https://theses.gla.ac.uk/

Theses Digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

Hepatitis B Virus Core Gene Deletions

by

David James Williams

A thesis presented for the degree of Doctor of Philosophy

in

The Faculty of Science
at the
University of Glasgow

Division of Virology
Church Street
Glasgow
G11 5JR

June 1997

ProQuest Number: 13815342

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 13815342

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ACKNOWLEDGEMENTS

I am grateful to Prof. J.H. Subak-Sharpe, Dr. H. Marsden and Profs. Clements and McGeoch, who have all contributed to running the Institute of Virology.

I would like to thank Dr. Bill Carman for his supervision of this project; for the lively journal club discussions (perhaps one day they will find e antigen binding to ε. If they do, I'm sure he'll be the first to let me know); and for leaving me alone to get on with the work. Thanks too for showing us that it is possible to work hard and still enjoy ourselves (*e.g.* 'Independence Day', 'Star Wars', *etc.*, *etc.*).

The present cosmopolitan Lab 109 (Jeffrey Mphahlele, Ashraf Basuni, Keisuke Hino and Siew Lin Ngui); remnants of G18 (Ed Dornan, Ricky van Deursen, Richard Smith and Ania Owsianka); and the die-hards from Lab 303 (Kathryn Colman, Myles Marshall, Winnie Boner and Lesley Wallace) all deserve a special mention. Their help, advice, discussions, attitudes and general nonsense talk all made our working environment so enjoyable.

I am very thankful for the practical help of Adrian Abbots (pET expressions) and Dr. John McLauchlan (pp65 epitope tag) and for the helpful advice from my second supervisor, Dr. Chris Preston. Dr. John McLauchlan also read and corrected my results section very critically, for which I am especially grateful. Dr. A. Ackrill, Prof. W.H. Gerlich, Dr. B. Ferns, Prof. K. Murray and Dr. N. Naoumov all kindly contributed reagents used in this work. The Institute Washroom, Media, Store and Front Office staff have cheered me up on numerous occasions, and Jim and Paul should consider a double-act!

The last special mentions go to my family, my flat-mates, Mahmuda and Michayla. Thanks to my family for their help and support during all my time in Glasgow. My flat-mates have had to put up with my irregular hours, gloomy moods, absences, *etc.* (Andrew has now escaped to UAE as a consequence!). Mahmuda (or should I now say 'Dr. Yasmin') has been an incredible example during our time together and a real friend, putting up with me sitting beside her for three years and keeping us all cheerful. Michayla is the last on the list but deserves the biggest thanks: a) for all her sympathy and encouragement on the many occasions when experiments gave only negative results; b) for correcting and proof-reading this thesis; and c) for marrying me.

"Indeed, HBV might eventuate in a ***cyborg mutant****...."*

Yamamoto *et al.*, 1994.

The author was the recipient of a Wellcome Trust studentship. Unless otherwise stated, all of the work presented in this thesis was carried out by the author.

David Williams June 1997

SUMMARY

The hepatitis B virus nucleocapsid, or core particle, forms the major internal structural component of the infectious viral Dane particle. Core particles play essential roles in the morphogenesis of the virion and in reverse transcription of pregenomic RNA, one of the initial steps in genome replication. The 183 amino acid core protein is the sole component of the core particle. In a disulphide-bonded dimeric form, the core protein accumulates to reach a critical threshold concentration leading to the co-operative core particle assembly process. The core protein consists of two domains: the N-terminal 140 amino acids and the C-terminal 43 amino acids. The latter region can be removed without affecting particle assembly. This region also contains a high proportion of basic residues and is tethered in the particle interior, where it interacts with the viral genome. The core protein is also an important immunological target during HBV infection. It contains B- and T-cell epitopes throughout its length and the core-specific helper T-cell response is thought to be an important factor in resolution of infection.

During HBV infection subpopulations of viral variants arise with mutations in their nucleotide sequences. Variants with deletions in genomic regions are also detected, with deletions within the core gene being the most common. Determination of the functional ability of these variants was important as they are proposed to represent immune escape variants in patients and, as such, may be responsible for the chronic hepatitis observed.

To allow detection of core deletion proteins with the immunodominant B-cell epitope (amino acids 74-89) removed, a ten amino acid epitope tag was inserted at the C-terminus of all proteins. The tag sequence did not affect particle assembly by full-length protein and allowed detection of bacterially-expressed core deletion proteins Δ84-109, Δ81-121, Δ60-117 and Δ79-125 (numbers indicate the residues deleted). Although, after sucrose gradient centrifugation, core deletion proteins were present in regions of the gradient where particles were expected, no particulate structures were detectable by electron microscopy. It was concluded that these deletion proteins were incapable of stable particle assembly and, instead, only formed large non-specific aggregates. Similar results were obtained when samples used for electron microscopy were tested in an agarose gel assay for core particles. The inability of the four core deletion proteins to assemble into core particles was not due to a failure of the proteins to dimerise, which is the first step of the particle assembly pathway.

The deleted region of the Δ81-121 gene was replaced by an identical-sized region from the HBV surface gene and the resulting Δcore-surface hybrid protein purified by density gradient centrifugation. This protein was also unable to form stable particles, indicating that the central region of the core protein is important for the formation of core particles.

Step-wise replacement of the original core gene sequence to the Δ81-121 core gene deletion, from both the 5' and 3' ends of the deleted sequence, was carried out. This gave two sets of genes encoding three 'fill-in' proteins (5'F10, 20, 30 and 3'F10, 20, 30) with decreasing deletions (31 amino acids to 11 amino acids). However, all proteins were unable to form stable particles when purified as before.

Results published during this project showed that central deletions of the core protein disrupted a motif consisting of a heptad repeat of hydrophobic residues. These results showed that even small disturbances of, or within, this motif prevented core particle assembly.

As viral variants with core gene deletions are found co-existing with wild-type virus the possibility that both forms of protein could co-assemble to form novel mixed particles was examined. Density gradient purification of extracts from cells co-expressing full-length protein and the Δ84-109 or Δ81-121 proteins showed that both proteins were present in the same fractions, with a gradient profile similar to that of particles formed from full-length protein. These profiles were also different to those obtained for Δ84-109 and Δ81-121 proteins expressed in isolation. Therefore it seemed that mixed particles were able to assemble. In experiments where full-length and 5'F30 proteins were co-expressed similar results were not obtained, suggesting that mixed particles were unable to form, despite the smaller size of the deletion.

In mammalian cells, core protein shows a cell-cycle dependent cellular localisation. The protein is present in the nucleus during the G_1 and G_2 phases of the cell cycle and in the cytoplasm during S phase. Although the mechanism and function of this cell cycle-dependent localisation are unknown, a decision was made to investigate whether or not the core deletion proteins were subject to the same regulation. Using aphidicolin to arrest cells at the G_1/S boundary, the location of the core deletion proteins in transfected HepG2 cells was examined by immunofluorescence. All four core deletion proteins displayed a localisation which was both nuclear and cytoplasmic within the same cell. This altered pattern was probably due to the presence of the deletion upstream to the nuclear localisation signal in the C-terminus of the core protein. The significance of this result is unknown at present, but may be relevant to the disease state observed in these patients.

Sequencing of full-length core genes from viral DNA extracted from the serum of a chronic active hepatitis patient with core gene deletions (supplied by Dr. N. Naoumov) was carried out to examine whether or not the full-length viral subpopulation was responsible for the nature of the infection. The ten clones isolated showed a number of mutations throughout the pre-core and core regions, but these did not cluster in immunologically important epitopes.

The results from this project have shown that HBV variants with core gene deletions are not functionally viable by themselves. However, the interactions between core deletion proteins and full-length protein require further study, as does the

significance of the altered core deletion protein localisation. Results from these studies may explain the disease severity observed in the affected patients. Alternatively, further studies on the functional differences displayed by full-length core protein variants will yield more information on the mechanism of chronic hepatitis development. Studies on the structures of the core particle and the core protein published recently should also allow a more rational and directed approach to identifying regions of the core protein required for core particle assembly.

TABLE OF CONTENTS

ABBREVIATIONS

A:	adenine; absorbance
ALT:	alanine aminotransferase
APS:	ammonium persulphate
ASC:	asymptomatic carrier
ATP:	adenosine 5'-triphosphate
β-ME:	β-mercaptoethanol
bp:	base pairs
BSA:	bovine serum albumin
C:	cytosine; carboxy (-terminal end of protein)
CAH:	chronic active hepatitis
ccc:	covalently-closed circular
Ci:	Curie
CPH:	chronic persistent hepatitis
CsCl:	caesium chloride
Da:	Dalton
dATP:	2'-deoxyadenosine 5'-triphosphate
dCTP:	2'-deoxycytidine 5'-triphosphate
ddATP:	2'3'-dideoxyadenosine 5'-triphosphate
ddCTP:	2'3'-dideoxycytidine 5'-triphosphate
ddGTP:	2'3'-dideoxyguanosine 5'-triphosphate
ddTTP:	2'3'-dideoxythymidine 5'-triphosphate
dGTP:	2'-deoxyguanosine 5'-triphosphate
DHBV:	duck hepatitis B virus
DIP:	defective interfering particle
DNase:	deoxyribonuclease
dNTP:	2'-deoxynucleoside 5'-triphosphate
DTT:	dithiothreitol
dTTP:	2'-deoxythymidine 5'-triphosphate
ECL:	enhanced chemiluminescence
EDTA:	ethylenediaminetetra-acetic acid (disodium salt)
ER:	endoplasmic reticulum
EtBr:	ethidium bromide
FITC:	fluorescein isothiocyanate
gp:	glycoprotein
G:	guanine
GSHV:	ground squirrel hepatitis B virus

HBS:	HEPES-buffered saline
HBV:	hepatitis B virus
HCC:	hepatocellular carcinoma
HCMV:	human cytomegalovirus
HEPES:	N-2-hydroxyethyl piperazine-N'-2-ethanesulphonic acid
HLA:	human leukocyte antigen
HNF:	hepatocyte nuclear factor
HRP:	horse radish peroxidase
HSV:	herpes simplex virus type 1
IPTG:	isopropyl-β-D-thiogalactoside
kb:	kilobase pairs
LHBs:	large hepatitis B virus surface protein
mAb:	monoclonal antibody
MHBs:	middle hepatitis B virus surface protein
moi:	multiplicity of infection
N:	amino (-terminal end of protein)
NLS:	nuclear localisation signal
NP40:	Nonidet P40
OD:	optical density
ORF:	open reading frame
pAb:	polyclonal antiserum
PAGE:	polyacrylamide gel electrophoresis
PBS:	phosphate-buffered saline
PCR:	polymerase chain reaction
PEG:	polyethylene glycol
PEI:	polyethyleneimine
pfu:	plaque forming units
pg:	pre-genomic
pi:	post-infection
PMSF:	phenylmethylsulphonyl fluoride
RGB:	resolving gel buffer
RNase:	ribonuclease
rpm:	revolutions per minute
SDS:	sodium dodecyl sulphate
SGB:	stacking gel buffer
SHBs:	small hepatitis B virus surface protein
T:	thymine
TBS:	tris-buffered saline

TEMED:	N, N, N', N'-tetramethylethylenediamine
Tris:	tris (hydroxymethyl) aminomethane
Tween 20:	polyoxyethylene sorbitan monolaurate
UV:	ultraviolet
WHV:	woodchuck hepatitis B virus
X-gal:	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

