exhibiting two widely differing dissociation constants, suggests the existence of two classes of binding sites between the molecules. Since two populations of HBcAg particles, large and small, are present in the system, it may have been thought that each species displayed different binding sites. To investigate this further, the binding experiment was carried out with fractionated large and small HBcAg particles. The Scatchard plots for the binding of L-HBsAg to large and small HBcAg particles (Fig. 30a and b) showed the same biphasic pattern to that of unfractionated HBcAg particles (Fig. 30c). The K_d^{rel} values determined by curve fitting the data points to a double hyperbolic function (Fig. 30d) are summarised in Table 6.

HBcAg particles	K_{d1}^{rel} (nM)	K_{d2}^{rel} (μM)
unfractionated	4.7 ± 0.9	1.4 ± 0.9
large	3.6 ± 0.9	1.9 ± 1.0
small	5.0 ± 1.4	0.4 ± 0.2

Table 6: Relative dissociation constants of L-HBsAg binding to unfractionated, large, and small HBcAg particles

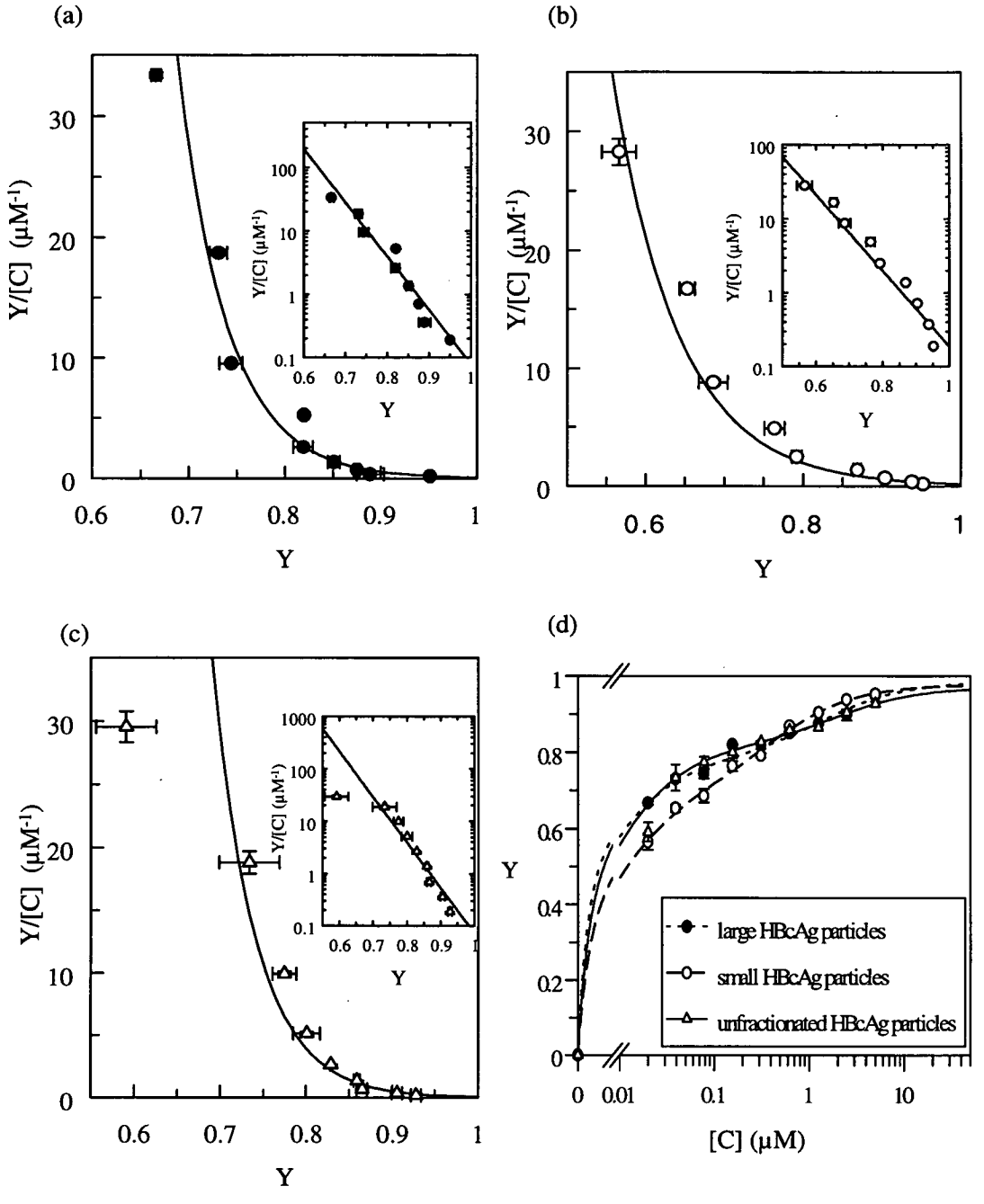


Fig. 30: Graphical presentation of the binding data collected for the interaction of large, small and unfractionated HBcAg particles with L-HBsAg

Scatchard plots for the binding of L-HBsAg to large (a), small (b) and unfractionated (c) HBcAg particles. The curves were fitted to the points using the single exponential decay equation of GraFit™ software. The Y axes of the graphs were also plotted as log_{10} as shown in insets. The three sets of binding data were combined in a semi-logarithmic graph (d) and the points were curve fitted with the equation for two binding sites.

5.2.5 Binding of fusion phage to large, small and unfractionated HBcAg particles

Dyson and colleagues (1995) showed that the interaction of fusion phages with unfractionated HBcAg particles gave rise to linear Scatchard plots, indicating an identical set of non-interacting binding sites present in the system. To find out whether the separated large and small HBcAg particles have the same docking sites for fusion phages, phage particles carrying peptide RSLLGRMK (designated 2A-8; residues R and S preceding the fusion hexapeptide, LLGRMK, were mutated from G and A of the gpIII protein; Dyson M.R., unpublished data) were used to interact with the separated core particles. Scatchard analysis of the binding data yielded linear plots (Fig. 31a) and many of the data points at high concentration of HBcAg were squeezed into a small region of the figure near the intercept on the Y axis. In order to present the experimental data uniformly, the fraction of free phage remaining in solution was plotted against the $\log_{10}$ [C], and the curves were fitted to the points using the hyperbolic function for one binding site (Fig. 31b). A comparison between K_d^{rel} values calculated by the two methods is shown in Table 7.

HBcAg particles	$K_d^{rel\,(a)}$ (nM)	$K_d^{rel\,(b)}$ (nM)
unfractionated	0.11 ± 0.01	0.10 ± 0.01
large	0.11 ± 0.01	0.11 ± 0.02
small	0.07 ± 0.01	0.06 ± 0.01

Table 7: Relative dissociation constants of phage 2A8 binding to unfractionated, large and small HBcAg particles

[a] Relative dissociation constants calculated by linear regression of the Scatchard plots(Fig. 31a). [b] Relative dissociation constants calculated by curve fitting to a hyperbolic function (Fig. 31b).

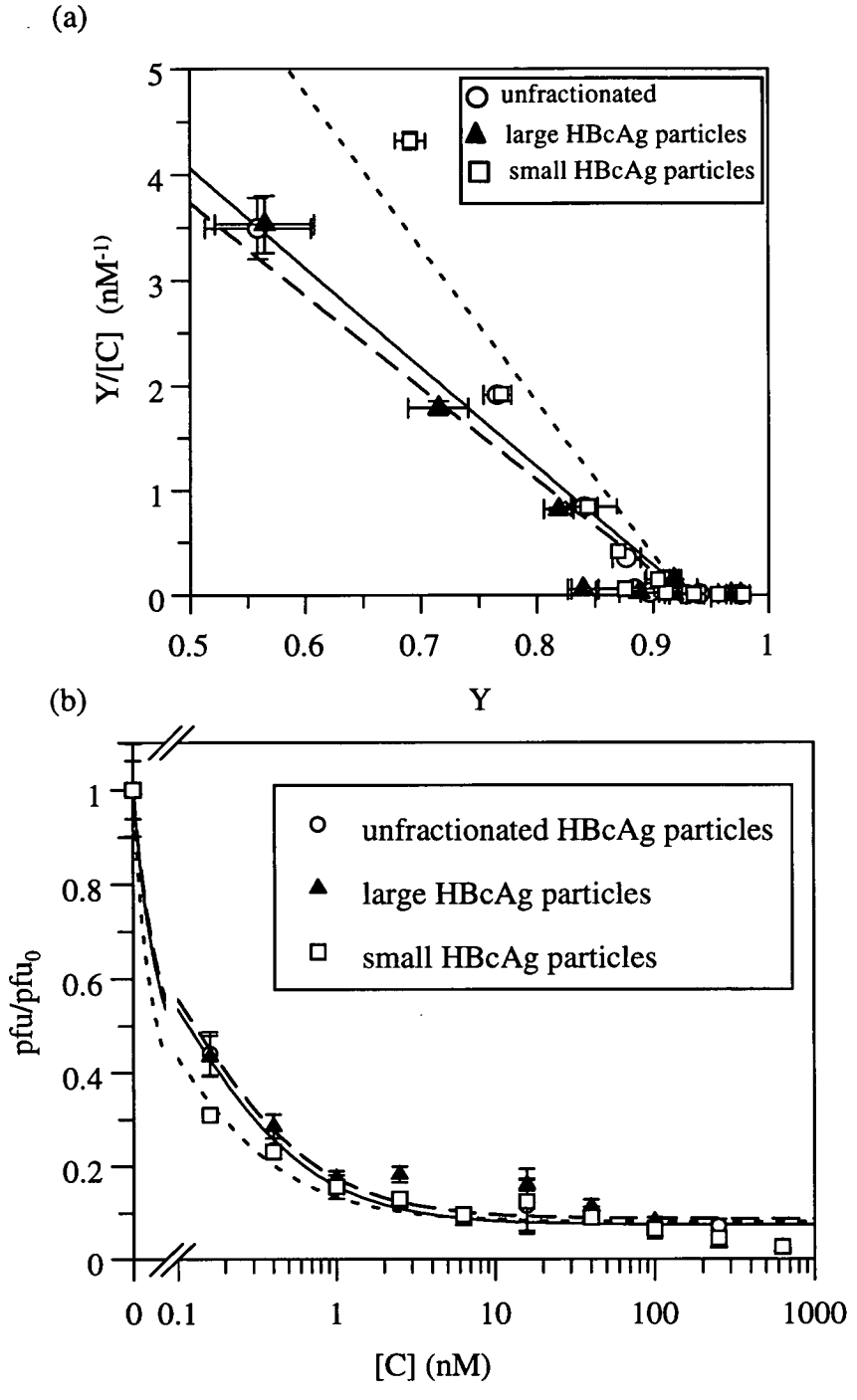


Fig. 31: Graphical presentation of the binding of phage 2A-8 to large, small and unfractionated HBcAg particles.

Data points at high concentration of HBcAg are extremely compressed in the Scatchard plots (a), but they are uniformly distributed in the semi-logarithmic graph (b). Data points are mean ± standard deviation of triplicate determinations.

5.2.6 Interaction of N- and C-terminal deletion mutants with unfractionated HBcAg particles

Following expression in rabbit reticulocyte lysates supplemented with [^{35}S]-methionine, the mutated proteins were subjected to the equilibrium binding assay in order to evaluate the effects of the mutations on the binding to HBcAg. The collected binding data were analysed with the Scatchard plot and non-linear curve fitting, and the K_d^{rel} values derived are given in Table 8. The data points were fitted to the curves shown using equations for one binding site (Fig. 32a and b) or two binding sites (Fig. 33). The results show that mutants NΔ24, CΔ263, CΔ322, CΔ346 and CΔ371 bind to HBcAg with high- and low-affinities, whereas, only one affinity was observed for the binding of mutants NΔ40, NΔ54, NΔ60, NΔ77, NΔ126, CΔ191, CΔ210, and CΔ242 to HBcAg. Those mutants showing two affinities gave rise to biphasic Scatchard plots as was shown by full-length L-HBsAg (Fig. 34). Linear Scatchard plots were obtained for those mutants showing one affinity (Fig. 35).