ONE AND THREE LETTER AMINO ACID ABBREVIATIONS

Amino acid	Three letter code	One letter code	Codons
Alanine	Ala	A	GCU GCC GCA GCG
Arginine	Arg	R	AGA AGG CGU CGC CGA CGG
Asparagine	Asn	N	AAU AAC
Aspartic acid	Asp	D	GAU GAC
Cysteine	Cys	C	UGU UGC
Glutamine	Gln	Q	CAA CAG
Glutamic acid	Glu	E	GAA GAG
Glycine	Gly	G	GGU GGC GGA GGG
Histidine	His	H	CAU CAC
Isoleucine	Ile	I	AUU AUC AUA
Leucine	Leu	L	UUA UUG CUU CUC CUA CUG
Lysine	Lys	K	AAA AAG
Methionine	Met	M	AUG
Phenylalanine	Phe	F	UUU UUC
Proline	Pro	P	CCU CCC CCA CCG
Serine	Ser	S	AGU AGC UCU UCC UCA UCG
Threonine	Thr	T	ACU ACC ACA ACG
Tryptophan	Trp	W	UGG
Tyrosine	Tyr	Y	UAU UAC
Valine	Val	V	GUU GUC GUA GUG

LIST OF TABLES AND FIGURES

CHAPTER 1 INTRODUCTION

1A THE *HEPADNAVIRIDAE*

1A1 Hepatitis B Virus

1) Discovery and characterisation

The characterisation of the hepatitis B virus (HBV) followed the detection of the 'Australia antigen' (AuAg) in serum from an Australian aborigine and in patients with acute leukaemia (Blumberg *et al.*, 1965). Detection of the antigen in cases of viral and post-transfusion hepatitis led to the proposal that it was associated with the causative agent (Blumberg *et al.*, 1967; Prince, 1968; Bayer *et al.*, 1968; Giles *et al.*, 1969; Gocke and Kavey, 1969). This antigen was subsequently shown to be present in the serum of acute and chronic hepatitis B patients, as single or aggregated 20nm diameter particles (either spheres or long filaments). Electron microscopic studies (Dane *et al.*, 1970; Jokelainen *et al.*, 1970) using sera from AuAg-positive hepatitis patients, showed that these 20nm diameter spheres and filaments aggregated with larger 42nm diameter particles when incubated with antiserum. The 42nm particle, known as the Dane particle, was proposed to be the hepatitis B virion and was shown to have a greater density than the smaller particles, probably due to the presence of the viral DNA. The particle possessed a 7nm thick outer envelope which surrounded a 28nm icosahedral inner core (see Figure 1A1). Virions and the smaller particles aggregated due to the presence of a common protein (or proteins) in their outer layers and the 20nm particles were proposed to represent non-infectious particles composed of surplus virion envelope proteins.

In 1973, Kaplan *et al.* discovered an endogenous DNA polymerase activity associated with 28nm core particles of the virion, which was either released from virions spontaneously or by NP40 treatment. This provided further evidence that the Dane particle was the HBV virion and that the inner core particles probably represented viral nucleocapsids. Conclusive evidence was provided by Robinson and Greenman (1974), who showed that this polymerase activity could be immunoprecipitated by antisera specific for the viral surface protein (anti-HBs) but not by antisera specific for the core protein (anti-HBc) before NP40 treatment of purified virions and by anti-HBc but not anti-HBs antisera after NP40 treatment. The structure of the viral DNA was initially determined by electron microscopic studies to be exclusively circular double-stranded DNA, smaller in size than any other double-stranded DNA virus and probably with a very limited coding capacity (Robinson *et al.*, 1974). Using both molecular biological and electron microscopic studies the nature of the viral DNA was further elucidated and was shown to contain a single-stranded region, which was converted to a fully double-stranded region by the endogenous polymerase reaction (Summers *et al.*, 1975; Hruska *et*

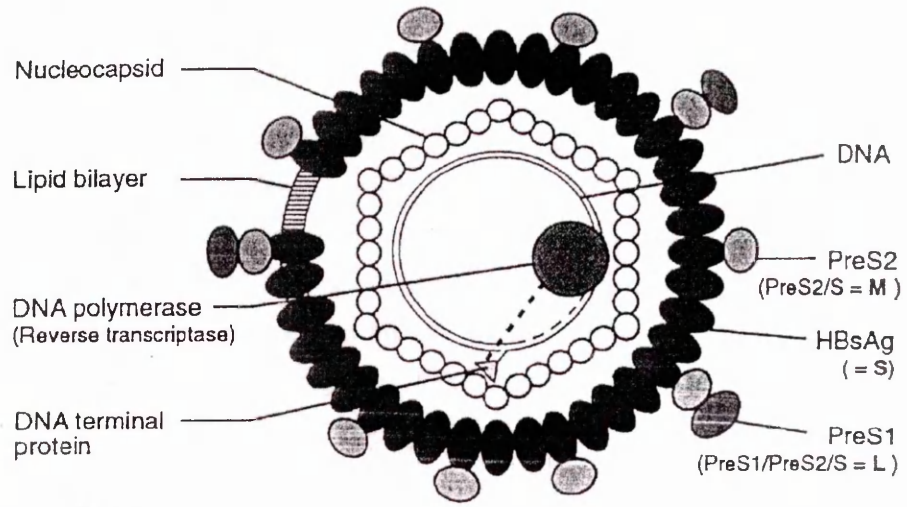


Figure 1A1: Structural organisation of the HBV virion
The infectious virion consists of an outer membrane containing the three surface gene products: SHBs (HBsAg), MHBs (PreS2+HBsAg) and LHBs (PreS1+preS2+HBsAg). The internal nucleocapsid, or core particle, consists entirely of core protein dimers and encloses the partially double-stranded circular DNA genome and the polymerase, covalently linked to the minus-strand DNA via its terminal protein domain. (Taken from Nassal and Schaller, 1993)

al., 1977). By convention, the full length strand, being complementary to the viral transcripts, is termed the minus-strand and the shorter strand is termed the plus-strand. This single-stranded region was observed in >99% of genomes examined and, although variable in size, had a favoured minimum length of 650-700 nucleotides (Delius *et al.*, 1983). In contrast to its 3' end, the 5' end of the plus-strand was fixed in position with respect to a unique *Eco* RI site in the genome and a nick present in the minus-strand. The partial duplex nature of the genome was maintained by a cohesive region of approximately 210-310bp at the 5' termini of both DNA strands (Sattler and Robinson, 1979; Siddiqui *et al.*, 1979). In addition to differing in length, the two DNA strands differed in the nature of their 5' termini. The minus-strand was shown to have a protein bound to its 5' terminus, explaining the extraction of viral DNA to the organic phase upon phenol extraction (Gerlich and Robinson, 1980). After treatment of the protein:DNA complex with 0.1M NaOH or by heating with SDS to 90°C this attachment remained, suggesting a covalent linkage. Both strands of the DNA were unable to be phosphorylated by polynucleotide kinase, the minus-strand because of the attached protein. However, no protein was detected at the 5' terminus of the plus-strand. This puzzle was solved when it was shown that a small capped oligoribonucleotide approximately 17 nucleotides long

was attached to this terminus (Will *et al.*, 1987). These two different attachments were later shown to result from the viral replication mechanism (See section 1B2).

2) Genomic Organisation

The cloning of Dane particle-extracted HBV DNA allowed complete determination of the nucleotide sequence (Galibert *et al.*, 1979; Pasek *et al.*, 1979; Ono *et al.*, 1983; see Figure 1A2). Although the plus-strand was predicted to contain several small open reading frames (ORFs), no protein products have ever been detected. The minus-strand was shown to have four major ORFs, one of which covered approximately 80% of the genome. The other ORFs had multiple in-frame start codons and were presumed to code for more than one protein.

ORF S was identified as coding for the surface proteins of the virus, present in the virion envelope. This was achieved by comparison of both the molecular weight and partial amino acid sequence of the surface protein with the predicted translated nucleotide sequence of the ORF. The gene had three in-frame start codons and a single stop codon, with the sequences upstream of the small surface gene named pre-S1 and pre-S2.

The core, or C, ORF was identified by the similarity in molecular weights of the core protein and the calculated molecular weight of the ORF C product. When this genomic region was expressed in *E.coli* and injected into rabbits, the sera reacted in immunodiffusion assays with human HBcAg isolated from human liver (Pasek *et al.*, 1979; Burrell *et al.*, 1979).

The largest open reading frame, ORF P, gave a protein of predicted molecular weight of approximately 90kDa. This was within the correct molecular weight range expected for a DNA polymerase or reverse transcriptase, thus putatively identifying this gene as coding for the viral polymerase.

The smallest open reading frame was designated ORF X, with a predicted size of 17kDa. No viral protein of this size has yet been identified in infected cells. This may be due to different sized proteins being produced from this ORF (Kwee *et al.*, 1992). The precise role of ORF X and its possible related products in the viral lifecycle is still unclear, but a number of functions have been suggested (see sections 1B3.3 and 1B4.2).

All four ORFs overlap to some extent: the surface gene is contained entirely within the region of the polymerase gene, although is in a different reading frame. The regulatory elements for transcription are also present within coding sequences, allowing further genomic compactness.

Two repeats of an 11bp sequence, termed DR1 (nucleotides 1590-1600) and DR2, (nucleotides 1824-1834) were identified at the ends of the viral genome. These direct repeats have significance for both the viral replicative mechanism and in specific integration of viral DNA into chromosomal DNA in hepatocellular carcinoma (Dejean *et al.*, 1984).

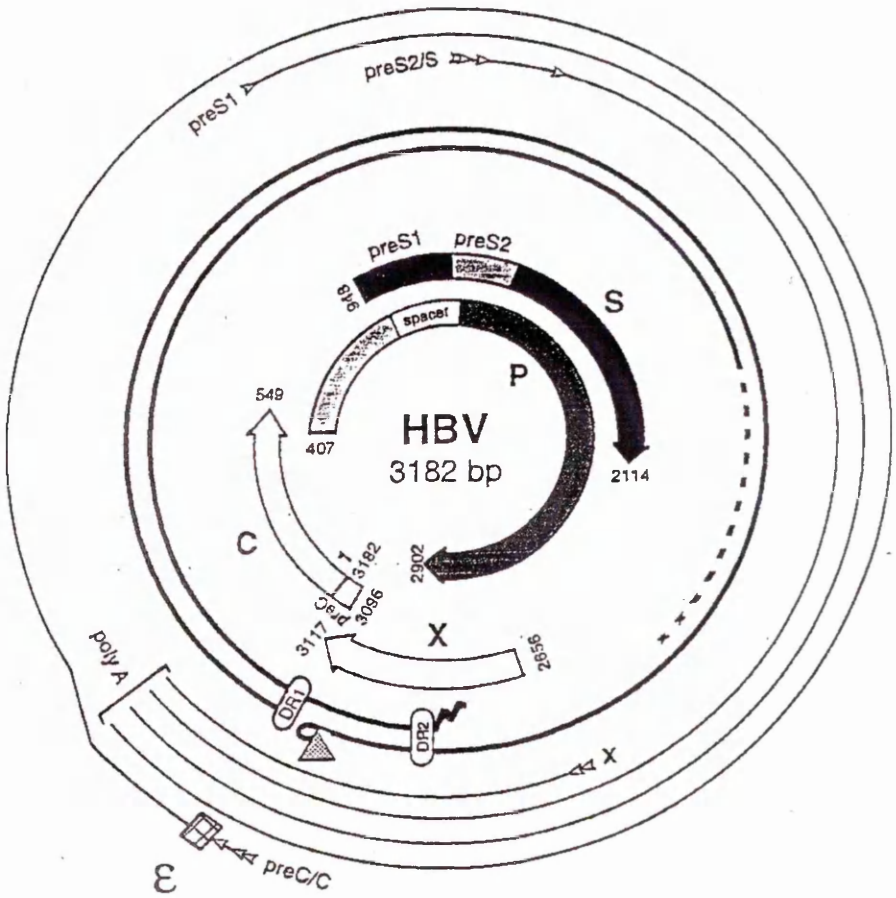


Figure 1A2: Genomic organisation of HBV

The outer lines represent the different classes of transcripts and the bold inner circles the DNA genome, as found in the virion. The oligoribonucleotide primer found at the 5' end of the plus-strand DNA is shown as a wavy line and the polymerase/terminal protein at the minus-strand terminus as a triangle. Encapsidation signal (ε) sequences are represented by a hatched rectangle. The four major ORFs (C, S, X and P) are indicated in the centre. The numbering system shown in this diagram differs from that used in the rest of this thesis, which follows the system of Ono *et al.*, (1983) where position 1 is identified as the first T nucleotide in the unique *Eco* RI site, at the 5' end of the pre-S2 region (nucleotide 1322 in the above diagram). (Taken from Nassal and Schaller, 1993)

1A2 Woodchuck Hepatitis Virus

The woodchuck hepatitis virus was the second hepadnavirus discovered shortly after the isolation of HBV. The existence of this virus was suggested by the high incidence of severe chronic hepatitis and primary hepatocellular carcinoma in a colony of woodchucks

in Penrose Research Facility at Philadelphia Zoo (Summers *et al.*, 1978). The serum of these animals contained large amounts of excess viral envelope proteins in the form of approximately 25nm diameter spheres or filamentous particles, similar to those described in humans. These particles displayed weak serological cross-reactivity with the analogous HBsAg particles, whilst core particles isolated from livers in the two species showed higher levels of cross-reactivity (Werner *et al.*, 1979). Many double-layered 55nm particles were also detected which contained DNA and a DNA polymerase activity analogous to that of the human virion. These particles were presumed to be woodchuck hepatitis B virus (WHV) virions. The viral genome showed significant cross-hybridisation to the HBV genome. Sequencing of full-length cloned genomes showed that the number and sizes of the open reading frames, along with their location relative to the nick in the minus-strand, were also similar to the HBV genome structure. Comparison of HBV and WHV nucleotide sequences showed 62-70% homology. The least homologous region was that surrounding the unique *Eco* RI site, corresponding to sequences common to the pre-S region and the polymerase gene (Galibert *et al.*, 1982). A region corresponding to the 3' end of the X gene also showed a lower homology. Viral genomes were predominantly found in the liver, although smaller quantities of viral DNA and RNA were detected in peripheral blood leukocytes, spleen and kidney (Summers *et al.*, 1978; Korba *et al.*, 1986).

1A3 Ground Squirrel Hepatitis Virus

A virus similar to HBV was isolated from seemingly healthy Beechey ground squirrels (Marion *et al.*, 1980). The virus showed marked hepatotropism and a narrow host range, with rats, mice, guinea pigs and hamsters showing no detectable antigenemia after parenteral administration of the virus (Ganem *et al.*, 1982a). Despite high serum titres of viral surface proteins and 47nm virions, no pathology was observed in infected ground squirrels (Marion *et al.*, 1980; Ganem *et al.*, 1982a). Using hybridisation studies, physical mapping and complete genome sequencing, the genome showed extensive homology to both HBV (55% nucleotide homology) and WHV (82% nucleotide homology). The number and arrangement of the ORFs was also very similar to both HBV and WHV (Siddiqui *et al.*, 1981; Ganem *et al.*, 1982b; Seeger *et al.*, 1984). The surface and core proteins were shown to be cross-reactive with and biochemically related to analogous HBV and WHV proteins (Cote and Gerin, 1983; Feitelson *et al.*, 1982).

Similar hepadnaviruses have also been isolated from Arctic ground squirrels (Testut *et al.*, 1996) and gibbons (Norder *et al.*, 1996), although these are less well characterised.

1A4 Avian Hepadnaviruses

HBV-like viruses have been discovered in Pekin Ducks, in some instances only one day after birth, suggesting viral infection and multiplication in embryonic eggs (Mason *et*

al., 1980). Virions of the same size and appearance as HBV were detected in sera from these birds, which preferentially localised to the liver.

Determination of the nucleotide sequence of the viral DNA by Mandart *et al.* (1984) showed the genome to be 3021 nucleotides in length, which is comparable in size to HBV. However, compared with HBV, a lower homology was shown (<40% for the largest part of the sequence). Also, the genome of duck hepatitis B virus (DHBV) has a different organisation, lacking a homologue of the mammalian hepadnaviral X gene. Instead only a larger, novel core gene was present with C-terminal basic amino acid repeats which are similar to those found in the mammalian hepadnaviral core proteins. The surface gene homologue was found to be approximately fifty amino acids shorter than the HBV protein, due to the deletion of a region corresponding to amino acids 105-155. The 40nm virion particle co-migrated on CsCl gradients with a DNA polymerase activity (Mason *et al.*, 1980), suggesting that DHBV also possessed an endogenous DNA polymerase. Amino acid sequence comparisons showed the predicted polymerase protein was related to the equivalent protein in HBV and WHV (Mandart *et al.*, 1984).

Another recently characterised avian hepadnavirus, heron HBV (HHBV), was isolated from grey herons in Germany. HHBV showed similar virion morphology, genome size and core and surface protein sizes when compared to DHBV (Sprengel *et al.*, 1988). Sequencing of the genome showed a high homology to DHBV (78.5%) and a similar organisation, with the mammalian hepadnaviral X gene missing and a larger core gene present instead. The differing genomic organisation and lower sequence homology for avian hepadnaviruses, compared to the mammalian viruses, suggests that these viruses form a distinct group, but still with specific host ranges, as HHBV was found to be non-infectious to Pekin Ducks (Sprengel *et al.*, 1988).

1B VIRAL LIFECYCLE

1B1 Hepatocyte Binding

To initiate infection, a virus must attach to a host cell receptor via proteins present on the virion surface (see Figure 1B1). The exact identity of these proteins and the mechanism of virus attachment and subsequent penetration into human hepatocytes in the lifecycle of HBV are unknown. The large, middle and small surface proteins (LHBs, MHBs and SHBs, respectively (see Figure 1A1 and Section 1B4.4)) are all present in the virion but it is unclear as to which particular region/s of these proteins are important for hepatocyte binding. However, candidate host cell proteins which bind to all three virion surface proteins have been identified.

The pre-S1 region of LHBs has been reported to bind to hepatocyte plasma membranes through specific interactions involving residues between amino acids 21-47

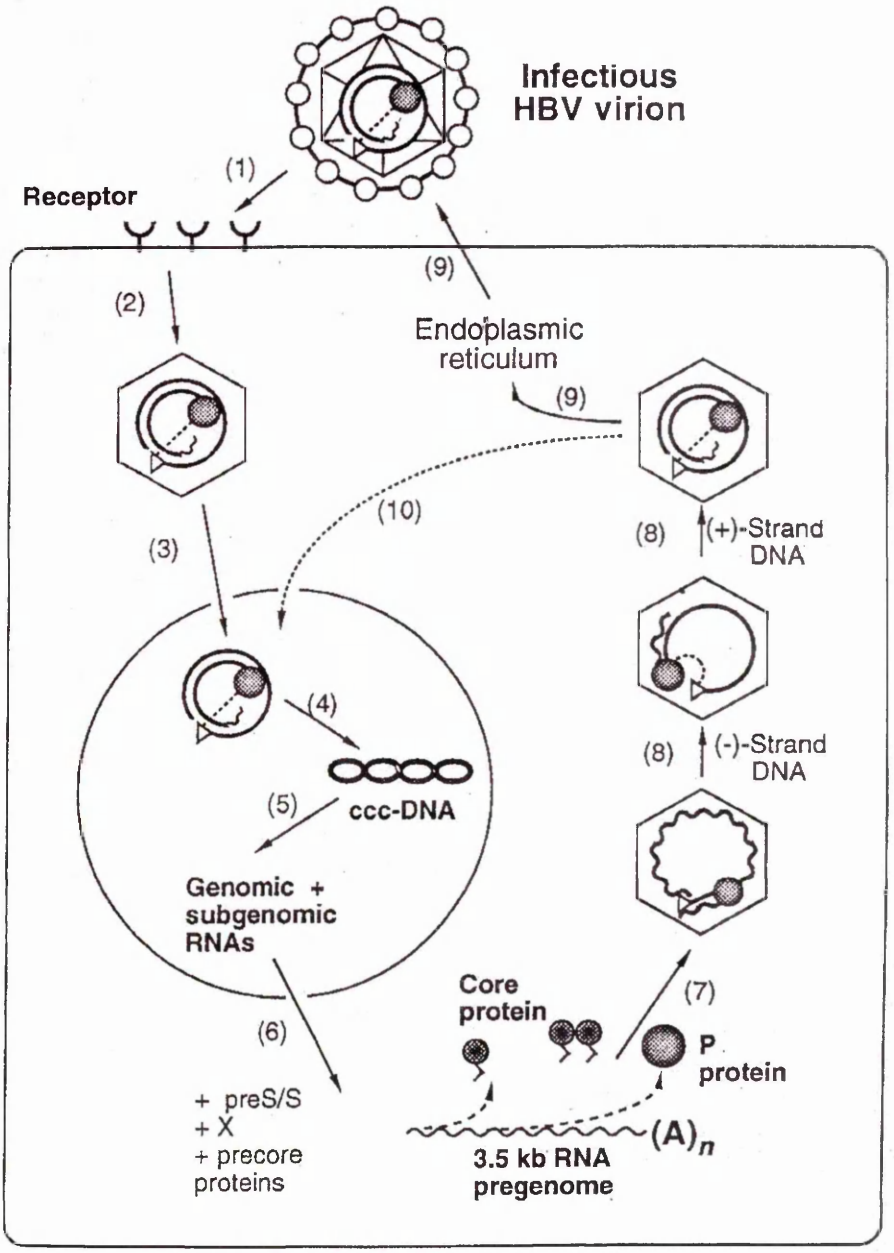


Figure 1B1: Basic life-cycle of HBV

(1) Attachment to and (2) penetration of the host cell. (3) Uncoating and transport of the viral genome to the nucleus. (4) Conversion into covalently-closed circular (ccc) DNA as the template for transcription. (5) RNA synthesis and transport to the cytoplasm. (6) Translation of the gene products. (7) Nucleocapsid assembly and encapsidation of pgRNA. (8) Reverse transcription of pgRNA and DNA synthesis. (9) Export from the cell as enveloped virions. Alternatively, the nucleic acid can be reimported to the nucleus for amplification of cccDNA (10). (Taken from Nassal and Schaller, 1993)

(Neurath *et al.*, 1986; Pontisso *et al.*, 1989a; Petit *et al.*, 1991). Interacting host cell membrane proteins have been detected which may have a role in the initial stages of viral infection. A region of the IgA α-1 chain, with limited amino acid similarity to pre-S1 amino acids 21-47, competed with pre-S1 residues for liver plasma membrane binding, suggesting the presence of a common receptor (Pontisso *et al.*, 1992). The use of anti-idiotypic antibodies to pre-S1 residues identified several liver cell membrane proteins which may function as HBV receptors (Petit *et al.*, 1992). The liver membrane asialoglycoprotein receptor was also shown to bind virions, with binding completely inhibited by anti-pre-S1, but not anti-pre-S2 monoclonal antibodies (Treichel *et al.*, 1994).