5.2.7 Interaction of amino acid substitution mutants with unfractionated HBcAg

The effects of single and multiple mutations in the L-HBsAg on the binding to HBcAg were also evaluated with the equilibrium binding assay. Table 9 shows the K_d^{rel} values obtained with the non-linear curve fitting method . All of the single amino acid substitution mutants displayed high- and low-affinities with K_{d1}^{rel} values ranging from 2 to 34 nM, while K_{d2}^{rel} values appeared to be more than 500 nM. Replacement of both Arg 92 and 102 together by Ala gave rise to one affinity (K_d^{rel} = 51 ± 8 nM). Multiple mutation with Leu 63 to His and Arg 88, 92, 102 and 187 to Ala caused more than an 80-fold reduction in the affinity. Furthermore, only one affinity was detected for these mutants.

mutant	two binding sites[a]		one binding site		weak[d] affinity
	K_{d1}^{rel} (nM)	K_{d2}^{rel} (μM)	K_d^{rel} (μM)[b]	K_d^{rel} (μM)[c]	(μM)
wild-type	3 ± 1	1.1± 0.5			
NΔ24	2 ± 0.4	4.5 ± 2.5			
NΔ40			2.2 ± 0.4	2.5 ± 0.5	
NΔ54			1.5 ± 0.1	1.5 ± 0.2	
NΔ60			2.2 ± 0.3	2.1 ± 0.1	
NΔ77			2.7 ± 0.2	3.0 ± 0.3	
NΔ126			6.0 ± 0.8	5.4 ± 0.6	
S-HBsAg					> 80
CΔ108					> 80
CΔ136					> 15
CΔ163					> 15
CΔ191			4.4 ± 0.4	4.4 ± 0.3	
CΔ210			11.0 ± 0.3	10.9 ± 0.1	
CΔ242			1.6 ± 0.2	1.5 ± 0.1	
CΔ263	31 ± 8	10.9 ± 3.0			
CΔ322	3 ± 1	1.2 ± 0.3			
CΔ346	3 ± 2	1.2 ± 0.7			
CΔ371	3 ± 2	1.3 ± 0.3			

Table 8: Relative dissociation constants of N- and C-terminal deletion mutants for HBcAg

The values were derived from the equilibrium binding assay in solution at 4°C and represent mean ± standard error. The K_d^{rel} values for wild type are the mean of four determinants performed in triplicate.

[a] For mutants displaying two affinities, K_{d1}^{rel} and K_{d2}^{rel} values were determined by curve-fitting the binding data with the equation for two binding sites.

For mutants showing one affinity, two methods were used to calculate the K_d^{rel} values:

[b] curve-fitting the data points with one binding site equation (Fig. 32)and,

[c] linear regression analysis of the data points of the Scatchard plots (Fig. 35).

[d] the K_d^{rel} values for weak affinity mutants (> 15 μM) were estimated from the concentration of HBcAg required to inhibit half of the total concentration of HBsAg.

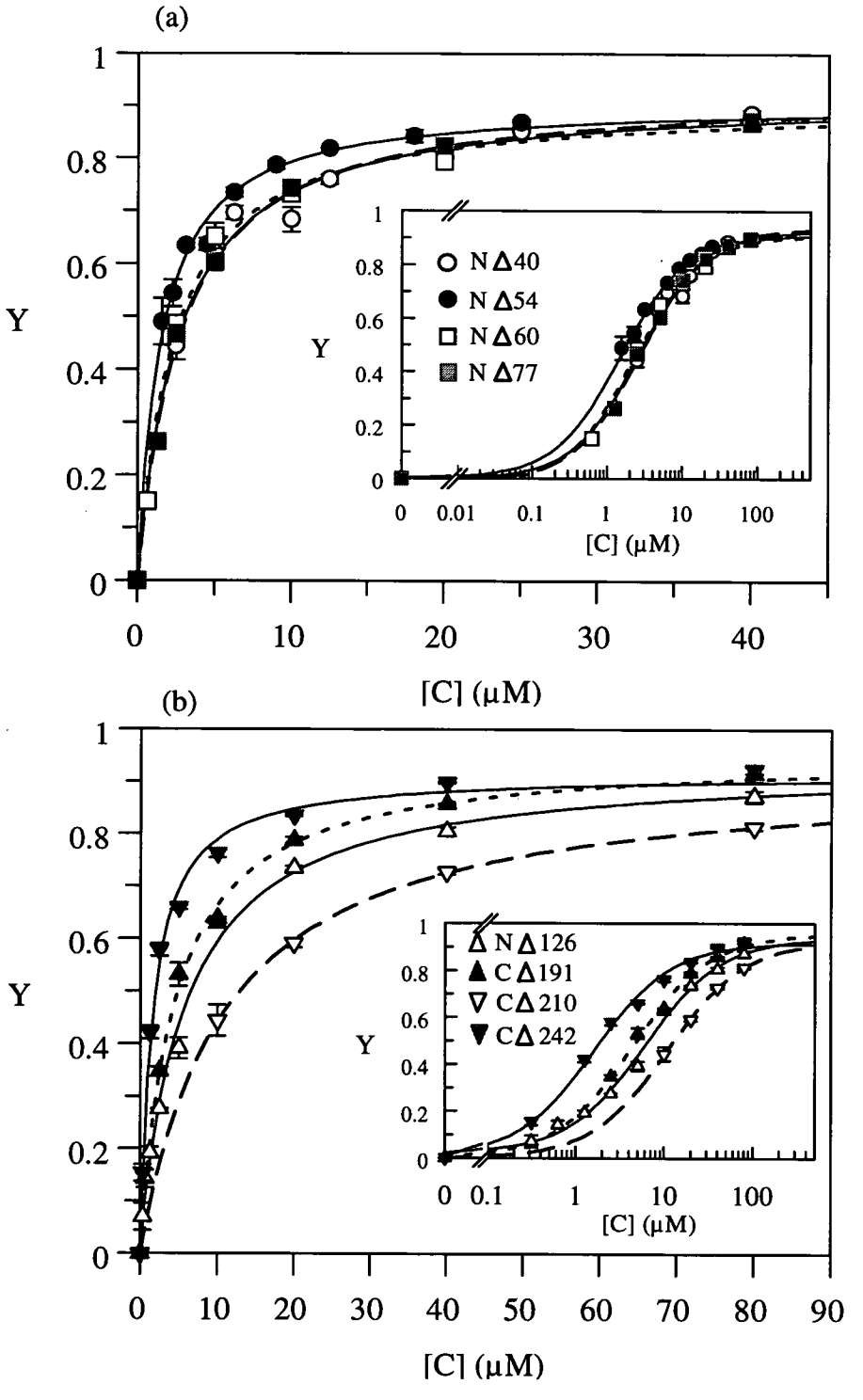


Fig. 32: N- and C-terminal deletion mutants showing one affinity

The binding data collected (mean ± standard deviation of triplicate) for the mutants as indicated on the left of the inserts were plotted as linear and semi-log scales (inserts). The curves were fitted to the points with the hyperbolic function for one binding site equation and the K_d^{rel} values generated are given in Table 8.

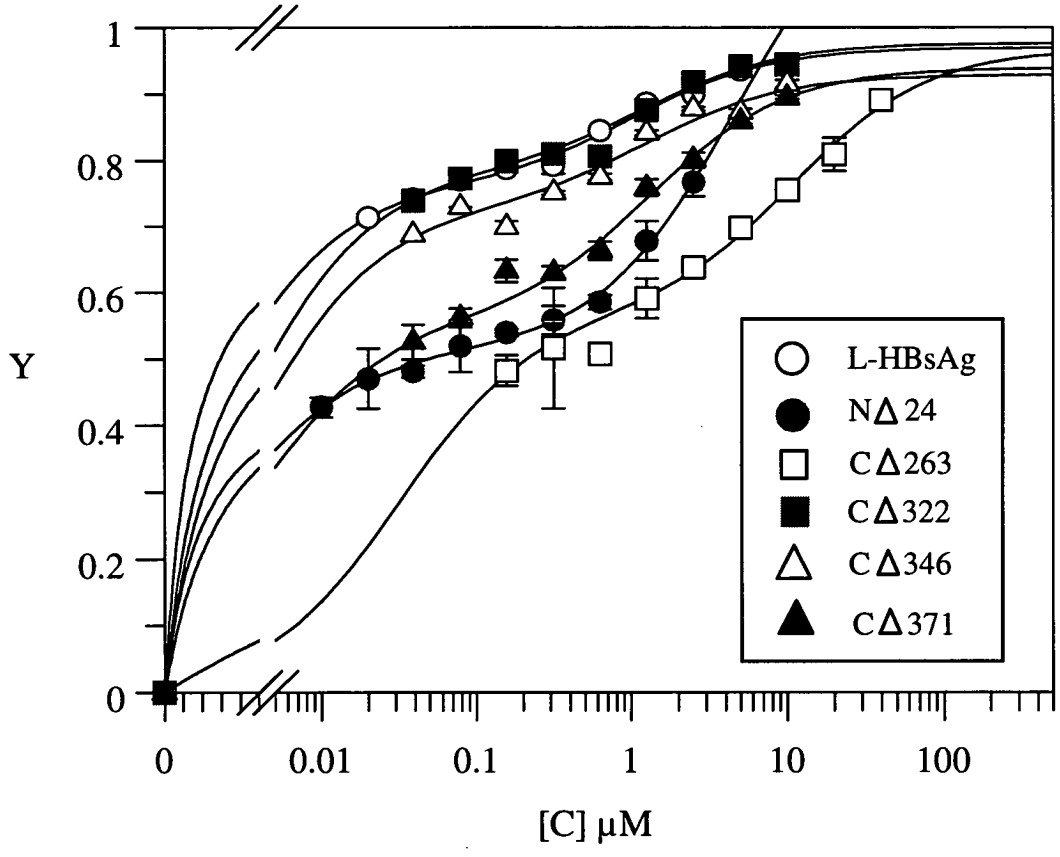
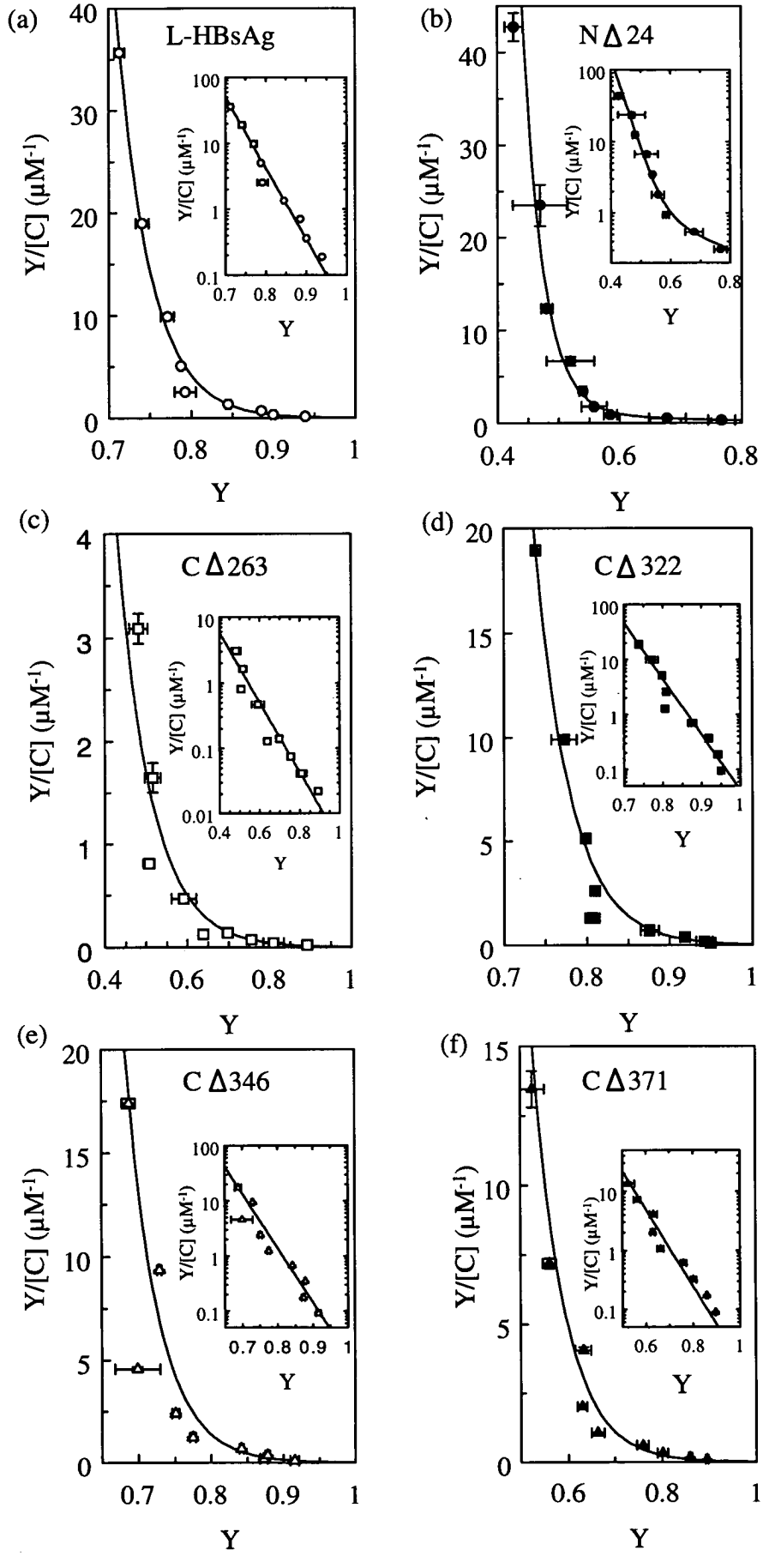


Fig. 33: Semi-logarithmic representation of the binding of N- and C- terminal deletion mutants to HBcAg with two different affinities

The double hyperbolic curves were fitted to the binding data with the equation for two binding sites. K_{d1}^{rel} and K_{d2}^{rel} values derived from this operation are presented in Table 8. Data points are mean ± standard deviation of triplicate determinations.

Fig. 34: Biphasic Scatchard plots for N- and C-terminal deletion mutants showing two different affinities.