Pre-S2 residues do not mediate binding to hepatocyte membranes by themselves, but binding was possible in the presence of polymerised human serum albumin (pHSA), which also binds liver cell membranes (Pontisso *et al.*, 1989b). However conflicting reports occur as to the relevance of pre-S2:pHSA-mediated hepatocyte binding: pre-S2 residues corresponding to the pHSA binding site are inaccessible to monoclonal antibodies when present in virions (Petit *et al.*, 1987) and the occurrence of gluteraldehyde-linked pHSA, the most efficient polymerised form for binding (Yu *et al.*, 1985), *in vivo* seems unlikely. In contrast to these findings which cast doubt on the relevance of pre-S2-mediated hepatocyte binding, pre-S2 residues were shown to be capable of generating neutralising and protecting antibodies (Neurath *et al.*, 1986; Itoh *et al.*, 1983).

Although early reports suggested SHBs was unable to bind to hepatocytes, or that these interactions were irrelevant, recent experiments by the group of Yap have shown this may not be the case. SHBs particles (see Section 1B4.4) bound to intact human hepatocytes (Leenders *et al.*, 1990) through specific interactions with endonexin II (annexin V), a 34kDa member of Ca^{2+}-dependent phospholipid binding proteins (Hertogs *et al.*, 1993). The production of anti-idiotypic anti-HBs antibodies in rabbits immunised with human liver endonexin II gave evidence for the ligand:receptor nature of this interaction (Hertogs *et al.*, 1994).

In order to reconcile these seemingly conflicting results, it is not unreasonable to suggest that more than one surface/pre-S protein region is required for efficient organ- and species-specific virus attachment. This process may involve a number of steps, each utilising a different surface protein region (Leenders *et al.*, 1990). SHBs protein may be a more likely candidate for hepatocyte binding, with pre-S regions providing species-specific interactions due to their low cross-species homology. Alternatively, virus attachment and entry into cells may occur by a non-specific process, despite the observation of these specific interactions between surface proteins and hepatocyte membrane proteins. If this was the case, though, the efficiency would have to be very high to allow successful propagation of a small inoculum of virus. Also, another

explanation would be required to explain the organ- and species-specificity of the hepadnaviruses.

The entry of enveloped viruses into cells often requires proteolysis of the envelope proteins to expose a fusion domain. This domain enables the fusion of viral membranes with either host cell membranes or endosome membranes if virions are endocytosed. In studies with staphylococcal V8 protease-treated virions, Lu *et al.* (1996) showed that at acid pH, virus particles were able to infect normally non-susceptible HepG2 cells. The V8 protease treatment exposed a putative consensus fusion domain between amino acids 182-188 of pre-S1 protein: FLGXLLV. The pre-S1 residues required for attachment, as described before, did not seem to be necessary for these experiments; V8 protease treatment removes essentially all pre-S residues from the virion. Surface residues were also the only requirement for entry of SHBs particles into primary human hepatocytes via endonexin II attachments (de Bruin *et al.*, 1995). After incubation with gold-conjugated SHBs particles at 4°C, cells warmed to 37°C showed invagination of plasma membrane coated pit regions at SHBs binding sites and formation of endocytotic vesicles. With the different mechanisms for viral entry inferred by both sets of results, fusion peptide-mediated and receptor-mediated endocytosis respectively, the problem remains unsolved.

Less still is known about the processes that occur following virion entry into host cells. The precise mechanism allowing delivery of the viral genome to the nucleus remains undetermined, although clues have come from an *in vitro* model derived from WHV (Kann *et al.*, 1997). In this model the polymerase:DNA complex was efficiently transported into the nucleus after release from core particles. These possibly disassemble before reaching the nuclear membrane (Bock *et al.*, 1996), as core particles themselves do not cross intact nuclear membranes (Guidotti *et al.*, 1994a). In addition, phosphorylation of core protein C-terminal serine residues by encapsidated protein kinase C (Kann *et al.*, 1993) may act to destabilise the core particle and allow virion genome release (Kann and Gerlich, 1994).

1B2 Viral Replication

The unusual structure of the HBV genome is a consequence of the mode of replication employed by the virus. HBV utilises a reverse transcription step to synthesise minus-strand DNA, in contrast to almost all other DNA viruses (Summers and Mason, 1982). HBV replication occurs in four main stages, which are listed and described below:

1) formation of covalently-closed circular DNA (cccDNA)
2) production of pregenomic RNA (pgRNA)
3) minus-strand DNA synthesis, by reverse transcription
4) plus-strand DNA synthesis

Stages 1 and 2 occur in the host cell nucleus, whereas stages 3 and 4 occur after the

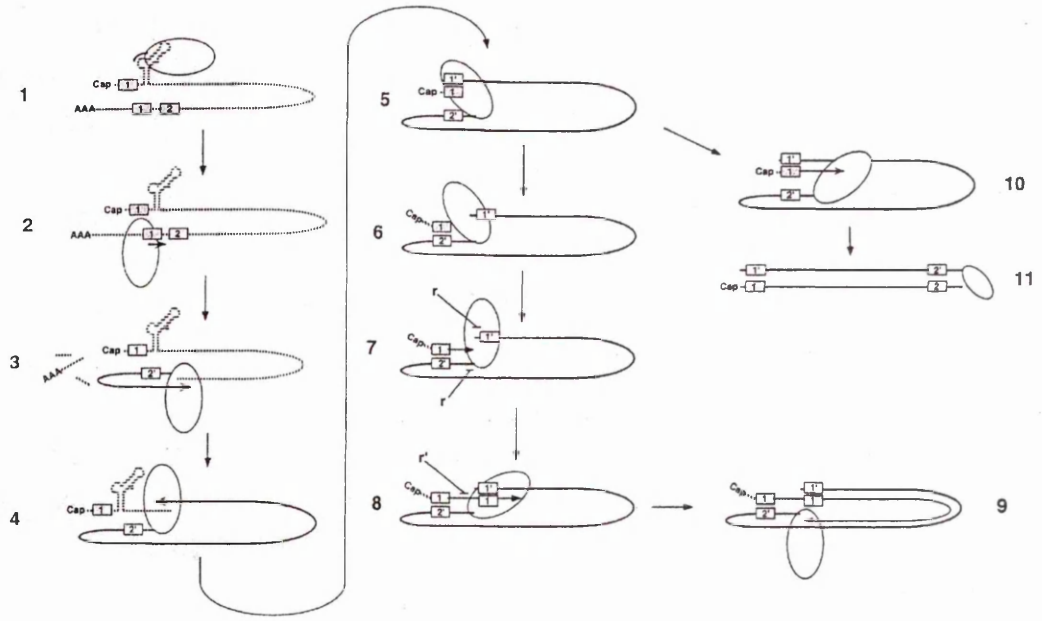


Figure 1B2: Model for HBV replication

(1) Encapsidation and initiation of minus-strand DNA synthesis. The dotted line represents pgRNA; boxes 1 and 2 represent DR1 and DR2, respectively. The polymerase (oval shape) binds the 5' ε stem loop and facilitates pgRNA encapsidation. Four nucleotides of minus-strand DNA are synthesised (solid line) from the bulge region. **(2)** Minus-strand template switch. The polymerase-tetranucleotide complex switches to complementary sequences in the 3' copy of DR1 and minus-strand DNA synthesis continues. **(3** and **4)** Synthesis of minus-strand DNA. As minus-strand DNA (solid line) is synthesised, pgRNA is degraded by the RNase H activity of the polymerase. **(5)** Completion of minus-strand DNA synthesis and generation of plus-strand RNA primer. The final RNase H cleavage generates the plus-strand primer, which is derived from the first 18 nucleotides of the pgRNA (from cap structure to 3' of DR1). **(6)** Translocation of the plus-strand primer. The primer anneals to complementary DNA sequences in DR2, which is juxtaposed to DR1 upon completion of minus-strand synthesis. **(7)** Initiation and elongation of plus-strand DNA synthesis. Approximately 50 nucleotides of plus-strand DNA is synthesised. The minus-strand DNA terminal redundancy is labelled 'r'. **(8)** Intra-strand template switch and elongation of plus-strand DNA. Plus-strand DNA synthesis switches templates from the 5' to 3' end of minus-strand DNA, facilitated by the terminal redundancy 'r'. Plus-strand DNA elongation continues. **(9)** Elongation and completion of plus-strand DNA synthesis results in a relaxed circular DNA. **(10** and **11)** *In situ* priming of plus-strand DNA synthesis. A fraction of plus-strands may not initiate at DR2, but at DR1 instead, due to plus-strand primer not being translocated. Elongation results in duplex, linear DNA. (Taken from Loeb *et al.*, 1997)

production of pgRNA, which is complexed to the viral polymerase and is encapsidated into an assembling core particle (see Section 1B5). A summary diagram of HBV replication is shown in Figure 1B2.

1) Formation of cccDNA

Several modifications to both strands of the virion DNA are required before it can be used as a transcriptional template: the protein must be dissociated from the 5' end of the minus-strand; eight terminally redundant nucleotides removed from the 5' end of the minus-strand and the 5' and 3' ends ligated; the 17-mer oligoribonucleotide must be removed from the 5' end of the plus-strand and plus-strand synthesis completed. The resulting circular supercoiled molecule is termed cccDNA. Host cell enzymes are thought to be responsible for all these stages, but this has only been demonstrated experimentally for the last stage mentioned (Köck and Schlicht, 1993). Amplification of cccDNA occurs through an intracellular pathway and this form of DNA accumulates in the nucleus (Miller and Robinson, 1984; Tuttleman *et al.*, 1986).

2) Production of pregenomic RNA

The 3.5kb pregenomic RNA (pgRNA) represents one of the major polyadenylated transcripts found in infected hepatocytes (Cattaneo *et al.*, 1984). It is transcribed from one region of the core promoter by host cell RNA polymerase II and, in addition to its role in viral replication, also acts as the template for translation of core and polymerase proteins. The 5' end of this transcript was mapped to the pre-core region nucleotide 1818, five nucleotides downstream of the pre-core ATG; this explains why this transcript does not code for e antigen. The transcript terminates at the common site for all hepadnaviral mRNAs, approximately twenty nucleotides downstream of the conserved hexanucleotide UAUAAA. This transcription termination signal is read-through at the first pass, leaving the pgRNA with a terminal redundancy of approximately 120 nucleotides (Will *et al.*, 1987). An 85 nucleotide sequence at the 5' end of pgRNA was identified as being necessary for efficient encapsidation of pgRNA into core particles, following interaction of the RNA with the viral polymerase (Junker-Niepmann *et al.*, 1990; see Section 1B5). This *cis*-acting encapsidation signal was termed ε and contained several inverted repeat sequences suggesting a high degree of secondary structure. These repeat sequences were similar in all hepadnaviruses. The secondary structure of ε was determined by RNase and nuclease probing and shown to be a bipartite stem-loop structure, with lower and upper stems interrupted by a six nucleotide bulge. The upper stem contained a single unpaired U and a six nucleotide loop (Knaus and Nassal, 1993; Pollack and Ganem, 1993; see Figure 1D2). Due to the terminal redundancy, two copies of ε and the DR1 repeat are present in pgRNA.