The mutants are indicated in the upper part of each graph. All curves were fitted to the points with the single exponential decay function (Grafit™) except mutant NΔ24 which was fitted to the double exponential decay function. Inserts show the Y axis of individual graphs plotted as a function of log_{10}.


(a)
L-HBsAg
Y/[C] (μM-1)
Y
(b)
NΔ24
(c)
CΔ263
(d)
CΔ322
(e)
CΔ346
(f)
CΔ371

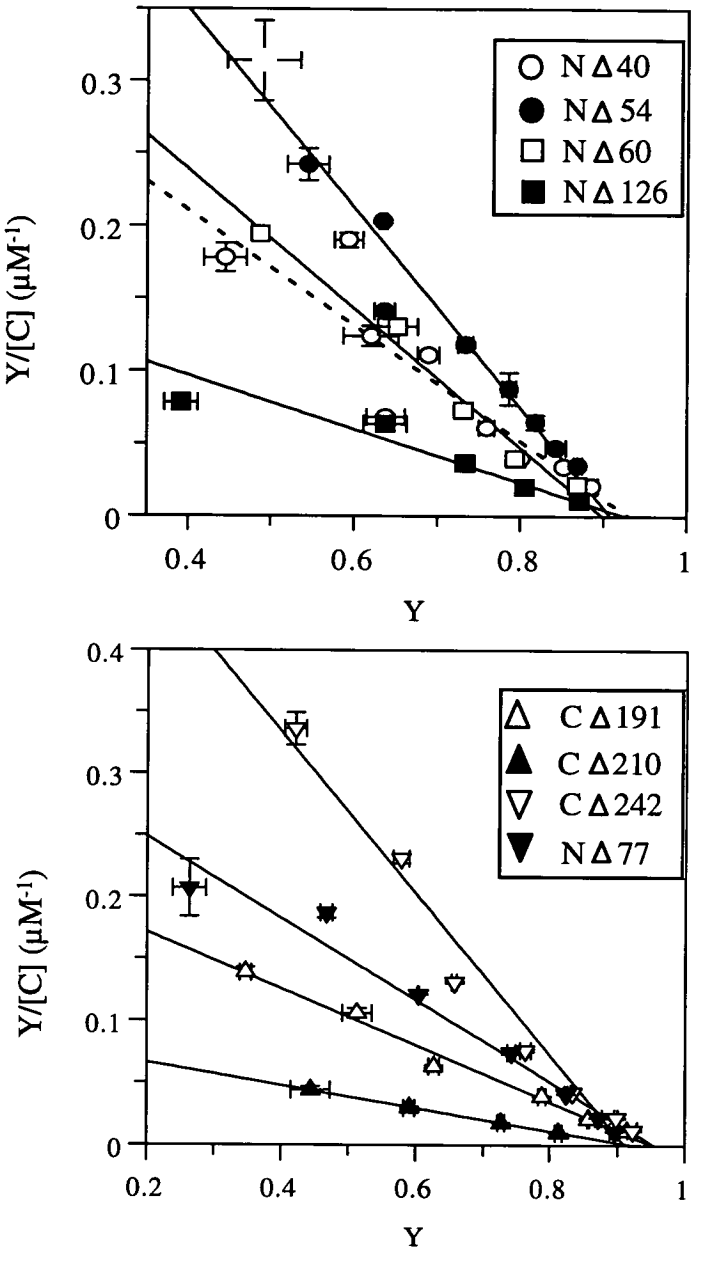


Fig. 35: linear Scatchard graphs for N- and C-terminal deletion mutants

The experimental data shown in Fig. 32 were converted into Scatchard plots. The symbols representing the mutants are indicated at the upper right of the graphs. Straight lines were put through the points using linear regression of GraFit™. The K_d^{rel} values derived from the slopes of the lines are shown in Table 8.

amino acid substitution	two binding sites		one binding site
	K_{d1}^{rel} (nM)	K_{d2}^{rel} (μM)	K_d^{rel} (nM)
wild type	3 ± 1	1.0 ± 0.6	
K38A	8 ± 5	1.1 ± 0.7	
L63H	11 ± 4	2.4 ± 0.9	
L64E	12 ± 7	1.8 ± 1.2	
G65R	2 ± 0.4	0.7 ± 0.1	
R88A	4 ± 1	0.9 ± 0.4	
R92A	34 ± 3	1.0 ± 0.5	
R102A	12 ± 3	0.6 ± 0.3	
R124A	14 ± 2	1.6 ± 1.0	
R126A	12 ± 3	1.0 ± 0.9	
R156A	15 ± 1	3.2 ± 2.0	
R187A	4 ± 1	0.7 ± 0.3	
L63A/G65E	23 ± 5	1.2 ± 0.8	
L63H/R187A	18 ± 7	2.0 ± 0.9	
R92A/R102A			51 ± 8
L63H/R88A/R102A/R187A			362 ± 96
L63H/R88A/R92A/R102A/R187A			288 ± 54
L63H/R88A/R92A/R102A/R187A+4a.a			234 ± 23

Table 9: Binding affinities of amino acid substitution mutants for HBcAg

K_{d1}^{rel} and K_{d2}^{rel} values were obtained by curve fitting the binding data with the equation for two binding sites, whereas K_d^{rel} values for mutants showing one affinity were determined with the equation for one binding site. The values for wild type are the mean of three determinants performed in triplicate.

5.2.8 Inhibition of L-HBsAg binding to HBcAg by synthetic peptides

Dyson and Murray (1995) showed that the synthetic peptide LDPAFR, corresponding to amino acids 19-24 of the PreS1 and also the epitope for monoclonal antibody MA18/7 (Heermann *et al.*, 1984), inhibits the association of L-HBsAg with HBcAg at a 50% inhibition concentration of around 350 μM. More recently, the structure of large HBcAg particles has been resolved to 7.4 Å with electron cryo-microscopy and image reconstruction, and a schematic diagram of the amino acid positions in the monomer has also been proposed (Böttcher *et al.*, 1997; see Fig.4, pg. 10). Interestingly, residues LEDPASR (corresponding to residues 76-82 of subtype *adyw*) located at the tip of the monomer show some similarities with the PreS1 epitope sequence. Furthermore, this region constitutes a big portion of the immunodominant loop (residues 74-89) of the protein and was speculated to be involved in the interaction with the surface proteins. It is, therefore, of interest to investigate whether the peptide LDPAFR may resemble the tip of the nucleocapsid spike and bind to a region on L-HBsAg, thus inhibiting the binding of HBsAg to HBcAg.

Two peptides of differents length, EDPASR and LEDPASR, which correspond to the residues at the tip of the nucleocapsid spike were synthesised and used in the inhibition experiments, and the results are shown in Fig. 36. The peptide LLGRMKG, selected from the fusion phage display library against HBcAg and included as a positive control, inhibited the interaction between L-HBsAg and HBcAg with 50% inhibition at a peptide concentration of ~100 μM. Clearly, both of the peptides EDPASR and LEDPASR did not inhibit the interaction, even at 1 mM concentration. The negative effect of peptide EDPASR was strongly supported by the fact that a combination of this peptide with LLGRMKG did not further inhibit the interaction. However, peptide LDPAFR showed some inhibition at 1 mM concentration, but this was a three-fold higher concentration than that observed by Dyson and Murray (1995). This difference could be due to reaction conditions used; i.e., incubation for 24 h (equilibrium) rather than 2.5 h that used by them.

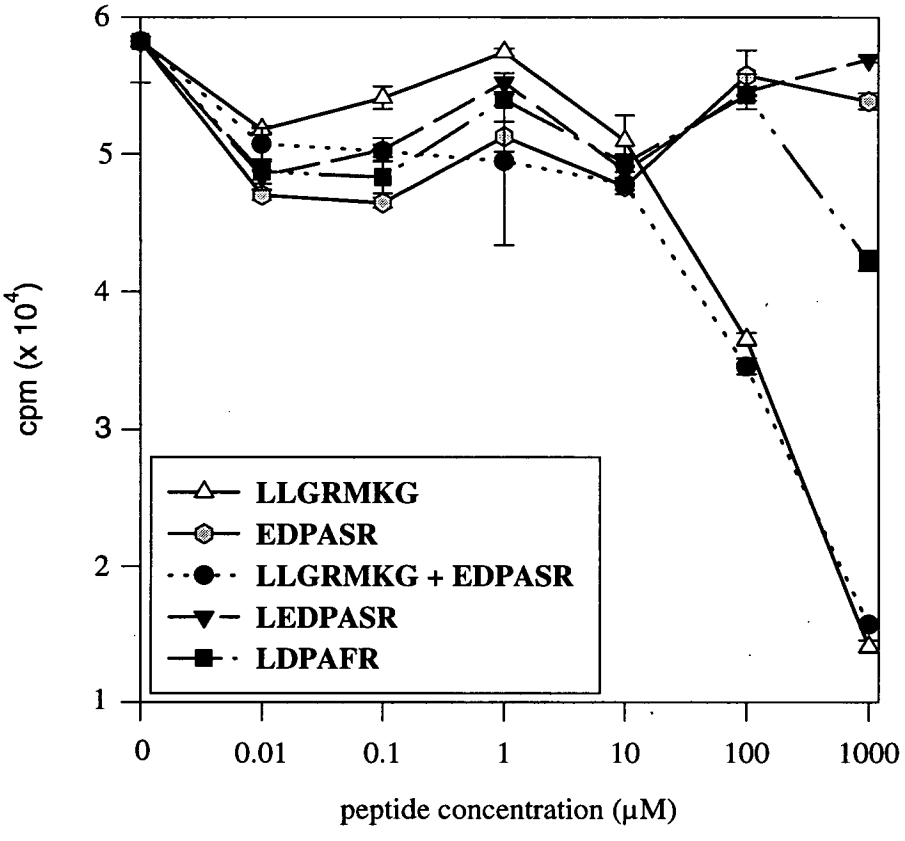


Fig. 36: Inhibition of L-HBsAg binding to HBcAg by synthetic peptides

A constant concentration of [^{35}S]L-HBsAg (100-fold dilution of a translation reaction) was mixed with serial dilutions of peptides ranging from 0 to 1 mM. The mixtures were then incubated in wells coated with HBcAg for 24 h, at 4 °C. The wells were then washed and radioactivity determined by scintillation counting. The symbols representing the peptides are indicated in the inserted box. The points are arithmetic mean ± standard deviation of triplicate determinations.

5.3 Discussion

5.3.1 Methods to study HBsAg-HBcAg interaction: from qualitative to quantitative

A major limitation to studies of the interaction between HBsAg and HBcAg is that one of the interacting components, HBsAg, cannot be expressed to a substantial amount in rabbit reticulocyte lysates to give adequate quantities of the protein for analysis in an ideal condition with most of the conventional protein-protein interaction methods such as surface plasmon resonance (SPR), nuclear magnetic resonance (NMR), light scattering, chemical cross linking, chemical modification or protection, etc. To overcome this problem, the parent and mutant proteins were radiolabelled with [^{35}S]-methionine and their binding to HBcAg was identified qualitatively with immunoprecipitation and studied quantitatively by a newly established binding assay.

In combination with phosphoimager analysis, the first method provides estimations of the binding affinity relative to the wild-type. However the data obtained could not be used for quantitative evaluation because the anti-core antibodies may compete for the same binding site with HBsAg. Furthermore, this method is not sensitive enough to discriminate between binding affinities that differ in magnitude by two- or three-fold, as may often result from a single amino acid substitution. Therefore attempts were taken to set up an equilibrium binding assay that allows quantitative description of the binding affinities of truncation and amino acid substitution mutants in order to gain a more complete picture of the contact sites that are involved.

The original technique described by Friguet *et al.* (1985) to determine the dissociation constants between antigens and antibodies was based on an enzyme-linked immunosorbent assay (ELISA) for measuring free antibody in solution using immobilised antigens. In the study described here, the sensitivity of the detection method was improved by labelling HBsAg with [^{35}S] and measuring radioactivity by scintillation counting. This permits the measurement of a very small concentration

of HBsAg, as low as 10^{-11} M, and thus eliminated the main constraint regarding the low yield of the protein synthesised in a cell free system.

Friguet *et al.*, (1995) also showed that the method can be used reliably to measure values of equilibrium constants without the need to purify the antibody to homogeneity or to titrate the antibody in the preparation, provided that the total antigen concentration is at least 10-fold in excess over the total antibody concentration. This is particularly useful for the case of HBsAg, because the protein cannot be overexpressed and it would be difficult, laborious, and time consuming to purify each of the mutant proteins. It is, therefore, the analogous assay that was performed with crude rabbit reticulocyte lysates and the concentration of total HBcAg [C_t] was always kept sufficiently in excess (>500-fold) of the HBsAg concentration. This requirement, however limits a precise measurement of K_d values smaller than 10^{-10} M.