3) Minus-strand DNA synthesis by reverse transcription

Recently the ε encapsidation signal has also been shown to be involved in the initial stages of minus-strand DNA synthesis. From the results of primer extension analysis it was previously believed that minus-strand DNA synthesis initiated at the 3' copy of DR1 (DR1*) and proceeded in a continuous manner, without the need of a template switch (Will *et al.*, 1987). However, experiments with DHBV polymerase expressed in a yeast TY1 retrotransposon system (Tavis and Ganem, 1993) showed that the minus-strand origin mapped within the stem-loop of ε. This origin had been previously described as a cryptic site which was used when DR1* was deleted from pgRNA (Condreay *et al.*, 1992). As the 3' copy of ε and the 5' copy of DR1 were functionally silenced in pgRNA (Seeger and Maragos, 1990), the hypothesised mechanism for minus strand synthesis involved a template switch by nascent minus-strand DNA from the 5' copy of ε to DR1*. This mechanism has now been demonstrated for both DHBV (Wang and Seeger, 1993; Tavis *et al.*, 1994) and HBV (Rieger and Nassal, 1996). The replication origin was more precisely mapped to the six nucleotide bulge of ε and particularly to the four distal nucleotides UUCA (Nassal and Rieger, 1996).

Evidence from studies on DHBV replication indicated that reverse transcription was initiated via a mechanism unique to hepadnaviruses: the reaction was primed by a protein molecule which remains covalently bound to the 5' end of minus-strand DNA throughout the viral lifecycle (Molnar-Kimber *et al.*, 1983). The initiation step involved the formation of a phosphodiester bond between the Tyr-96 residue in the terminal protein domain of the polymerase and, in the case of DHBV, a dGTP nucleotide (Bartenschlager and Schaller, 1988; Zoulim and Seeger, 1994; Weber *et al.*, 1994). This nucleotide was specified by the 3' terminal nucleotide in the bulge of the ε sequence. The use of HBV polymerase expressed in insect cells has confirmed a similar mechanism for this stage of HBV replication. HBV polymerase was expressed in the presence of RNA containing ε and DR1 sequences to examine HBV minus-strand DNA initiation. The minus-strand DNA synthesised initiated with a dTTP nucleotide (Lanford *et al.*, 1995), complementary to the 3' terminal nucleotide in the bulge of the HBV ε sequence. However, minus-strand transfer to the DR1 sequence was found to be more efficient than for DHBV. Another difference from DHBV replication was that Tyr-63 was the site of the DNA:polymerase linkage (Lanford *et al.*, 1997).

After the addition of four nucleotides complementary to the bulge sequence, primer elongation halts, possibly because the polymerase encounters the structurally rigid lower stem of ε. The polymerase-primer complex then translocates to a complementary tetranucleotide sequence in DR1* and minus-strand synthesis is completed in a continuous manner as initially hypothesised (Will *et al.*, 1987). The sequence complementarity only partially explains the specificity of the translocation reaction, as the tetranucleotide sequence occurs eighteen times in a unit-length genome and sequences

complementary by only two nucleotides are capable of translocating efficiently (Nassal and Rieger, 1996). In addition, when multiple complementary tetranucleotide sequences were introduced around DR1*, the original sequence was still used preferentially (Loeb and Tian, 1995). The conformation of the pgRNA may place the 5' ε and DR1* sequences in close spatial proximity, presumably by adopting a circularised shape, bringing the ends of the pregenome together. This conformation may be stabilised by RNA:protein interactions, which could possibly involve additional polymerase binding sites (Seifer and Standring, 1995) or cellular proteins such as the 65kDa nuclear protein which binds pyrimidine-rich regions present at both ends of pgRNA in the terminally redundant sequences (Perri and Ganem, 1996). Kidd and Kidd-Ljunggren (1996) carried out computer-based modelling studies of the folding of the pgRNA template and predicted that the 3' end of pgRNA contained an RNA superstructure with DR1* exposed on a loop. This superstructure is not present at the 5' end of pgRNA and so may direct correct positioning of the translocating polymerase-primer complex. As the tetranucleotide translocation site for this complex in DR1* is located eight nucleotides downstream of the 5' end of the pgRNA, this results in a terminal redundancy of eight nucleotides in the completed minus-strand DNA.

Digestion of the pgRNA occurs simultaneously with reverse transcription and is also catalysed by the viral polymerase. Mutational analysis of the polymerase protein showed that RNase H activity mapped to the C-terminal region, between amino acids 680-832 (Radziwill *et al.*, 1990).

4) Plus-strand synthesis

The 5' end of plus-strand DNA was found to contain a covalently-linked seventeen nucleotide long oligoribonucleotide with sequence identical to both DR1 and nucleotides 5' of DR1 (Will *et al.*, 1987; Seeger *et al.*, 1986). This oligoribonucleotide was capped in the DHBV genome (Lien *et al.*, 1986) and was proposed to be the undigested 5' end of the pgRNA. This would have been translocated to the DR2 sequence in the newly-synthesised minus-strand DNA, which is where the 5' end of plus-strand DNA was mapped. Again, the translocation mechanism is unknown, but the spatial conformation of the minus-strand is probably critical for correct positioning of the translocated oligoribonucleotide.

Oligoribonucleotide-primed plus-strand synthesis proceeds up to the 3' end of the minus-strand template after which an intra-molecular template switch must take place. This switch is facilitated by the eight nucleotide redundancy in the 5' and 3' ends of minus-strand DNA, but the exact mechanism remains to be determined. Recent studies on DHBV replication have suggested that factors other than the terminal redundancy are required for this template switch (Loeb *et al.*, 1997).

Plus-strand synthesis halts before completion resulting in the circular partially double-stranded genome observed in virions. The reasons for this are unclear, but possibilities include spatial constraints within the nucleocapsid or an insufficient pool of available dNTPs after the mature virion is released into the extracellular medium.

1B3 HBV Gene Expression

The subject of HBV gene expression results in the production of a large amount of literature, some of which can be difficult to integrate and may be contradictory. This section gives only a basic summary of the main points; further, more detailed information is available in recent reviews by both Schaller and Fischer (1991) and Yen (1993). The positions and lengths of the transcripts are represented in Figure 1A2, where the pgRNA is labelled 'C'.

1B3.1 Transcripts

Viral messages are transcribed from the episomal cccDNA template (see Section 1B2(1)) by the host cell transcriptional apparatus. Four classes of transcripts have been detected: 3.5kb, 2.4kb, 2.1kb and 0.9kb. The transcripts are initiated from four viral promoters: the core promoter, the surface promoters SpI and SpII and the X promoter. All transcripts terminate downstream of the single polyadenylation site, UAUAAA, which is located after the core gene AUG start codon.

The core promoter gives rise to three different 3.5kb transcripts (Yaginuma *et al.*, 1987; Honigwachs *et al.*, 1989). Two of these are the long pre-core mRNAs, initiated 20-30 nucleotides upstream of the pre-core ATG at 1783/4 and 1790±1 and which encode e antigen. The third is a shorter, more abundant core mRNA, or pgRNA. This is initiated at 1819±1 and can be reverse transcribed or used as message for core and polymerase proteins. This promoter contains binding sites for the liver-enriched factors HNF3 and HNF4 (Johnson *et al.*, 1995; Guo *et al.*, 1993; Raney *et al.*, 1997). However, a recent report by Yu and Mertz (1996) has shown that synthesis of pgRNA and pre-core mRNAs was actually under the control of two discrete promoters which, although partially overlapping, were genetically separable. The activities of the two promoters were also shown to be differentially regulated, which may explain the differing relative abundance of the pgRNA and pre-core transcripts in infected liver cells.

As these RNAs are all of supergenomic length, the polyadenylation signal must be ignored when first encountered by the RNA polymerase. Efficient recognition of the downstream polyadenylation signal, in the case of ground squirrel hepatitis virus (GSHV) at least, requires three sequences located in the 397 nucleotides upstream of the AAUAAA sequence (Russnak and Ganem, 1990). The inefficient usage of the signal on the first pass is a result of pgRNA containing only one of these regions between its 5' end and the

polyadenylation signal.

The SpI promoter directs the synthesis of a single 2.4kb transcript (the pre-S1 transcript) which is initiated at nucleotide 2807 and encodes LHBs (Will *et al.*, 1987). This is the only HBV promoter to contain an upstream TATA box, which explains the lack of heterogeneity of the 5' end of LHBs-encoding transcripts. The SpI promoter shows a strong preference for differentiated hepatocytes and contains binding sites for HNF1 and HNF3 (Chang *et al.*, 1989; Raney *et al.*, 1995).

The SpII promoter directs the synthesis of a number of mRNAs with start sites on either side of the pre-S2 ATG (Cattaneo *et al.*, 1983, 1984; Standring *et al.*, 1984; Yaginuma *et al.*, 1987). Those mRNAs which initiate 5' to this ATG encode MHBs, while those initiating 3' to this ATG encode SHBs. Thus the SpII promoter is responsible for the production of both MHBs and SHBs proteins.

The X promoter directs the synthesis of viral mRNAs which encode the X protein. These transcripts have heterogeneous 5' termini and are approximately 0.7-0.9kb in length (Treinin and Laub, 1987). However, these transcripts have not been detected in HBV-infected liver cells, although the X protein has been detected in liver biopsies (Haruna *et al.*, 1991; Wang *et al.*, 1991a) and sera (Feitelson and Clayton, 1990) from HBV-infected patients.

The 3.5kb pgRNA transcript undergoes splicing to produce two separate mRNAs, approximately 2.2kb in size, which have been detected in infected liver cells and in HepG2 cells transfected with genomic constructs (Wu *et al.*, 1991a). The 2.2kb mRNAs are either singly or doubly spliced: singly spliced transcripts have 1224 nucleotides removed and are predicted to encode a core protein lacking its terminal cysteine, an N-terminally truncated polymerase and smaller surface proteins. Doubly spliced transcripts contain a 282 nucleotide in-frame deletion from the middle of the core gene and a 1016 nucleotide deletion from the terminal codon of the core gene. These splicing events would remove the C-terminal cysteine of the core protein and the N-terminus of SHBs, also producing a truncated polymerase. A 2.7kb spliced RNA has also been detected in transfected HepG2 cells with the same splice acceptor site as the 2.2kb RNA, but a splice donor site 500bp downstream of the 2.2kb donor site (Suzuki *et al.*, 1990). The relevance of these splicing events is unclear as they seem to play no major role in the viral lifecycle, as determined by mutational inactivation of the donor and acceptor sites (Suzuki *et al.*, 1990). However as they seem to be more prominent in patients who have progressed from acute to chronic hepatitis, as compared to patients who recover from acute hepatitis, a role has been postulated for them in the cause of chronic HBV infection (Rosmorduc *et al.*, 1995).

1B3.2 Transcriptional regulation by *cis*-acting elements

1) Enhancers

In addition to the four promoters and the polyadenylation site, the HBV genome also contains two enhancer elements, I and II, which regulate promoter activities. Enhancer I (nucleotides 1074 and 1234) is located in the upstream region of the X promoter, and can function in an orientation-independent manner (Shaul *et al.*, 1985). Enhancer I sequences were capable of increasing the rate of transcription from the core and X promoters, but had little effect on the SpII promoter (Zhang *et al.*, 1992; Raney *et al.*, 1989). The mechanism of action of enhancer I is likely to contribute to the hepatotropism of HBV as it showed higher levels of activity in differentiated hepatocyte cell lines (Antonucci and Rutter, 1989; Honigwachs *et al.*, 1989; Zhang *et al.*, 1992). Both footprinting and gel retardation analyses have shown the presence of multiple binding sites in enhancer I for the liver-enriched factors HBLF, HNF3 and HNF4 (Trujillo *et al.*, 1991; Chen *et al.*, 1994).

Enhancer II (nucleotides 1636-1741) is positioned upstream of the core promoter between DR1 and DR2 and can function in an orientation- and position-independent manner, with increased activity in differentiated liver cells (Yee, 1989). Enhancer II was capable of upregulating transcription from both the SpI and SpII promoters (Yuh and Ting, 1990). The functional ability of enhancer II is strictly dependent on a bipartite structure consisting of a 23bp box α (1646-1668) and a 12bp box β (1704-1715) (Yuh and Ting, 1990; Yuh *et al.*, 1992). Liver cell-specific proteins have been shown to bind to enhancer II, again explaining the high activity of enhancer II in hepatocytes, compared with other cell types (Yuh and Ting 1993).

The glucocorticoid response element, located approximately 730bp upstream to enhancer I, has been shown to bind purified glucocorticoid receptor and to augment the activity of enhancer I in the presence of dexamethasone (Tur-Kaspa *et al.*, 1988). This may explain the increased expression of surface and core proteins observed in patients on corticosteroid treatment (Scullard *et al.*, 1981; Wu *et al.*, 1982; Ohtsu *et al.*, 1991).

2) Negative regulation of transcription

In addition to the action of the enhancer elements, viral transcription is also negatively regulated by *cis*-acting elements. In transient transfection assays, 2.4kb surface gene transcripts were difficult to detect, in contrast to the abundant 2.1kb surface mRNAs (Bulla and Siddiqui, 1989). This was due to differences in the levels of transcription from the SpI promoter and the SpII promoter. Deletion analysis of the surface gene regulatory region showed that the decreased SpI activity was due to a negative regulatory element which contained the downstream SpII region.

A second negative regulatory element is located directly upstream of enhancer II (Lo and Ting, 1994). This represses both the stimulatory effects of enhancer II and the

stimulatory effect which sequences overlapping with enhancer II have on the core promoter (Yuh *et al.*, 1992).

1B3.3 Transcriptional regulation by the X protein

The X protein transactivates a large number of viral and cellular promoters (reviewed by Rossner, 1992), including all four HBV promoter elements (Colgrove *et al.*, 1989; Siddiqui *et al.*, 1989; Raney *et al.*, 1990; Unger and Shaul, 1990). Using nuclear run-on transcription assays, increased gene expression was shown to be due to an increased rate of transcription (Colgrove *et al.*, 1989). As these experiments also utilised recircularised HBV genomes to express the X protein at *in vivo* levels, it seems likely that the role of this protein during the viral life cycle is to increase the levels of viral mRNAs required for viral replication. The precise mechanism for this transcriptional transactivation is unknown at present, but it appears that the X protein effects this by performing a dual role, depending on its subcellular location. The X protein has been reported to transactivate gene expression in two ways: (i) by directly interacting with components of the transcription machinery at the promoter (Haviv *et al.*, 1996; Cheong *et al.*, 1995; Qadri *et al.*, 1995, 1996; Maguire *et al.*, 1991; Unger and Shaul, 1990; Seto *et al.*, 1990) or (ii) by acting indirectly on transcription factors, modifying their activity through cellular signalling pathways (Benn *et al.*, 1996; Su and Schneider, 1996; Benn and Schneider, 1994; Natoli *et al.*, 1994; Kekule *et al*, 1993; Cross *et al.*, 1993; Lucito and Schneider, 1992). Experiments carried out by Doria *et al.* (1995) showed that the X protein could both stimulate signal transduction pathways (when located in the cytoplasm) and transactivate transcription elements, such as the HBV enhancer I (when located in the nucleus). Therefore the activation of gene expression by the X protein may be regulated by proteins which affect its subcellular localisation.

1B4 Viral Proteins

The four classes of HBV mRNAs act as the templates for the translation of seven viral proteins, the structures and functions of which are described below. Probably only one of these proteins, the HBe protein or e antigen, is not found in the virion; instead it is detected mainly in the serum of patients. The question as to whether the X protein is located within the virion or is synthesised only after viral entry and transcription have occurred remains to be resolved. Analysis of the kinetics of the X gene transcripts showed that they were degraded more rapidly than other viral mRNAs (Wu *et al.*, 1991b). Analysis of the X protein expressed in HepG2 cells from a recombinant vaccinia virus showed that the protein had a half-life of 3hr (Schek *et al.*, 1991). The X protein may not, therefore, remain in the cell long enough to be encapsidated, a hypothesis disputed by Wu *et al.* (1990), who claimed to detect the X protein in virions.

1B4.1 The e antigen

The 3.5kb pre-core mRNAs serve as the templates for e antigen production, as they are the only viral RNAs to include the pre-core ATG. The e antigen is translated as a 25kDa precursor protein, p25e (Jean-Jean *et al.*, 1989a; Yang *et al.*, 1992). The p25e protein includes all of the core protein residues plus an additional twenty-nine N-terminal amino acids which are largely hydrophobic. The nineteen N-terminal residues constituted a signal peptide sequence, which directed p25e to the membranes of the endoplasmic reticulum (ER) (Ou *et al.*, 1986). Co-translational signal peptide cleavage of the transmembrane protein occurred at pre-core residue Ala-19, producing p22e in the ER lumen (Standring *et al.*, 1988; Bruss and Gerlich, 1988). As p22e passes through the cellular secretory pathway it is further cleaved to remove part or all of the 37 C-terminal amino acids, producing p16e, the 16kDa secreted e antigen detectable in the serum (Magnius and Espmark, 1972; Takahashi *et al.*, 1983). As p22e was present in both the ER and the Golgi apparatus, yet p16e was found only in the Golgi, this C-terminal processing most likely occurred in the Golgi compartment. This was confirmed by suppression of the processing by brefeldin A, an inhibitor of protein transport between the ER and the Golgi (Wang *et al.*, 1991b). Aspartyl proteases were found to be responsible for the processing, as it was inhibited by pepstatin, a competitive inhibitor of these enzymes (Jean-Jean *et al.*, 1989a). Although the core protein contains an aspartyl protease-like motif, mutagenesis studies of this motif showed that the core protein was not responsible for the C-terminal processing of e antigen (Jean-Jean *et al.*, 1989b; Nassal *et al.*, 1989). These same experiments also showed that p22e could not process itself, therefore cellular enzymes are thought to be responsible. Although e antigen contains the core protein sequences necessary for particle assembly (see Section 1E1), the protein is found only as a monomer and is antigenically different from the core protein. These biophysical and antigenic differences are due to the ten pre-core amino acids remaining after signal peptide cleavage. These include a cysteine residue and a hydrophobic triad motif (WLW). The cysteine residue forms a disulphide bond with Cys-61 of the core protein (Nassal and Rieger, 1993) which, along with the hydrophobic triad motif, forces e antigen into a conformation incompatible with aggregation (Wasenauer *et al.*, 1992). This results in the synthesis of the secretory form of e antigen. Although viewed primarily as a secreted protein, e antigen and its precursors are also found in cellular compartments, presumably by evading certain translocation or processing stages. These include the nucleus (Ou *et al.*, 1989; Yang *et al.*, 1992), cytoplasm (Garcia *et al.*, 1988; Yang *et al.*, 1992) and the plasma membrane (Schlicht and Schaller, 1989) Despite understanding in detail the synthesis and biochemical properties of e antigen, ascribing a function to the protein has proved to be more difficult; as its amino acid similarity to core protein is high and it is of secretory nature, it has been proposed to act as an immune-modulatory protein (Milich *et al.*, 1990; Hsu *et al.*, 1992). However, its expression on the cell surface could

play an opposing role by allowing antibody-mediated elimination of infected cells.

Recent results from the groups of Wands and Chisari have shown that p22e may regulate HBV replication. Transient or stable overexpression of p22e reduced the levels of HBV replicative intermediates produced (Scaglioni *et al.*, 1997). Furthermore, expression of physiological levels of p22e in *trans*, from genomic constructs in transfected hepatocytes, could also reduce the levels of replicative intermediates. In transgenic studies, mice expressing pre-core protein were crossed with HBV genomic transgenic mice, which showed hepatic viral replication. Viral replication in the resulting progeny was inhibited in a dominant negative manner (Guidotti *et al.*, 1996). This was not the case in progeny resulting from the cross of surface or X transgenic mice with the HBV genomic transgenic mice. These results correlated with similar observations made in earlier transfection assays using either a genome with an authentic pre-core gene or a pre-core-minus genome, with a nonsense mutation at codon 28 (Lamberts *et al.*, 1993). These genomes showed low and high replication levels, respectively. It is probable that the pre-core protein interacts with core protein to form heterodimers which interfere with core particle assembly and pgRNA encapsidation, therefore affecting the subsequent replication steps.

1B4.2 The X protein

Although the transactivational properties of the X protein have been described previously (Section 1B3.3), other functions have also been attributed to this protein. However, for most of these it has not been possible to determine the relevance of such properties to the lifecycle of the virus. The other activities of the X protein include a novel serine/threonine kinase activity (Wu *et al.*, 1990), a Kunitz-type serine protease inhibitor activity (Takada *et al.*, 1994; Koike and Takada, 1995) and an AMP kinase activity (Dopheide and Azad, 1996). The X protein has also been shown to interact with a number of cellular proteins *in vitro*, including a probable DNA repair protein (Lee *et al.*, 1995) and a novel subunit of the proteasome complex (Huang *et al.*, 1996). However the importance of these interactions in the role of the X protein during HBV infection also remains unsolved.

The X protein may be a major determinant of the pathogenicity of HBV infection leading to the development of HCC, as X-transgenic mice develop carcinomas specifically in liver tissue (Kim *et al.*, 1991; Koike *et al.*, 1994a). The interaction of the X protein with p53, resulting in alterations in p53 function, probably contributes to its role in HCC development (Feitelson *et al.*, 1993; Wang *et al.*, 1994; Truant *et al.*, 1995; Ueda *et al.*, 1995). The expression of X protein in transfection assays induced quiescent mouse fibroblasts to enter the cell cycle (Koike *et al.*, 1994b) and expression of X protein in Chang cells, from a replication-defective adenovirus vector, deregulated the normal cell cycle checkpoint controls (Benn and Schneider, 1995). Both of these activities probably

contribute to the uncontrolled cellular proliferation involved in HCC.

1B4.3 Polymerase/Reverse Transcriptase

The viral polymerase is encoded by the 3.5kb pgRNA, but not the pre-core RNAs (Ou *et al.*, 1990; see Section 1B3.1). The P gene was putatively identified as encoding the viral polymerase/reverse transcriptase (pol/RTase) due to the size and the amino acid homology of the predicted product to the reverse transcriptases of RSV and Mo-MuLV (Toh *et al.*, 1983). This included conservation of the reverse transcriptase YMDD motif. This gene was later shown to encode the reverse transcriptase and all other activities required for viral replication; Laub and Bavand, 1988; Bavand *et al.*, 1989). Mutational analysis of the P gene showed that it encodes a protein with three functional domains (Radziwill *et al.*, 1990). These domains are termed terminal protein, pol/RTase and RNase H, from the N to the C terminus, respectively. There is a non-essential spacer region between the terminal protein and pol/RTase domains. This analysis also showed that the pol/RTase domain appeared to be divided by poorly conserved residues into two subdomains, possibly identifying the C-terminal subdomain as containing the reverse transcriptase activity, due to the presence of the YMDD motif.

The initiation of polymerase gene translation at the polymerase AUG appears to occur by a mechanism involving leaky ribosomal scanning (Lin and Lo, 1992; Fouillot *et al.*, 1993). However pre-core mRNA with the common pre-core mutation producing a stop codon at residue 28 (see Section 1D1) also has the ability to encode polymerase and core proteins (Fouillot and Rossignol, 1996). This occurs by translational reinitiation after aborted translation of the pre-core region. The importance of this phenomenon is unclear at present.

In addition to having a role in the major viral genome replication steps, the polymerase protein also plays an essential structural role during the pregenome encapsidation process (see Section 1B5).