In addition to the minor modification of the detection method itself, non-linear regression and Scatchard analyses were applied concurrently to analyse the binding data and to extract dissociation constants from them. Other than that, the experimental and theoretical basis for the equilibrium binding assay remain the same as those of the prototype assay. The dissociation constants obtained with this method were regarded as relative rather than true values because: (i) the interaction was assumed to be bimolecular, although there is a possibility that one molecule of HBsAg binds to a dimer of HBcAg because HBcAg particles are made from dimer intermediates. If this is the case, the number of HBsAg molecules remaining free in the solution will be twice the amount calculated above and thus the dissociation constants will be two-fold lower; (ii) the free HBcAg concentration [C] was approximated by the total HBcAg concentration [C_t] because the HBcAg was in large excess over HBsAg. These assumptions were used throughout this study to calculate dissociation constants, and should affect all the values to a similar degree, and thus a relative comparison should be valid.

5.3.2 Do fusion phage and L-HBsAg bind at the same docking site on HBcAg?

When Scatchard analysis was performed on the binding data of L-HBsAg with HBcAg, an upward-curved Scatchard plot was obtained. As described in several reports (Rosenthal, 1967; Dalquist, 1978; Zierler, 1989; Bordbar *et al.*, 1996), such phenomenon could be due to one or both of the following factors: (i) the system has two or more classes of binding sites with different affinities but occupancy to one site does not change the affinity of the binding to any other; (ii) there is a negative cooperativity in binding, in which binding at one site decreases the affinity of others. Examples of interactions displaying this kind of plot are: sodium n-dodecyl sulphate (SDS) with bovine catalase (Beyhl, 1986) and rabbit anti-pazobenzoate with its hapten p-iodobenzoate (Nisonoff and Pressman, 1958).

Clearly, the phenomenon is unlikely to be due to different species of HBcAg particles because the Scatchard plots of the fractionated small and large HBcAg particles have the same pattern, and their high affinities determined by non-linear regression were in good agreement with that of unfractionated HBcAg. This observation suggests that although HBcAg assembles into two sizes of particles with different triangulation number, these particles still possess the same binding sites for L-HBsAg.

Unlike L-HBsAg, the phage particles selected from a random hexapeptide display library against HBcAg particles showed linear Scatchard plots (Dyson *et al.*, 1995) when their binding constants were analysed with a method that also was adapted from Friguet *et al.* (1985). Employing the same method with one of the phage derivatives carrying the peptide RSLLGRMK, but with fractionated large and small HBcAg particles, revealed no significant difference in the K_{d1}^{rel} values. Furthermore, only one affinity was observed for the interactions. These results, together with the observations with L-HBsAg, suggest that the phages displaying specific ligands on their gpIII proteins bind to HBcAg in a different manner from that of L-HBsAg. This is no doubt due to the fact that the ligand mimics part of the contact areas in HBsAg, perhaps fitting into a 'cleft' on the surface of HBcAg. Whereas L-HBsAg and HBcAg might interact with each other over a large area,

small peptides derived from the selected phages may mimic part of the essential contact areas and will therefore hinder L-HBsAg from binding to HBcAg.

Peptides EDPASR and LEDPASR, which correspond to a linear sequence at the tip of HBcAg monomer, did not inhibit L-HBsAg from binding to HBcAg, but the peptide LDPAFR, which coincides with the epitope for monoclonal antibody MA18/7, did inhibit the interaction, but with a small effect at 1 mM concentration. Clearly, this result rules out the speculation that peptide LDPAFR resembles the structure of the tip of HBcAg monomer and somehow associates with L-HBsAg to prevent it from interacting with HBcAg. It is more likely that the peptide shares the same binding site on HBcAg with the sequences selected from the phage display library (LLGRMK, YLLRFR, LLGRFK, LLGRFR and other related sequences; Dyson and Murray, 1995) because the first and the last two residues of the hexapeptide appear in several of the selected sequences. Moreover, it has been shown by Dyson and Murray (1995) that the peptide LDPAFR not only inhibits L-HBsAg, but also the phage bearing amino acid sequence LLGRMK from binding to HBcAg.

5.3.3 Dissection of two binding domains in L-HBsAg

Recent analysis of the transmembrane topology of L-HBsAg in transfected cells or *in vitro* indicated that this protein adopts more than one transmembrane orientation, so that the PreS regions are disposed to both the cytoplasmic and luminal sides of the ER membrane (Prange and Streeck, 1995). It is most likely that the cytosolly disposed PreS regions offer docking sites for the envelopment of the viral nucleocapsid, while the luminally disposed PreS regions become exposed on the surface to serve as a receptor binding domain. Immunoprecipitation of the N-terminal deletion mutants showed that a 40 amino acids deletion from the N-terminus of L-HBsAg dramatically reduced its interaction with HBcAg, and a 77 amino acids deletion caused a further reduction in the interaction. This suggests that the PreS1 region plays a major role in the interaction. On the other hand, immunoprecipitation of the C-terminal deletion mutants indicated that the PreS

regions alone are not sufficient for interaction with HBcAg, an additional 28 amino acids of the S regions including the hydrophobic helix I being required to form a better docking site for HBcAg.

Equilibrium binding assays in solution showed that the interaction between L-HBsAg and HBcAg is mediated by two binding sites which extend beyond the PreS regions. Binding of mutants containing only the PreS regions (CΔ108, CΔ136, and CΔ163) was too weak for accurate measurement by curve fitting or Scatchard analysis (Fig. 37). However, the affinity increased to 4.4 μM (still a relatively weak interaction) for mutant CΔ191 which contains the PreS regions and helix I of the S-HBsAg. Also this truncation was judged to bind HBcAg in the immunoprecipitation assay (Fig. 24). This suggests that the additional 28 amino acids at the C-terminus of the PreS regions provide extra contact sites or perhaps form a better docking site for HBcAg than can be provided by the PreS regions alone. However, mutant CΔ210, which has 19 amino acids more than CΔ191 and was expected to exhibit stronger binding, actually showed a 2.5-fold reduction in K_d^{rel} value compared with CΔ191. This further suggests that the N-terminal 191 amino acids of the L-HBsAg form a binding domain that was sterically hindered by the additional 19 amino acids at the C-terminus of this fragment. Nevertheless, this domain might contribute towards part of the interactions provided by the full length L-HBsAg because mutant CΔ191 binds to HBcAg with a significantly weaker affinity, K_d^{rel} =4.4 μM, which is around 1000-fold weaker than the K_{d1}^{rel} of the wild type, suggesting that the N-terminal binding domain fills part or one of the binding sites on HBcAg and that additional contacts might be made with a second binding domain. The existence of the first binding domain is supported indirectly, perhaps, by several individual reports showing that the entire PreS regions are not co-translationally translocated into the ER lumen and remain at the cytosolic side (Ostapchuk *et al.*, 1994; Bruss *et al.*, 1994; and Prange and Streeck, 1995).

To locate the second binding domain, stop codons were introduced at various positions further downstream in the L-HBsAg gene, before and after the postulated hydrophobic helices II, III, and IV (mutants CΔ242, CΔ263, CΔ322, CΔ346 and CΔ371). Interestingly, mutants with a stop codon after helix II (see Fig. 37)

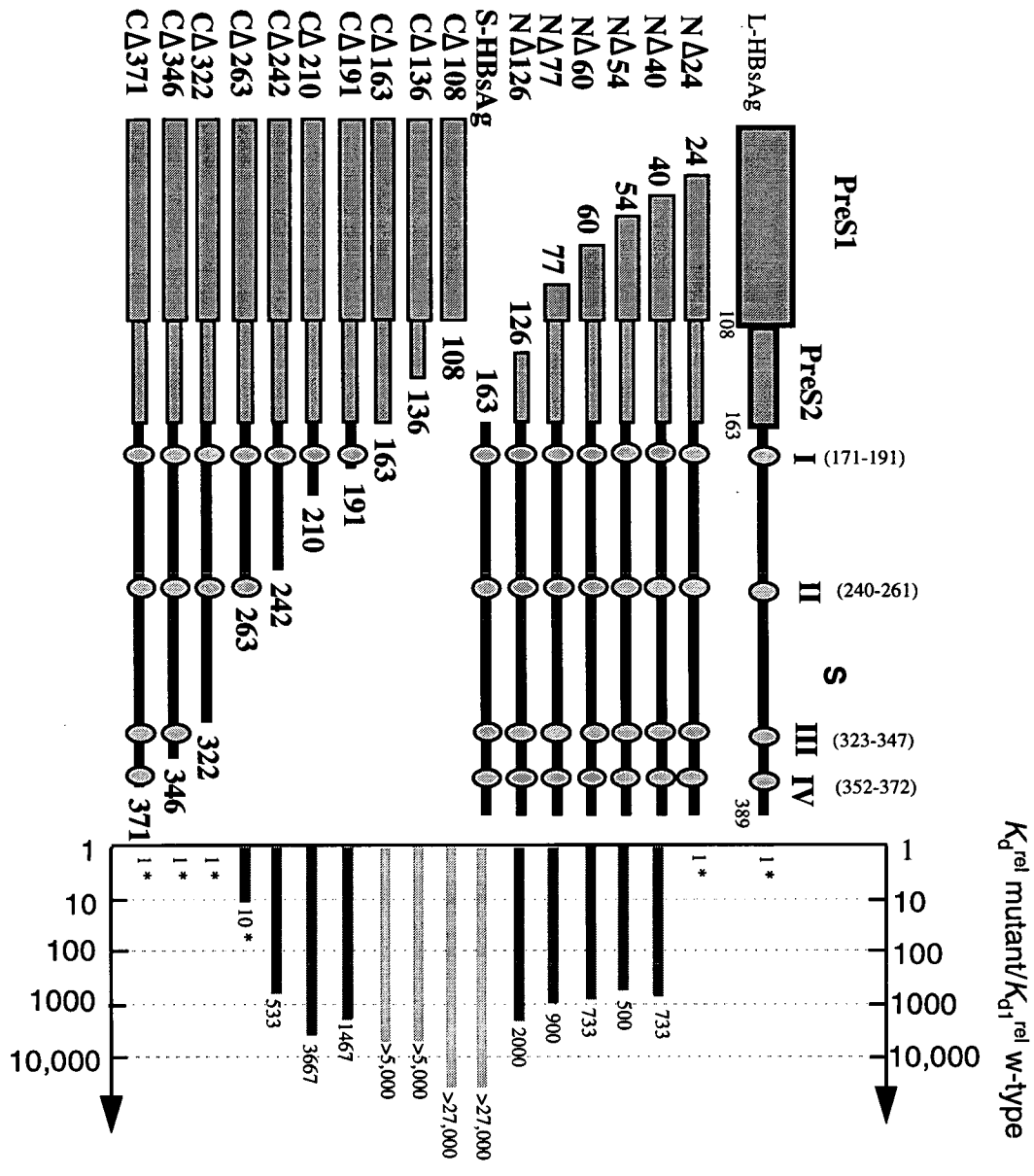


Fig. 37: Change in the relative dissociation constant relative to wild type for N- and C- terminal deletion mutants.

The diagram at the top represents domain organisation of L-HBsAg and, the N- and C-terminal deletion mutants. Four predicted transmembrane helices in the S-region are indicated by ovals, and their amino acid positions are printed in parenthesis. The vertical lines beneath each mutant represent the ratio of the K_d^{rel} of the corresponding mutant over the high-affinity (K_{d1}^{rel}) of L-HBsAg. Mutants displaying two affinities are indicated by asterisk and their K_{d1}^{rel} values were used in comparison with the high affinity of the wild type. Please refer to Table 8 (pg. 106) for the K_d^{rel} values for the mutants.

exhibited two binding affinities indicating the existence of a second binding domain. However, the wild type-affinities were only fully restored in mutants (CΔ322, CΔ346 and CΔ371) which contain the second hydrophilic loop of the S region.

These results, together with those from other smaller C-terminal deletions, suggest that some of the amino acids between hydrophobic helices I and III are necessary for constituting a second functional binding domain which may lie next to, and perhaps overlap or interact with the N-terminal binding domain in order to optimise interaction with corresponding binding sites on the core particle. Each domain may provide a principal binding site for its corresponding target site on HBcAg and its partner may contribute additional stability to the association. Disruption of either one of these binding domains by N- or C-terminal deletions from full-length L-HBsAg , as shown with mutants NΔ40, NΔ54, NΔ60, NΔ77, NΔ126, CΔ191, CΔ210, and CΔ242, resulted in linear Scatchard plots with much weaker binding affinities (more than 500-fold weaker) than those of full-length L-HBsAg. This further supports the existence of a synergistic interaction between the two binding domains which governs the high affinity binding with HBcAg. Recently, Poisson *et al.*(1997) demonstrated that two synthetic peptides, one corresponding to the 13 C-terminal amino acids of PreS1 plus the 8 N-terminal amino acids of PreS2, and the other to amino acids 56 to 80 of the S region bound directly to HBcAg. Again this supports the argument that two distinct binding domains occur in L-HBsAg.