1B4.4 Surface proteins

1) Small surface protein (SHBs)

The SHBs protein plays an important role in HBV infection and virion formation: it is the main component of the 20nm subviral particles found in serum (Bayer et al., 1968); the host of a possible virion attachment site to hepatocytes; and a structural component of the virion envelope. Translation of the S gene produces a 226 amino acid protein with a molecular weight of 24kDa, which can be N-glycosylated at Asn-146 to produce the gp27 glycoprotein. This modification is common to all three surface proteins (Peterson *et al.*, 1982). Infected liver cells (Gerber *et al.*, 1974) and cells stably transfected with the S gene (Patzer *et al.*, 1986) accumulate 20nm spheres and filaments (characteristic of SHBs) within the ER but not on the cell surface. SHBs probably inserts into the ER

membrane and particles bud into the ER lumen. Structural modelling studies have added to this by predicting a transmembrane conformation for SHBs, with α-helices spanning the membrane four times and both the N and C termini of the protein exposed on the virion surface (Stirk *et al.*, 1992; Berting *et al.*, 1995). This transmembrane orientation is due to the targeting action of two hydrophobic signal sequences located between amino acids 4-28 and 84-136 (with residues 80-98 being highly conserved), termed signals I and II respectively (Eble *et al.*, 1987; Bruss and Ganem, 1991a). However, in contrast to conventional signal sequences, signals I and II are not cleaved in the ER. Signal I directs the translocation of the N-terminus of SHBs across the ER membrane and is therefore also important during MHBs synthesis (Eble *et al.*, 1986, 1990). Signal II directs the translocation of the C-terminus of SHBs, resulting in a surface protein with two transmembrane spans. Whether or not the C-terminal amino acids after residue 160 form another two transmembrane spans, which would be in accordance with the modelling studies, is unclear.

After adopting this orientation, SHBs proteins first dimerise and then oligomerise before budding to form subviral particles (Simon *et al.*, 1988); this last stage occurs in a post-ER, pre-Golgi compartment (Huovila *et al.*, 1992) before particles are secreted via the cellular pathway.

The transmembrane topology of SHBs produces two major hydrophilic loops, consisting of amino acids 28-77 and 101-159. The first loop is exposed to the cellular cytoplasm during synthesis and lies along the inner surface of the envelope of the virion. Mutational analysis of the cysteine residues in this region showed that Cys-48, Cys-65 and Cys-69 were all essential for secretion of SHBs particles, as was His-60 (Mangold and Streeck, 1993). The second hydrophilic loop, as the site of SHBs glycosylation, is located in the ER lumen and is therefore exposed on the external surface of the virion. This region contains eight highly conserved cysteine residues which participate in disulphide bonds. These bonds dictate the complex structure of SHBs and are important for its antigenicity (Ashton-Rickardt and Murray, 1989a; Bruce and Murray, 1995). The immunodominant B-cell epitope, the *a* determinant, is located within the conformational structure of this region, as are the *d/y* and *w/r* subtype determinants: lysine or arginine at residue 122 (le Bouvier, 1971; Peterson *et al.*, 1984; Okamoto *et al.*, 1987a) and lysine or arginine at residue 160 (Bancroft *et al.*, 1972; Okamoto *et al.*, 1987a), respectively. However, other residues may also affect the reactivity of this region to subtype-specific antibodies (Ashton-Rickardt and Murray, 1989b; Okamoto *et al.*, 1989). Progress towards elucidating the conformation of this region has come from observing the protective efficacy of peptide vaccines covering amino acids 117-137, which had been cyclised by a disulphide bond between Cys-124 and Cys-137 (Dreesman *et al.*, 1982) and the use of phage display libraries (Chen *et al.*, 1996). The model for SHBs produced using the latter technique proposes disulphide bonds between cysteines 107-138, 137-

149, 139-147 and 121-124. However, earlier work showed that although cysteine residues at 124, 147, and 149 were essential for maintenance of the conformation of this region, mutation of Cys-138 had little or no effect (Ashton-Rickardt and Murray, 1989a; Bruce and Murray, 1995). These results do not nullify the proposed model, as Cys-138 is not an essential residue: it is claimed that the structure of SHBs would not alter drastically if cysteine residues 107, 137, 138 and 149 did not form intra-chain disulphide bonds, or if they formed disulphide bonds with other monomers within the particle. All cysteine residues in this region have been shown to contribute to the efficiency of SHBs particle secretion (Mangold and Streeck, 1993).

2) Middle surface protein (MHBs)

The MHBs protein incorporates the entire SHBs amino acid sequence plus an additional 55 N-terminal amino acids, termed the pre-S2 region (Stibbe and Gerlich, 1983). The pre-S2 domain contains an additional site for N-linked glycosylation at Asn-4 which, in combination with the common site at Asn-146, gives rise to the two MHBs glycoproteins, gp33 and gp36. MHBs is a component of both HBV virions and surface antigen particles (Stibbe and Gerlich, 1982; Heermann *et al.*, 1984; Machida *et al.*, 1983). The pre-S2 amino acids are thought to be exposed on the virion surface, as antibodies raised in chimpanzees to residues 14-32 were able to bind virions and protect the immunised animals from 10^6 infectious doses of HBV (Itoh *et al.*, 1986). However, whether it is MHBs or LHBs that is responsible for these effects *in vivo* is unclear, as both proteins contain the pre-S2 region. It has been proposed that during M protein production, 20nm particles are formed in the ER lumen before being transported to a post-ER/pre-Golgi compartment and fully glycosylated. It may be that some MHBs can evade this second glycosylation stage by being directed to the plasma membrane and secreted instead (Sheu and Lo, 1994).

3) Large surface protein (LHBs)

The largest protein encoded by the S gene, translated from the 2.4kb mRNA, is LHBs. This protein includes S, pre-S2 and the N-terminal pre-S1 amino acids; the latter region is encoded for by a highly variable region of the HBV genome. The pre-S1 domain is 108-119 amino acids long, depending on the viral subtype, and translation initiating at the pre-S1 start codon produces a 39kDa protein. LHBs can be present in both 20nm spheres and filaments (more commonly in the latter), but is more prevalent in virions, which contain up to twenty times more LHBs than subviral particles (Heermann *et al.*, 1984). LHBs is also N-glycosylated to produce gp42. The pre-S1 amino acids are exposed on the surface of the virion, as antibodies raised against whole virions are capable of binding to pre-S1 residues. The pre-S1 region probably masks the S region in LHBs, as digestion of 20nm particles with trypsin increases their reactivity with anti-

SHBs antibodies.

As LHBs was finally present on the virion surface, it was presumed that its biosynthesis followed the same pathway as the S and M proteins, with the N-terminus co-translationally translocated into the ER lumen. However, results from *in vitro* and cellular assays showed that the N-terminal regions of virtually all LHBs proteins (possibly up to the signal II sequence), were initially located in the cytoplasm (Ostapchuk *et al.*, 1994; Bruss *et al.*, 1994). A post-translational translocation event then occurs to reorientate pre-S1 residues externally. The mechanism for retention of the pre-S1 residues in the cytoplasm is unclear, but does not seem to be due to membrane anchorage mediated by the myristylated N-terminal glycine (Persing *et al.*, 1987; Prange and Streeck, 1995). Signal I residues in the S region may be unable to direct the translocation of the 163-174 amino acid pre-S residues, which have no signal of their own (Ostapchuk *et al.*, 1994). The C-terminal pre-S1 region (amino acids 70-107), however, has been shown to interfere with co-translational translocation of LHBs protein, possibly by interacting with cytoplasmic cellular proteins and inhibiting the translocation process (Prange and Streeck, 1995). These residues may, therefore, be important for regulating the later translocation event. The LHBs topology switch is likely to involve a reorganisation of lipid membranes, but insufficient studies have concentrated on these to allow determination of the mechanism. The initial cytoplasmic orientation of LHBs may also explain why it is not glycosylated in the pre-S2 region. It is possible that both SHBs and the S region of LHBs have a similar dual topology in the second hydrophilic loop, thus explaining the partial glycosylation observed at Asn-146. However reports examining this possibility are in conflict at present (Prange and Streeck, 1995; Wunderlich and Bruss, 1996).

1B4.5 Core protein

As this protein is the subject of this thesis (see Section 1E) its role in the viral lifecycle will be described only briefly here.

The 21kDa core protein is the sole component of the viral nucleocapsid, or core particle, as was shown by the production of 27nm diameter particles when the core gene was expressed in *E.coli* (Cohen and Richmond, 1982). The icosahedral core particle is composed of 90 or 120 core protein dimers (Crowther *et al.*, 1994) and assembles to encapsidate the polymerase/pgRNA complex (see Section 1B5). Reverse transcription of the pgRNA and the subsequent replication stages occur within core particles, which are present in the cytoplasm before entering the virion assembly pathway.

1B5 Pregenome Encapsidation and Virion Assembly

As well as being the initiation site for reverse transcription of pgRNA, the 5' ε signal is an essential element for the encapsidation of pgRNA (Junker-Niepmann *et al.*, 1990). The virion assembly process is initiated by the specific interaction of ε with the viral

polymerase (Bartenschlager *et al.*, 1990; Roychoudhury *et al.*, 1991; Bartenschlager and Schaller, 1992). This interaction may then be further stabilised by interactions with core protein molecules until particle assembly occurs. The pgRNA is then packaged within the lumen of the core particle with the polymerase-ε complex intact. Mutational analysis of the nucleotide sequence of ε showed that the only regions not important for polymerase binding were the four distal bulge nucleotides (although the bulge structure itself was required for this interaction) and the lower region of the stem. This suggested a model in which the polymerase covers almost the entire upper portion of ε (Pollack and Ganem, 1993; Fallows and Goff, 1995; Rieger and Nassal, 1995). A more complex scenario has arisen for the DHBV encapsidation process, which has been studied in more detail. It appears that the polymerase-ε interaction is insufficient for the encapsidation and priming of DNA synthesis and other factors, such as Hsp90 and p23 are required (Pollack and Ganem, 1994; Tavis and Ganem, 1996; Hu and Seeger, 1996; Hu *et al.*, 1997). It remains to be seen whether similar additional factors are required for HBV.

Although pre-core mRNAs also contain ε signals, pgRNA is selectively encapsidated because the stem-loop structure of ε on pre-core mRNA is disrupted by the 80S translating ribosomes (Nassal *et al.*, 1990). Although pgRNA also acts as a translational template, its translation is carried out by 40S scanning ribosomes. These are probably blocked by the polymerase-ε interaction and core protein molecules stabilising this latter interaction. Therefore the concentrations of both polymerase and core proteins are able to regulate initiation of encapsidation.

Following encapsidation of the polymerase-pgRNA complex, the core particle is enveloped by a lipid bilayer containing the HBV surface proteins. As all three surface proteins are present in virions they must localise simultaneously in the same membrane region. With the exception of Ueda *et al.* (1991), most investigators conclude that, although present in virions, MHBs is dispensable for virion assembly (Bruss and Ganem, 1991b; Fernholz *et al.*, 1993). The amount of LHBs present, and hence the SHBs:LHBs ratio, determines the type of particle produced (Heermann *et al.*, 1984). The 20nm spherical surface proteins contained low levels of LHBs, whereas both filaments and virions contained higher levels. The amounts of LHBs protein in virions were up to twenty-fold higher than LHBs levels in the 20nm surface particles. However, overexpression of LHBs resulted in the inhibition of surface particle secretion (Persing *et al.*, 1986; Chisari *et al.*, 1986). It was proposed that this LHBs-mediated retention of surface proteins allowed all three proteins to aggregate in the ER membrane before the envelopment of core particles. However, this model was disproved when virion release was also shown to be inhibited by LHBs overexpression (Bruss and Ganem, 1991b). Pre-S1 residues, therefore, are important for virion production. In particular the C-terminal 16 amino acids (104-119) are necessary, as N-terminal deletions which removed them prevented virion release (Bruss and Thomssen, 1994).