The high affinities of mutants NΔ24, CΔ322, CΔ346 and CΔ371 for HBcAg are similar to that of the wild type, implying that the first 24 amino acids and the last 66 amino acids, which include the putative hydrophobic helices III and IV (Stirk *et al.*, 1992; Berting *et al.*, 1995) of S-HBsAg are not critical for the interaction. It also implies that the epitope of monoclonal antibody M18/7 with peptide sequence LDPAFR of residues 19-24 of the PreS1 region is unlikely to contribute to the interaction. When 40 residues of the L-HBsAg were deleted from the N-terminus, only one affinity was detected with a K_d^{rel} of 2.2 μM which is around 700-fold weaker than that of wild type. Further deletion of the PreS1 and PreS2 regions

(mutants NΔ40, NΔ54, NΔ60, NΔ77, NΔ126), further decreased but did not totally eliminate the interaction, as demonstrated by both the immunoprecipitation and equilibrium binding experiments. These results corroborate the finding that deletion of 102 amino acids of the PreS1 region of the HBV subtype *adw* (corresponding to subtype *adyw* amino acid 91) does not prevent virion maturation (Bruss and Thomssen, 1994). A possible explanation for this observation is that the second binding domain may still interact weakly with HBcAg to mediate assembly of the virus, although the first binding domain was partially disrupted. Surprisingly, our results show that S-HBsAg still binds to the HBcAg, but with a very weak affinity (> 80 μM) which suggests that the S region could provide an extra weak contact with HBcAg which is consistent with the finding that a 25-residue peptide corresponding to amino acid sequence 56-80 in the S region interacts with HBcAg (Poisson *et al*, 1997).

Binding results from the deletion mutants show that two interdependent binding domains which extend to the second hydrophilic loop of the S region (between helices II and III) are required for strong interaction with HBcAg. From simple topological consideration, the domains might protrude into the cytoplasm in order to interact efficiently with the corresponding binding sites on HBcAg. Therefore, it is likely that the first and second hydrophobic helices lie on the surface of the membrane or inside the binding domains, rather than spanning the membrane as predicted (Eble *et al.*, 1986, 1987; Simon *et al.*, 1988). The cytosolic location of these binding domains may influence the topology to the succeeding hydrophobic helices III and IV near the C-terminus of the S region, which were shown not to be involved in the interaction. However, this model is highly speculative, because the study was carried out in a cell free system. Nevertheless, it provides a straightforward explanation of the heterogeneous orientation of the S region of L-, M-, and S-HBsAg, in which the second hydrophilic loop carrying antigenic determinants can be both cytosolly and luminally disposed (Prange and Streeck, 1995). In addition, it accounts for the partial glycosylation of the envelope proteins at Asn-146 in the S region (Eble *et al.*, 1987).

5.3.4 Amino acid residues critical for the interaction between L-HBsAg and HBcAg

The interaction of the L-HBsAg with HBcAg was inhibited by the peptide ALLGRMKG and other related sequences selected from a random hexapeptide display library with purified HBcAg (Dyson and Murray, 1995). These sequences could reflect a region of the L-HBsAg that contacts the nucleocapsid, but they do not occur as a contiguous sequence in the L-HBsAg sequence and therefore they probably represent a mimic of the docking sites for HBcAg. However, three out of six matching residues from the sequence LLGRMK were identified at positions 63-65 (LLG) in the PreS1 region and residues 21-26 (LLTRIL) in hydrophobic helix I of the S region (Dyson and Murray, 1995). The latter peptide sequence is unlikely to contribute to the interaction because this is not inhibited by the peptide ALLTRILG, but the role of tripeptide LLG (63-65) in the interaction was tested by introducing point mutations in this region. Substitution of Leu-63 with His and Leu-64 with Glu each individually reduced the high affinity interaction 4-fold, whereas mutant G65R slightly enhanced it (Fig. 38). A double mutant, L63A/G65E displayed an 8-fold decrease in K_{d1}^{rel}. Since these mutations did not dramatically weaken the interaction with HBcAg, it is difficult to draw definite conclusions about the contribution of residues LLG (63-65) to the interaction.

The impact of positively charged amino acids within this domain was then examined because two positively charged amino acids appear in the fusion phage sequence (LLGRMK). To minimise the effects on protein folding, these amino acids were exchanged for Ala. Mutation of Lys-38, Arg-102, Arg-124, Arg-126 and Arg-156 all individually reduced the interaction by 2- to 5-fold (Fig. 38). However, the affinities of mutants R88A and R187A were very similar to that of the wild type, suggesting these two positively charged amino acids are not required for the interaction. Single substitution of Arg-92 to Ala decreased the binding affinities more than any other single mutation tested, around 11-fold relative to the wild type. It is thus likely that Arg-92 is involved in the interaction. This was further supported by double mutation of Arg-92 and Arg-102 to Ala which caused a further 7 fold decrease in affinity. The multiple mutations, L63H/R88A/R102A/R187A and

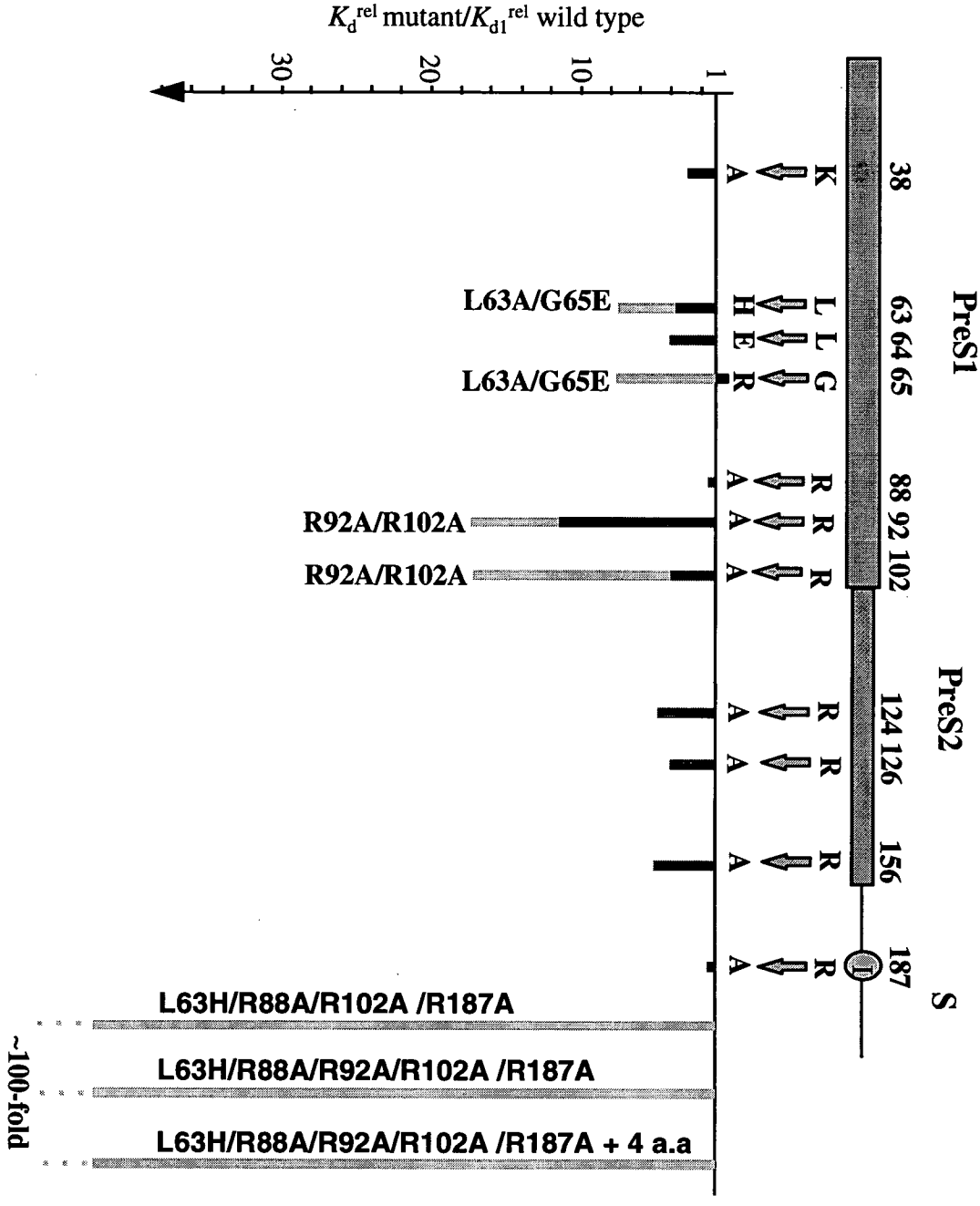


Fig. 38: The effect of single, double, and multiple mutations on the relative dissociation constant.

The diagram at the right represents the PreS and part of S regions. The positions of the amino acid substitutions are indicated on the right of the diagram. The changes made are indicated by arrows. The ratio of the high-affinity (K_{d1}^{rel}) of the single amino acid mutation over the high-affinity of the wild type is represented by a dark line on the left of each arrow. The impacts of double and multiple mutations are shown by light lines. Please refer to Table 9 (pg. 111) for the K_d^{rel} values.

L63H/R88A/R92A/R102A/R187A, reduced the affinity ~100-fold, but insertion of 4 amino acids (QPTP) at position 97 of the quintuple mutant, which was believed to disrupt the conformation of the PreS regions, did not decrease the affinity further. This implies that a second binding domain in the S region is involved in the interaction with HBcAg.

SUMMARY AND CONCLUSION

Envelopment of HBV is thought to be triggered by the interaction of cytoplasmic nucleocapsids with surface proteins inserted in the ER membrane of its host cell. However, very little is known about this event since the virus cannot be propagated in cell culture. A precise definition of the interaction between HBsAg and HBcAg requires a high resolution structure showing the locations of interacting groups at or near the interface. In the absence of structural information, identification of interacting residues with a mutagenesis approach is a bit like looking for a needle in a haystack, especially when the contacting area may be large and involves many residues. In order to map the binding sites in L-HBsAg, a series of N- and C-terminal deletion mutants were generated by *BAL 31* exonuclease and introduction of stop codons at various positions in the coding sequence. Site-directed mutagenesis was also applied to introduce specific amino acid substitutions in L-HBsAg, particularly in the PreS regions, to disrupt intermolecular interactions.

Biochemical and biophysical analyses of the interaction require the expression of the mutant proteins and their counterpart in appropriate systems. It is unfortunate that HBsAg cannot be expressed efficiently in *E.coli*. Alternatively, rabbit raticulocyte lysates, an eukaryotic cell-free system, were used to express the parent and mutant proteins. To compensate for the low yield of the biosynthetic products, they were labelled with [^{35}S] which permits their detection with autoradiography and scintillation counting. Unlike the HBV surface proteins, HBcAg can be produced easily in *E.coli*, and supplementation of the cell with T4 AGA tRNA significantly improved its yield. In the prokaryotic cell, HBcAg assembles into large and small particles. These particles can be separated from each other by sucrose gradient centrifugation and their size distribution was found to be essentially monodisperse by light scattering, but their monomers were heterogeneous at the C-terminal region, probably due to incomplete synthesis of the protein or protease degradation at the carboxy end.

Following mutagenesis and expression, the next challenge was to characterise the interactions of mutated HBsAg with the nucleocapsid. Two methods were employed to evaluate the impact of mutations on binding;

(i) immunoprecipitation of the interacting molecules with polyclonal anti-core antibodies and, (ii) measurement of relative dissociation constants with a newly established equilibrium binding assay in solution. Scatchard and non-linear regression analyses of the binding data of L-HBsAg with either fractionated large and small, or unfractionated HBcAg particles showed two widely differing affinities, suggesting the existence of two components to the binding sites between the molecules. In contrast to L-HBsAg, only one affinity was observed for the binding of phage bearing specific ligands to these particles, suggesting a different mode of interaction between the phage and L-HBsAg with HBcAg, in which the later displays a more complex behaviour. The interactions between HBsAg and HBcAg are believed to involve a large area which is mediated by two interdependent binding domains. Separately, these domains bind with low affinities to their respective binding sites on HBcAg, but an intact complex consists of these two domains bound to the HBcAg through high- and low-affinitiy interactions. In addition, point mutation studies show that the Arg residue at position 92 was particularly important for the interaction.

Elucidation of the regions and amino acids in L-HBsAg that are required to interact efficiently with HBcAg have furnished clues to the structure and function of these molecules, especially L-HBsAg, during assembly of the virus. However, many questions remain to be answered, particularly the regions on HBcAg that are involved with the binding domains in L-HBsAg. The ultimate goal of study of the interaction between the two molecules is to design anti-HBV drugs to inhibit the essential contact area. However, progress in this field has been hampered by the lack of structural information at high resolution. This goal will certainly occupy the experiments of many scientists in the fields of chemistry, biochemistry, molecular biology, structural biology and pharmacology who will pursue the problem in the next millennium.

REFERENCES

Aota, S.J., Gojobori, T., Ishibashi, F., Maruyama, T. and Ikemura, T. (1988). Codon usage tabulated from GeneBank sequence data. *Nucl. Acids Res.* **16** (suppl): r-315-r402.

Argos, P. and Fuller, S.D. (1988). A model for the hapatitis B virus core protein: prediction of antigenic sites and relationship to RNA virus capsid proteins. *EMBO J.* **7**: 819-824.

Avantaggiati, M.L., Natoli, G., Balsano, C., Chirillo, P., Artini, M., DeMarzio, E., Collepardo, D. and Levrero, M. (1993). The hepatitis B virus (HBV) PX trasnactivates the C-Fos promoter through multiple cis-acting elements. *Oncogene.* **8**: 1567-1574.

Bachmann, B.J. (1972). Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriol. Rev.* **36**: 525-557.

Balsano,C., Avantaggiati, M.L., Natoli, G., DeMarzio, E., Will, H., Perricaudet, M. and Levero, M. (1991). Full-length and truncated versions of the hepatitis B virus (HBV) X-protein (PX) transactivate the CMYC protooncogene at the transcriptional level. *Biochem. Biophys. Res.* **176**: 985-992.

Bartenschlager, R. and Schaller, H. (1988). The amino-terminal domain of hepanaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. *EMBO J.* **7**: 4185-4192.

Bass, S.H., Mulkerrin, M.G. and Wells, J.A. (1991). A systematic mutational analysis of hormon-binding determinants in the human growth hormone receptor. *Proc. Natl. Acad. Sci. USA.* **88**: 4498-4502.

Bayer, M., Blumberg, B. and Werner, B. (1968). Particles associated with Australia antigen in the sera of patients with leukemia, Down's Syndrome and hepatitis. *Nature.* **218**:1057-1059.

Benn, J. and Schneider, R.J. (1994). Hepatitis B virus HBx protein activates Ras-GTP complex formation and establishes a Ras, Raf, MAP kinase signaling cascade. *Proc. Natl. Acad. Sci. USA.* **91**: 10350-10354.

Berting, A., Hehnen, J., Kroeger, M. and Gerlich, W.H. (1995). Computer-aided studies on the spatial structure of the small hepatitis B surface protein. *Intervirol.* **38**: 8-15.

Beyhl, F.E. (1986). Interactions of detergents with microsomal enzymes: effects of some ionic and non-ionic detergents on glucose-6-phosphatase. *Med. Sci.* **14**: 417-418.

Birnbaum, F. and Nassal, M. (1990). Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. *J.virol.* **64**: 3319-3330.

Birnboim, H.C. (1983). A rapid alkaline extraction method for the isolation of plasmid DNA. *Method Enzymol.* **100**: 243-255.

Birboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**: 1513-1523.

Bonekamp, F. and Jensen, K.F. (1988). The AGG codon is translated very slowly in *E.coli* even at very low expression level. *Nucl. Acids Res.* **16**: 3013-3020.

Bordbar, A.K., Saboury, A.A. and Moosavi-Movahedi, A.A. (1996). The shapes of Scatchrd plots for systems with two sets of binding sites. *Biochem. Edu.* **24**: 172-175.

Borisova, G., Arya, B., Dislers, A., Borschukova, O., Tsibinogin, V., Skrastina, D., Eldarov, M.A., Pumpens, P., Skryabin, K.G. and Grens, E. (1993). Hybrid hepatitis B virus nucleocapsid bearing an immunodominant region from hepatitis B virus surface antigen. *J. Virol.* **67:** 3696-3701.

Borisova, G.P., Berzins, I., Pushko, P.M., Pumpen, P., Gren, E.J., Tsibinogin, V.V., Loseva, V., Ose, V., Ulrich, R., Siakkou. H. and Rosenthal, H.A. (1989). Recombinant core particles of hepatitis B virus exposing foreign antigenic determinants on their surface. *FEBS Letters.* **259**: 121-124.

Böttcher, B., Wynne, S.A. and Crowther, R.A. (1997). Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature.* **386**: 88-91.

Bowie, J.U., Reidhaar-Olsen, J.F., Lim, W.A. and Sauer, R.T. (1990). Deciphering the message in protein sequences: tolerence to amino acid substitutions. *Science.* **247**: 1306-1310.

Bruss, V. and Ganen, D. (1991a). Mutational analysis of hepatitis B surface antigen particle assembly and secretion. *J. virol.* **65**: 3813-3820.

Bruss, V. and Ganem, D. (1991b). The role of envelope proteins in hepatitis B virus assembly. *Proc. Natl. Acad. Sci. USA.* **88**: 1059-1063.

Bruss, V., Lu., X., Thomssen, R. and Gerlich, W.H. (1994). Post-translational alterations in transmembrane topology of the hepatitis B virus large envelope protein. *EMBO J.* **13**: 2273-2279.

Bruss, V. and Thomssen, R. (1994). Mapping a region of the large envelope protein required for hepatitis B virion maturation. *J. Virol.* **68**:1643-1650.

Bruss, V. and Vieluf, K. (1995). Functions of the internal Pre-S domain of the large surface protein in hepatitis B virus particle morphogenesis. *J. Virol.* **69**: 6652-6657.

Buendia, M.A. (1992). Hepatitis B viruses and hepatocellular carcinoma. *Adv. Cancer Res.* **59**: 167-226.

Bullock, W.O., Fermamdez, J.M. and Hunter, W.M. (1978). XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with β-galactosidase selection. *Biotechniques*. **5**: 376-378.

Burrell, C.J., MacKay, P., Greenaway, P.J., Hofshneider, P.H. and Murray K. (1979). Expression in *Escherichia coli* of hepatitis B virus DNA sequences cloned in plasmid pBR322. *Nature*. **279**: 43-47.

Cheng, K., Smith, G.L. and Moss, B. (1986). Hepatitis B virus large surface protein is not secreted but is immunogenic when selectively expressed by recombinant vaccinia virus. *J. Virol.* **60**: 337-344.

Cheong, J.H., Yi, M.K., Lin, Y. and Murakami, S. (1995). Human RPB5, a subunit shared by eukaryotic nuclear RNA polymerases, binds human hepatitis B virus X protein and may play a role in X transactivation. *EMBO J.* **14**: 143-150.

Chisari, F.V., Flippi, P., McLachlan, A., Millich, D.R., Riggs, M., Lee, S., Palmiter, R.D., Pinkert, C.A. and Brinster, R.L. (1986). Expression of hepatitis B virus large envelope polypeptide inhibits hepatitis B surface antigen secretion in transgenic mice. *J. Virol.* **60**: 880-887.

Chothia, C. (1976). The nature of the accessible and buried surfaces in proteins. *J. Mol. Biol.* **105:** 1-12.

Clarke, B.E., Newton, S.E., Carroll, A.R., Francis, M.J., Appleyard, G., Syred, A.D., Highfueld, P.E., Rowlands, D.J. and Brown , F. (1987). Improved immunogeneicity of a peptide epitope after fusion of hepatitis B core protein. *Nature*. **330**: 381-384.

Cleland, J.K., Jones, A.J.S. and Craik, C.S. (1995). Introduction to protein engineering. In Cleland, J.L. and Craik., C.S. (ed.) *Protein engineering: principles and practicle*. New York, USA:John Wiley and Sons, PP. 1-32.

Cohen, S.N., Chang, A.C.Y. and Hsu, L. (1972). Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc. Natl. Acad. Sci. USA*. **69**: 2110-2114

Cohen, B.J. and Richmond, J.E. (1982). Electron microscopy of hepatitis B core antigen synthesized in *Echerichia coli. Nature*. **296**: 677-678.

Cornish-Bowder, A. (1995). *Fundamentals of Enzyme Kinetics*. London, UK: Portland Press.

Creighton, T. E. (1984). *Proteins*. New York, USA: W.H. Freeman and Company.

Cross, J.L., Wen, P. and Rutter, W.J. (1993). Transactivation by hepatitis B virus X protein is promiscuous and dependent on mitogen-activated cellular serine threonine kinase. *Proc. Natl. Acad. Sci. USA*. **90**: 8078-8082.

Crowther, R.A., Kiselev, N.A., Böttcher, B., Berriman, J.A., Borisova, G.P., Ose, V. and Pumpens, P. (1994). Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. *Cell.* **77**: 943-950.

Dalquist, F.W. (1978). The meaning of Scatchard and Hill plots. *Method Enzymol.* **48**:270-299.

Dane, D.S., Cameron, C.H. and Briggs, M. (1970). Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet*, **I**: 695-698.

Darget, M. and Ehrlich, S.D. (1974). Prolonged incubation in calcium chloride improves competence of *Escherichia coli* cells. *Gene.* **6**: 23-28.

Dimmock, N.J. and Primrose, S.B. (1994). *Introduction to Modern Virology.* Cambridge, Great Britain, University Press.

Doria, M., Klein, N., Lucito, R. and Schneider, R.J. (1995). The hepatitis B virus HBx protein is a dual specificity cytoplasmic activator of Ras and nuclear activator of transcription factor. *EMBO J.* **14**: 4747-4757.

Dudley, F.J., Schever, P.J. and Sherlock, S. (1972). Natural history of hepatitis B associated antigen: positive chronic liver disease. *Lancet.* **ii**: 1388-1393.

Dyson, M.R., Germaschewski, V. and Murray, K. (1995). Direct measurement via phage titre of the dissociation constants in solution of fusion phage-substrate complexes. *Nucl. Acids Res.* **23**: 1531-1535.

Dyson, M.R. and Murray, K. (1995). Selection of peptide inhibitors of interactions involved in complex protein assemblies: association of the core and surface antigens of hepatitis B virus. *Proc. Natl. Acad. Sci. USA.*. **92**: 2194-2198.

Eble, B.E., Lingappa, V.R. and Ganem, D. (1986). Hepatitis B surface antigen and unusual secreted protein initially synthesized as a transmembrane polypeptide. *Mol. Cell Biol.* **6**: 1454-1463.

Eble, B.E., Lingappa, V.R. and Ganem, D. (1990). The N-terminal (PreS2) domain of a hepatitis B virus surface glycoprotein is translocated across membranes by downstream signal sequence. *J. Virol.* **64**: 1414-1419.

Eble, B., Macroe, D., Lingappa, V. and Ganem, D. (1987). Multiple topogenic sequences determine the transmembrane orientation of hepatitis B surface antigen. *Mol. Cell Biol.* **7**: 3591-3601.

Edman, J.C., Hallewell, R.A., Valenzuela, P., Goodman, H.M. and Rutter, W.J. (1981). Synthesis of hepatitis B surface and core antigen in *Escherichia coli. Nature.* **291**: 503-506.

Emilsson, V., Naslund, A.K. and Kurland, C.G. (1993). Growth rate dependent accumulation of twelve tRNA species in *Escherichia coli. J. Mol. Biol.* **230**: 483-491.

Fallows, D. and Goff. S.P. (1995). Mutations in the epsilon sequences of human hepatitis B virus affect both RNA encapcidation and reverse trnscription. *J.Virol.* **69**: 3067-3073.

Fernholz, D., Galle, P.R. Stemler, M., Brunetto, M., Bonino, F. and Will, H. (1993). Infectious hepatitis B virus variant defective in Pre-S2 protein expression in a chronic carrier. *Virol.* **194**: 137-148.

Friguet, B., Chaffotte, A.F., Djavadi-Ohaniance, L. and Goldberg, M.E. (1985). Measurements of the true affinity constant in solution of antigen-antibody complexes by enzyme-linked immunosorbent assay. *J. Immunol. Methods.* **77**: 305-319.

Gallina, A., Bonelli, F., Zentilin, L., Rindi, G., Muttini, M. and Milanesi, G. (1989). A recombinant hepatitis B core antigen polypeptide with the protamine-like domain deleted self-assembles into capsid particles but fails to bind nucleic acids. *J. virol.* **63**: 4645-4652.

Ganem, D. (1991). Assembly of Hepadnaviral virions and subviral particles. *Current Topics in Microbiol. and Immunol.* **168**: 61-63.

Ganem, D. and Varmus, H.E. (1987). The molecular biology of the hepatitis B viruses. *Annu. Rev. Biochem.* **56**: 651-693.

Gavilanes, F., Gonzalez-Ros, J.M. and Peterson, D.L. (1982). Structure of hepatitis B surface antigen: characterization of the lipid components and their association with the viral proteins. *J. Biol. Chem.* **257**: 7770-7777.

Gerin, J., Purcell, R., Hoggan, M., Holland, P. and Chanock, R. (1969). Biophysical properties of Australia antigen. *J. Virol.* **4**: 763-768.

Gibbs, C.S. (1995). Structure-functions relationships for protein design. In Cleland, J.L. and Craik., C.S. (ed.) *Protein engineering: principles and practicle.* New York, USA:John Wiley and Sons, PP. 317-347.

Gibson, T.J. (1984). Studies on Epstein Barr virus genome. PhD thesis, Cambridge university, U.K.

Goding, J. W. (1978). Use of staphylococcal protein A as an immunological reagent. *J. Immunol. Methods.* **20**: 241-253.

Gough, N.M. and Murray, K. (1982). Expression of HBV surface, core and e antigens by stable rat and mouse cell lines. *J. Mol. Biol.* **162**: 43-67.

Hancock, A.A., DeLean, A.L. and Lefkowitz R.J. (1979). Quantitative resolution of beta-adrenergic receptor subtypes by selective ligand-binding: application of a computerized model fitting technique. *Mol. Pharmacol.* **16**: 1-9.

Hatton, T., Zhou, S.L. and Standring, D.N. (1992). RNA-binding and DNA-binding activities in hepatitis B virus capsid protein: a model for their roles in viral replication. *J. Virol.* **66**: 5232-5241.

Heermann, K.H., Goldmann, U., Schwartz, W., Seyffarth, T., Baumgarten H. and Gerlich, W.H. (1984). Large surface proteins of hepatitis B virus containing the Pre-S sequence. *J. Virol.* **52**: 396-402.

Hilditch, C.M., Rogers, L.J. and Bishop, D.H.L. (1990). Physicochemical analysis of the hepatitis B virus core antigen produced by a baculovirus expression vector. *J. Gen. Virol.* **71**: 2755-2759.

Hoofnagle, T.H., Dusheiko, G.M., Seeff, L.B., Jones, E.A.Waggoner, J.G. and Bales, Z.B. (1981). Seroconversion from hepatitis B e antigen to antibody in chronic type B hepatitis. *Annu. Inter. Med.* **94**: 744-748.

Hu, K.Q., Vierling, J.M. and Siddiqui, A. (1990). Trans-activation of HLA-DR gene by hepatitis B virus X gene product. *Proc. Natl. Acad. Sci. USA.* **87**: 7140-7144.

Hu, K.Q., Yu, C.H. and Vierling, V.J. (1992). Up-regulation of intercellular-adhesion-molecule-1 transcription by hepatitis B virus X-protein. *Proc. Natl. Acad. Sci. USA.* **89**: 11441-11445.

Hulme, E.C. and Birdsall N.J.M. (1992). Strategy and tactics in receptor-binding studies. In Hulme, E.C. (ed) *Receptor-ligand interactions, a pratical approach.* Oxford: IRL press, pp. 63-176.

Huovila, A.J., Elder, A.M. and Fuller, S.D. (1992). Hepatitis B virus surface antigen assembles in a post-ER, pre-Golgi compartment. *J. Cell Biol.* **118:** 1305-1320.

Hutchison, C.A., Phillips, S., Edgell, M.H., Gillam, S., Jahnke, P. and Smith, M. (1978). Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **253**: 6551-6560.

Ikemura, T. (1982). Correlation between the abundance of *Escherichia coli* tRNAs and the occurrence of respective codons in its protein genes: a proposal for synonomous codon choice that is optimal for the *E.coli* translating system. *J. Mol. Biol.* **151**: 389-409.

Janin, J. (1979). Surface and inside volumes in globular proteins. *Nature.* **277**: 491-492.

Janin, J. and Chothia, C. (1990). The structure of protein-protein recognition sites. *J. Biol. Chem.* **265**: 16027-16030.

Kenny, J.M., Bonsdorff, C-H., Nassal, M. and Fuller, S.D. (1995). Evolutionary conservation in hepatitis B virus core structure: composition of human and duck cores. *Structure*. **3**: 1009-1019.

Kim, C.M., Koike, K., Saito, I., Miyamura, T. and Jay, G. (1991). HBx gene of hepatitis B virus induces liver cancer in transgenic mice. *Nature*. **361**: 742-745.

Khudyakov, Y.E. and Makhov, A.M. (1989). Prediction of terminal protein and ribonucleases H domains in the gene P products of hepadnaviruses. *FEBS Letters*. **243**: 115-118.

Klotz, K.M. and Hunston, D. (1971). Properties of graphical representations of multiple classes of binding sites. *Biochem*. **10**: 3065-3069.

Knaus, T. and Nassal, M. (1993). The encapsidation signal on the hepatitis B virus-RNA pregenome forms a stem-loop structure that is critical for its function. *Nucl. Acids. Res*. **21**: 3967-3975.

Kniskern, P.J., Hagopian, A., Montgomery, D.L., Burke, P., Dunn, N.R., Hofmann, K.J., Miller, W.J. and Ellis, R.W. (1986). Unusually high expression of a foreign gene (hepatitis B virus core antigen) in *Saccharomyces cerevisiae. Gene*. **46**: 135-141.

Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. U.S.A*. **82**: 488-492.

Kunkel, T.A., Roberts, S.D. and Zakour, R.A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Method Enzymol*. **154**: 367-382.

Kuroki, K., Floreani, M., Mimms, L.T. and Ganem, D. (1990). Epitope Mapping of the PreS1 Domain of the hepatitis B virus large surface protein. V*irol*. **176**: 620-624.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **227**: 680-685.

Lee, T.H., Elledge, S.J. and Butel, J.S. (1995). Hepatitis B virus X protein interacts with a probable cellular DNA repair protein. *J. Virol*. **69**: 1107-1114.

Lien, J.M., Aldrich, C.E. and Mason, W.S. (1986). Evidence that a capped oligonucletides is the primer for duck hepatitis B virus plus strand synthesis. *J. Virol*. **57**: 229-237.

Lis, J.T. (1980). Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Method Enzymol*. **65**: 347-353.

Lis, J.T. and Schleif, R. (1975). Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucl. Acids Res*. **2**: 383-389.

Liu, C.C., Yansura, D. and Levinson, A. (1982). Direct expression of hepatitis B surface antigen in monkey cells from an SV40 vector. *DNA*. **1**: 213-221.

Loeb, D.D., Hirsch, R.C. and Ganem, D. (1991). Sequence-independent RNA cleavages generate the primers for plus strand DNA synthesis in Hepatitis B virus-implications for other reverse transcribing elements. *EMBO J.* **10**: 3533-3540.

McAleer, W.J., Buynak, E.B., Maigetter, R.Z., Wampler, D.E., Miller, W.J. and Hilleman, M.R. (1984). Human hepatitis B vaccine from recombinant yeast. *Nature*. **307**: 178-180.

Machida, A., Kishimoto, S., Ohnuma H., Baba, K., Ito, Y., Miyamoto, H., Funatsu, G. and Oda, K. (1984). A polypeptide containing 55 amino acid residues coded by the PreS region of hepatitis B virus deoxyribonucleic bears the receptor for polymerized human as well as chimpanzee albumins. *Gastroenterology*, **86**: 910-918.

Machida, A., Ohnuma, H., Tsuda, F., Yoshikawa, A., Hoshi, Y., Tanaka, T., Khishimoto, S., Akahane, Y., Miyakawa, Y. and Mayumi, M. (1991). Phosphorylation in the carboxyl terminal domain of the capsid protein of hepatitis B virus: evaluation with a monoclonal antibody. *J. Virol.* **65**: 6024-6030.

Mahe, Y., Mukaida, N., Kuno, K., Akiyama, M., Ikeda, N., Matsushima, K. and Murakami, S. (1991). Hepatitis B virus X protein transactivates human interleukin-8 gene through acting on nuclear factor-κB and CCAAT enhancer binding protein-like cis elements. *J. Biol. Chem.* **266**: 13759-13763.

Matsuda, K., Satoh, S. and Ohori, H. (1988). DNA-binding activity of hepatitis B e antigen polypeptide lacking the protamine-like sequence of nucleocapsid protein of human hepatitis B virus. *J. Virol.* **62**: 3517-3521.

Mazzara, G.P., Plunkett, G. and McClain, W.H. (1981). DNA sequence of the transfer-RNA region of bacteriophage T4: implications for transfer-RNA synthesis. *Proc. Natl. Acad. Sci. USA*. **78**: 889-892.

Messing, J., Gronenborn, B., Muller-Hill, B. and Hofschneider, P.H. (1977). Filamentous coliphage M13 as a cloning vehicle: insertion of a *Hin*dII fragment of the *lac* regulatory region in M13 replicative form *in vitro*. *Proc. Natl. Acad. Sci. USA*. **74**: 3642-3646.

Mimms L.T., Floreani, M., Tyner, J., Whitters, E., Rosenlof, R., Wray L, Goetze, A., Sarin, V. and Eble, K. (1990). Discrimination of hepatitis B virus (HBV) subtypes using monoclonal-antibodies to the PreS1 and PreS2 domains of the viral envelope. *Virol.* **176**: 604-619.

Misra, R. and Reeves, P. (1995). Intermediates in the synthesis of TolC protein include an incomplete peptide stalled at a rare Arg codon. *Eur. J. Biochem.* **152**: 151-155.

Miyanohara, A., Imamura, T., Araki, M., Sugawara, K., Ohtomo, N. and Matsubara, K. (1986). Expression of hepatitis B virus core antigen gene in *Saccharomyces cerevisae*: synthesis of two polypeptide translated from different initiation codons. J. *Virol.* **59**: 176-180.

Moos, M., Nguyen, N.Y. and Liu, T.Y. (1988). Reproducible, high yield sequencing of proteins electrophoretically separated and transferred to an innert support. *J. Biol. Chem.* **263**: 6005-6008.

Munson, P.J. and Rodbard, D. (1980). Ligand: a versatile computerized approach for characterization of ligand-binding systems. *Anal. Biochem.* **107**: 220-239.

Murray, K., Bruce, S.A., Wingfield, P., van Eerd, P., de Reus, A. and Schellekens, H. (1984). Hepatitis B virus antigen made in microbial cells immunise against viral infection. *EMBO J.* **3**: 645-650.

Nassal, M.(1992). The arginine-rich domain of the hepatitis B virus core protein is required for precore encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. *J. virol.* **66**: 4107-4116.

Nassal, M. and Rieger, A. (1993). An intramolecular disulfide bridge between Cys-7 and Cys-61 determines the structure of the secretory core gene product (e antigen) of hepatitis B virus. *J. Virol.* **67**: 4307-4315.

Nassal, M. and Schaller, H. (1993). Hepatitis B virus replication. *Trends in Microbiol.* **1**: 221-228.

Natoli, G., Avantaggiati, M.L., Chirillo, P., Puri, P.L. Ianni, A., Balsano, C. and Levrero, M. (1994). Ras- and Raf-dependent activation of C-Jun transcriptional activity by the hepatitis B virus transactivator pX. *Oncogene.* **9**: 2837-2843.

Nelson, R.M. and Long, G.L. (1991). Solution-phase equilibrium binding interaction of human protein S with C4b-binding protein. *Biochem.* **30**: 2384-2390.

Neurath, A.R., Kent, S.B.H., Strick, N. and Parker, K. (1986). Identification and chemical synthesis of a host cell receptor binding site on hepatitis B virus. *Cell.* **46**: 429-436.

Neurath, A.R., Stirck, N., Seto, B. and Girard, M. (1989). Peptides from the Pre-S1 region of the hepatitis B virus envelope protein as components of polyvalent hybrid vaccines. In *Vaccines (89):modern approaches to new vaccines including prevention of AIDS*. New York: Cold Spring harbor Laboratory, pp. 473-478.

Nisonoff, A. and Pressman, D. (1958). Heterogeneity and average combining constants of antibodies from individual rabbits. *J. Immunol.* **80**: 417-428.

Onodera, S., Ohori, H., Yamaki, M. and Ishida, N. (1982). Electron microscopy of human hepatitis B virus cores by negative staining-carbon film technique. *J. Med. Virol.* **10**: 147-155.

Oostra, B.A., Harver, R., Ely, B.K., Markham, A.F. and Smith, A.E. (1983). Transforming activity of polyoma virus middle T antigen probed by site-directed mutagenesis. *Nature.* **304**: 456.

Ostapchuk, P., Hearing, P. and Ganem, D. (1994). A dramatic shift in the transmembrane topology of a viral envelope glycoprotein accompanies hepatitis B viral morphogenesis. *EMBO J.* **13**: 1048-1057.

Ou, J.-H., Laub, O. and Rutter, W.J. (1986). Hepatitis B virus gene function: the precore region targets the core antigen to cellular membranes and causes the secretion of the e antigen. *Proc. Natl. Acad. Sci. USA.* **83**: 1578-1582.

Pasek, M., Goto, T., Gilbert, W., Zink, B., Schaller, H., MacKay, P., Leadbetter, G. and Murray, K. (1979). Hepatitis B virus genes and their expression in *E.coli. Nature.* **282**: 575-579.

Patzer, E.J., Nakamura, E.R., Simonsen, C.C., Levinson, A.D. and Brands, R. (1986). Intracellular assembly and packaging of hepatitis B surface antigen particles occurs in the endoplasmic reticulum. *J. Virol.* **58**: 884-892.

Patzer, E.J., Nakamura, E.R. and Yaffe, A. (1984). Intracellular transport and secretion of hepatitis B surface antigen in mammalian cells. *J.Virol.* **51**: 346-354.

Persing, D.H., Varmus, H.E. and Ganem, D. (1987). The PreS1 protein of hepatitis B virus is acylated at its amino terminus with myristic acid. *J. virol.* **61**: 1672-1677.

Petit, M.-A. and Pillot, J. (1985). HBc and HBe antigenicity and DNA-binding activity of major core protein P22 in hepatitis B virus core particles isolated from the cytoplasm of human liver cells. *J. Virol.* **53**: 543-551.

Poisson, F., Severac, A., Hourioux, C., Goudeau, A. and Roingeard, P. (1997). Both PreS1 and S domains of hepatitis B virus envelope proteins interact with the core particle. *Virol.* **228**: 115-120.

Pollack, J.R. and Ganem, D. (1993a). An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. *J. Virol.* **67**: 3254-3263.

Pollack, J.R. and Ganem, D. (1993b). Site-specific RNA binding by a hepatitis B virus reverse transcriptase initiates two distinct reactions: RNA packaging and DNA synthesis. *J. Virol.* **68**: 5579-5587.

Pontisso, P., Petit, M.A., Bankowski, M.J. and Peeples, M.E. (1989). Human liver plasma membrane contain receptor for hepatitis B virus PreS1 region and, via polymerized human serum albumin for the PreS2 region. *J. Virol.* **63**: 1981-1988.

Prange, R., Nagel, R. and Streeck, R. (1992). Deletions in the hepatitis B virus small envelope protein: effect on assembly and secretion of surface antigen particles. *J. Virol.* **66**: 5832-5841.

Prange, R. and Streeck, R. E. (1995). Novel transmembrane topology of the hepatitis B virus envelope proteins. *EMBO J.* **14**: 247-256.

Prasad, V.R. and Goff, S.P. (1989). Linker insertion mutagenesis of the human immunodeficiency virus reverse transcriptase expressed in bacteria: defination of the minimal polymerase domain. *Proc. Natl. Acad. Sci. USA.* **86**: 3104-3108.

Qadri, I., Maguire, H.F. and Siddiqui, A. (1995). Hepatitis B virus transactivator protein X interacts with the TATA-binding protein. *Proc. Natl. Acad. Sci. USA.* **92**: 1003-1007.

Radziwill, G., Tucker, W. and Schaller, H. (1990). Mutational analysis of the hepatitis B virus P gene product domain structure and RNAs H activity. *J. Virol.* **64**: 613-620.

Raney, A.K., Milich, D.R. and McLachlam, A. (1989). Characterisation of hepatitis B virus major surface antigen gene transcription regulatory elements in differentiated hepatoma cell lines. *J. Virol.* **63**: 3919-3925.

Redeker, A.G. (1975). Viral hepatitis: clinical aspects. *Am J. Med. Sci.* **270**: 9-16.

Reich, J.G., Wangermann, G., Falck, M. and Rohde, K. (1972). A general strategy for parameter estimation from isosteric and allosterik-kinetic data and binding measurements. *Eur. J. Biochem.* **26**: 368-379.

Reiger, A. and Nassal, M. (1995). Distinct requirements for primary sequence in the 5'- and 3'- part of a bulge in the hepatitis B virus RNA encapsidation signal revealved by a combined *in vivo* selection/ *in vitro* amplication system. *Nucl. Acids Res.* **23**: 3909-3915.

Robinson, M., Lilley, R., Little, S., Emtage, J.S., Yarranton, G., Stephens, P., Millacan, A., Eaton, M. and Humphreys, G. (1984). Codon usage can affect translation rate in *Escherichia coli. Nucl. Acids. Res.* **12**: 6663-6671.

Robinson, W. S. and Luttwick, L.I. (1976). The virus of hepatitis type B. *New Eng. J. Med.* **295**: 1168-1175.

Rosenthal, H.E. (1967). A graphic method for the determination and presentation of binding parameters in a complex system. *Anal. Biochem.* **20**: 525-532.

Rossi, J. and Zoller, M. (1987). Site-specific and regionally directed mutagenesis of protein encoding sequences. In Oxender, D.L., and Fox, C.F. (ed.) *Protein Engineering*. New York, USA: Alan R. Liss, INC, pp 51-64.

Rossinck, M.J., Jameel, S., Loukin, S.H. and Siddiui, A. (1986). Expression of hepatitis B viral core region in mammalian cells. *Mol. Cellular Bio.* **6**: 1393-1400.

Rossner, M.T. (1982). Hepatitis B virus X gene product: a promiscuous transcriptional activator. *J. Med. Virol.* **36**: 101-117.

Salfeld, J., Pfaff, E., Noah, M. and Schaller, H. (1989). Antigenic determinants and functional domains in core antigen and e antigen from hepatitis B virus. *J. Virol.* **63**: 798-808.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular cloning: a laboratory manual.* Cold Spring Harbour, New York. Cold Spring Harbor Laboratory Press.

Sanger, F. and Coulson, A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **94**: 441-448.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**: 5463-5467.

Saxena, P. and Walker, J.R. (1992). Expression of ArgU: the *Escherichia coli* gene coding for a rare arginine tRNA. *J. Bacteriol.* **174**: 1956-1964.

Schaller (1993). Molecular biology of Hepatitis B virus. *ZMBH 1992-93 report.* 118-130.

Schaller, H. and Fisher, M. (1991). Transcriptional control of Hepadnavirus gene expression. *Current Topic in Microbiol. and Immunol.* **168**: 21-39.

Schoedel., F., Moriaty, A.M., Peterson, D.L., Zheng, J., Hughes, J.L., Will, H., Leturcq, D.J., McGee, J.S. and Milich, D.R. (1992). The position of heterologous epitopes inserted in hepatitis B virus core particles determines their immunogenicity. *J. Virol.* **66**: 106-114.

Schroder, R., Maassen, A., Lippoldt, A., Borner, T., Vonbaehr, R. and Dobrowolski, P. (1991). Expression of the core antigen gene of hepatitis B virus (HBV) in *Acetobacter methanolicus* using broad host range vectors. *Appl. Microb. Biotech.* **35**: 631-637.

Scott, J.K. and Smith, G.P. (1990). Searching for peptide ligands with an epitope library. *Science.* **249**: 386-390.

Seeger, C., Summers, J. and Mason, W.S. (1991). Viral DNA synthesis. *Current Topic in Microbiol. and Immunol.* **168**: 41-59.

Seto, E., Benedict, T.S., Peterlin, B.M. and Ou, J. H. (1988). Transactivation of the human immunodeficiency virus long terminal repeat by the hepatitis B virus X protein. *Proc. Natl. Acad. Sci. USA.* **85**: 8262-8290.

Siddiqui, A., Jameel, S. and Mapoles, J. (1987). Expression of the hepatitis B virus X gene in mammalian cells. *Proc. Natl. Acad. Sci. USA.* **84**: 2513-2517.

Simon, K., Lingappa, V.R. and Ganem, D. (1988). Secreted hepatitis B virus surface antigen polypeptides are derived from a transmembrane precursor. *J. Cell. Biol.* **107**: 2163-2168.

Slagle, B.L., Lee, P.-H. and Butel, J.S. (1992). Hepatitis B virus and hepatocellular carcinoma. *Prog. Med. Virol.* **39**: 167-203.

Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B.H. and Hood, L.E. (1986). Fluorescence detection in automated DNA-sequence analysis. *Nature.* **321**: 674-678.

Smith, G.P. and Scott, J.K. (1993). Libraries of peptides and proteins displayed on filamentous phage. *Method Enzymol.* **217**: 228-257.

Sondek, J. and Shortle, D. (1990). Accommodation of single amino acid insertions by the native state of staphylococcal nuclease. *Proteins.* **7**: 299-305.

Sorensen, M.A., Kurland, C. and Pedersen, S. (1989). Codon usage determines translation rate in *Escherichia coli. J. Mol. Biol.* **207**: 365-377.

Spanjaard, R.A. and van Duin, J.E. (1988). Translation of the sequence AGG-AGG yields 50% ribosomal frameshift. *Proc. Natl. Acad. Sci. USA.* **85**: 7967-7971.

Stahl, S., MacKay, P., Magazin, M., Bruce, S. and Murray, K. (1982). Hepatitis B virus core antigen: synthesis in *Escherichia coli* and application in diagnosis. *Proc. Natl. Acad. Sci. USA.* **79**: 1606-1610.

Stahl, S.J. and Murray, K. (1989). Immunogenicity of peptide fusions to hepatitis virus core antigen. *Proc. Natl. Acad. Sci. USA.* **86**: 6283-6287.

Stevens, C.E., Toy, P.T., Tong, M.J.,Taylor, P.E., Vyas, G.N. Nair, P.V., Gudavalli, M. and Krugman, S. (1985). Perinatal hepatitis B virus transmission in the United States. *J. American Med. Association.* **253**: 1740-1745.

Stewart, F.S. (1993). Mutations that affect the structure and interactions of the core antigen of hepatitis B virus. PhD Thesis, University of Edinburgh.

Stirk, H.J., Thornton, J.M. and Howard, C.R. (1992). A topological model for hepatitis B surface antigen. *Intervirol.* **339**: 148-158.

Szmuness, W., Stevens, C.E., Harley, E.J. Zhang, A., Olwszko, W.R., William, D.C., Sodovsky, R., Morrison, J.M. and Kellner, A. (1980). Hepatitis B vaccine: demonstration of efficacy in a controlled clinical trial in a high-risk population in the United States. *New England J. Med.* **303**: 833-841.

Tabor, S. and Richardson, C.C. (1985). A bacteriophage T7 RNA polymerase promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA.* **82**: 1074-1078.

Takahashi, K., Kishimoto, S., Ohori, K., Yoshizawa, H., Machida, A., Ohnuma, H., Tsuda, F., Munekata, E., Miyakawa, Y. and Mayumi, M. (1991). Molecular heterogeneity of e antigen polypeptides in sera from carriers of hepatitis B virus. *J. Immunol.* **147**: 3156-3160.

Takahashi, K., Machisa, A., Funatsu, G., Nomura, M., Usuda, S., Aoyagi, S., Tachibana, K., Miyamoto, H., Imai, M., Makamura, T., Miyakawa, Y. and Mayumi, M. (1983). Immunochemical structure of hepatitis B e antigen in the serum. *J. Immunol.* **130**: 2903-2907.

Thomas, H.C. and Scully, L.J. (1985). Antiviral therapy in chronic hepatitis B virus infection. *British Med. Bulletin.* **41**: 374-380.

Toh, H., Hayashida, H. and Miyata, T. (1983). Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus. *Nature*: **305**: 827-829.

Twu, J.-S, Lai, M.-Y., Chen, D.-S. and Robinson, W.S. (1993). Activation of protooncogene C-Jun by the X protein of hepatitis B virus. *Virol.* **192**: 346-350.

Twu, J.-S and Robinson, W.S. (1989). Hepatitis B virus X gene can transactivate heterologous viral sequences. *Proc. Natl. Acad. Sci. USA.* **86**: 2046-2050.

Twu, J.-S. and Schloemer, R.H. (1987). Transcriptional transactivating function of hepatitis B virus. *J. Virol.* **61**: 3448-3453.

Ueda, K., Tsurimoto, T. and Matsubara, K. (1991). Three envelope proteins of hepatitis B virus: large S, middle S and major S proteins needed for the formation of Dane particles. *J. Virol.* **65**: 3521-3529.

Ulrich, R., Meisel, H., Soza, A., Krüger, D.H., Ladhoff, A.M., Borisova, G. and Pumper, P. (1993). Characterization of chimeric core particles of hepatitis B virus containing foreign epitopes. In *Vaccine (93): a modern approaches to new vaccines including prevention of AIDS.* New York: Cold Spring Harbor press, pp. 323-328.

Uy, A., Bruss, V., Gerlich, W.H., Kochel. H.G. and Thomssen, R. (1986). Precore sequence of hepatitis B virus inducing e antigen and membrane-association of the viral core protein. *Virol.* **155**: 89-96.

Valenzuela, P., Medina, A., Ruffer, W.J., Ammerer, G. and Hall, B.D. (1982) Synthesis and assembly of hepatitis B virus surface antigen particles in yeast. *Nature*. **298**: 347-350.

Will, H., Reiser, W., Weimer, T., Pfaff, E., Buscher, M., Sprengel, R., Cattaneo, R. and Schaller, H. (1987). Replication strategy of human hepatitis B virus. *J. Virol.* **61**: 904-911.

Wingfield, P.T., Stahl, S.J., Williams, R.W. and Steven, A.C. (1995). Hepatitis core antigen produced in *Escherichia coli*: subunit conformational analysis and *in vitro* capsid assembly. *Biochem.* **34**: 4919-4932.

Wunderlich, G. and Bruss, V. (1996). Characterization of early hepatitis B virus surface protein oligomers. *Arch. Virol.* **141**: 1191-1205.

Yaginuma, K and Koike, K. (1989). Identification of a promoter region for 3.6 kilobase mRNA of hepatitis B virus and specific cellular binding protein. *J. Virol.* **63**: 2914-2921.

Zahn, K. and Landy, A. (1996). Modulation of lambda integrase synthesis by rare arginine tRNA. *Mol. Microbiol.* **21**: 69-76.

Zheng, J., Schödel, F. and Peterson, D.L. (1992). The structure of hepanaviral core antigens. *J. Bio. Chem.* **267**: 9422-9428.

Zhou, D.-X., Taraboulous, A., Ou, J.-H. and Yen, T.S.B. (1990). Activation of class I major histocompatibility complex gene expression by hepatitis B virus. *J. virol.* **64**: 4025-4028.

Zhou, S.L. and Standring, D.N. (1991). Production of hepatitis B virus nucleocapsid-like core particles in *Xenopus* oocytes: assembly occurs mainly in the cytoplasm and does not require the nucleus. *J. Virol.* **65**: 5457-5464.

Zierler, K. (1989). Misuse of nonlinear Scatchard Plots. *Trends Biochem. Sci.* **14**: 314-317.

Zlotnick, A., Cheng, N., Conway, J.F., Steven, A.C., Stahl, S.J. and Wingfield, P.T. (1996). Dimorphism of hepatitis B virus capsids is strongly influenced by the C-terminus of the capsid protein. *Biochem.* **35**: 7412-7421.