Due to the dual topology of LHBs, cytoplasmic pre-S1 residues may be involved in protein:protein interactions required for recognition of core particles prior to envelopment. The use of a phage display library suggested that sequences which bound to core particles mimicked pre-S1 residues 19-24 and may therefore be important for core:surface interactions (Dyson and Murray, 1995). However these residues are included in the N-terminal pre-S1 region which is dispensable for virion formation. Binding assays between core particles and synthetic peptides based on the amino acid sequences of pre-S1, pre-S2 and the cytosolic loop of S showed that the thirteen C-terminal pre-S1 residues were capable of efficient binding (Poisson *et al.*, 1997). This was more consistent with the results from the LHBs deletion experiments.

Virion assembly also appears to be regulated by viral DNA replication events occurring in the core particle interior (Gerelsaikhan *et al.*, 1996). A presently undefined signal essential for virion maturation appears to be displayed on the exterior of the core particle some time after the initiation of reverse transcription of the pregenome. This may be a conformational change induced in the particle structure, possibly due to the replacement of the structurally flexible pgRNA with a more rigid double-stranded DNA. Alternatively core protein phosphorylation by protein kinase C, which is also encapsidated (Kann *et al.*, 1993), may result in a conformational change. As the core protein kinase C target sequence within core protein is located within the C-terminal nucleic acid binding region (Kann and Gerlich, 1994), dissociation of the pgRNA during reverse transcription would allow protein kinase C access to this sequence. This would be followed by phosphorylation of the core protein and the induction of a conformational change.

Following the interaction of the core particle and surface proteins, virions are generated by budding of the core particles into the ER lumen, along with the surface protein-containing membrane. The resulting virions are then exported from the cell via the constitutive secretory pathway.

IC IMMUNOPATHOGENESIS OF HBV INFECTION

Upon infection with HBV most adults develop a self-limiting acute hepatitis and are able to eliminate infected hepatocytes. However, approximately 5% of infected adults are unable to clear the virus and develop a chronic infection with or without hepatic inflammation: chronic active hepatitis (CAH) and chronic persistent hepatitis (CPH), respectively (Chu *et al.*, 1985). Studies using various transgenic mouse lineages have shown that the pathologies observed are not due to a direct viral cytopathic effect (Farza *et al.*, 1988; Araki *et al.*, 1989; Guidotti *et al.*, 1995), but rather are caused by the subsequent immune response directed against viral proteins (Moriyama *et al.*, 1990). The outcome of infection seems to depend primarily on the strength of the response mounted

by the host immune system against the virus. The humoral response against virion surface proteins neutralises viral infectivity, whilst the cellular response against these and other viral antigens destroys infected hepatocytes. Individuals who develop chronic infections are thought to be especially deficient in this latter response.

1C1 Humoral Response

Antibodies to all viral proteins can be detected in HBV-infected individuals. However, with the exception of anti-HBs antibodies, the roles played by the other antibodies produced during the course of virus infection are not fully understood. Anti-polymerase antibodies are induced in both acute and chronic hepatitis but seem to serve only as a marker of advanced liver disease and ongoing viral replication (Yuki *et al.*, 1990; Weimer *et al.*, 1990). Titres of anti-HBx antibodies were found to be higher and occurred more frequently in the sera of patients with chronic hepatitis infections compared to acute cases or asymptomatic carriers (Stemler *et al.*, 1990).

Although e antigen serves as a standard serological marker for HBV replication and seroconversion to anti-HBe antibody is usually predictive of a remission in liver disease (Hoofnagle *et al.*, 1981), the role played by anti-HBe antibodies in these processes is unclear. Antibodies to both e antigen and core proteins are produced in high titres in patients with acute and chronic infection in spite of them having no obvious virus-neutralising properties. The ability of core particles to act as both a T helper cell (Th)-independent and a Th cell-dependent antigen (Milich and McLachlan, 1986) and the high cross reactivity of core and e proteins at the T cell level (Milich *et al.*, 1987a; 1988) probably explains the high antibody levels observed.

Early antibody binding studies on both liver-derived and bacterially-expressed core particles showed that a single immunodominant epitope was present, which was recognised by both monoclonal antibodies and human sera (Waters *et al.*, 1986; Ferns and Tedder, 1986). However, further studies have mapped numerous antibody binding epitopes in the core protein (see Figure 1C1). This is possibly a reflection of the differences in the immunogens used to raise antibodies and the binding assay formats. The use of core particles or core protein, in the form of fusion proteins, suggested that the dominant epitope was located in the centre of the protein, around amino acid 74 (Bichko *et al.*, 1993; Salfeld *et al.*, 1989; Schödel *et al.*, 1992). This epitope was shown to be exposed on the surface of the particle and conformational. The conformational nature of core antigenicity had been previously demonstrated by protease studies on core particles (Mackay *et al.*, 1981). Limited proteolysis of particles under dissociating conditions resulted in decreased core antigenicity, with a concomitant increase in e antigenicity. This suggested that the particulate core antigen was being disrupted to produce e antigenicity. Salfeld *et al.* (1989) proposed that two epitopes existed on e antigen: HBe1, a predominantly linear epitope from amino acids 76-89, which could also be detected on the

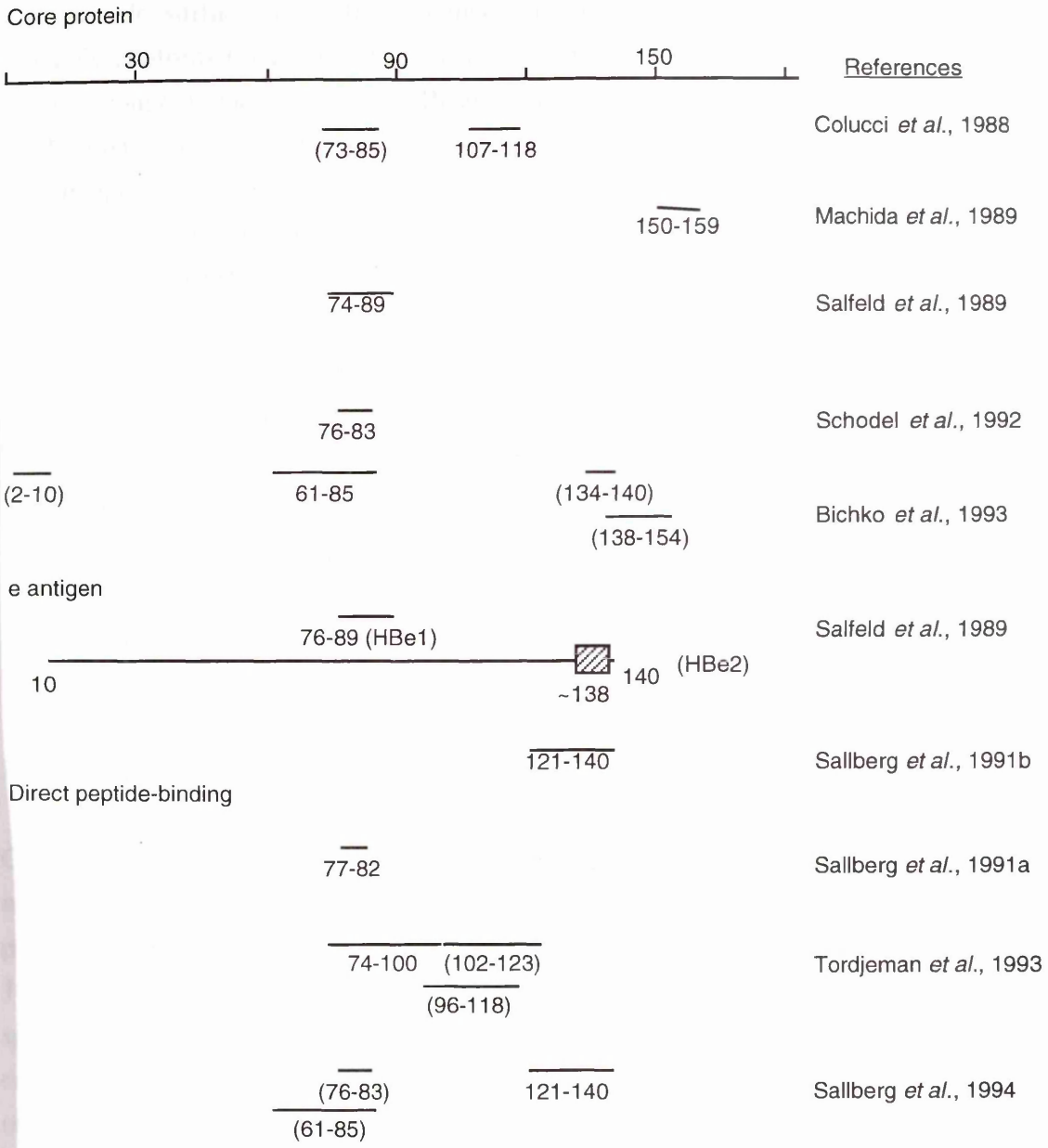


Figure 1C1: B-cell epitopes in core protein

The positions of mapped B-cell epitopes in the core protein sequence are shown along with the relevant reference. Bracketed positions indicate that the epitope was recognised less frequently or was a minor epitope. The shaded box in the HBe2 epitope indicates that this region is important for antigenicity, but intra-molecular participation from the extended sequence shown is also required. Epitopes in the lower region of the figure were identified by antibodies binding synthetic peptides.

core particle surface, and HBe2, which required most of the polypeptide and was probably conformational in nature. HBe1 overlaps with the conformational core protein epitope around amino acids 74-89. Results from Colluci *et al.* (1988) and Sallberg *et al.* (1991a) have confirmed the presence of linear epitopes within the conformational core protein epitope.

Other epitopes have been identified using synthetic peptides (Tordjeman *et al.*, 1993; Sallberg *et al.*, 1994), but as these may lack any conformation they may not be significant *in vivo.*

Antibodies directed against the surface proteins, and in particular the *a* determinant, have been shown to provide protection against virus challenge in chimpanzees (Iwarson *et al.*, 1985; Itoh *et al.*, 1986). As these antibodies are detectable in patients who have cleared the virus and are undetectable in chronic HBV infections, they are thought to play an essential role in virus neutralisation (Pontisso *et al.*, 1989a). The production of anti-HBs antibodies is dependent on Th cells (Roberts *et al.*, 1975) and is also linked to the activation of Th cells by core protein epitopes, which can elicit production of both anti-HBc and anti-HBs antibodies (Milich *et al.*, 1987b).

1C2 HLA Class II-restricted T Cell Response

In acute hepatitis cases, strong human leukocyte antigen (HLA) class II-restricted $CD4^+$ Th responses to core protein and e antigen are detectable in peripheral blood mononuclear cells (PBMCs). A number of Th epitopes have been identified in core protein, some of which are unrelated to HLA-type (Ferrari *et al.*, 1990, 1991; Jung *et al.*, 1991, 1995; Diepolder *et al.*, 1996; see Figure 1C2). In contrast, the surface protein-specific Th response is much weaker, and although this response has been detected in the early incubation phase of acute infections (Vento *et al.*, 1987) the basis for the weaker response in the symptomatic phase is unclear. The possible depressing effect of the abundance of surface particles on surface-specific Th cells may provide a partial explanation. The lack of surface-specific Th cell stimulation for the production of anti-HBs antibodies is compensated for by core protein epitopes, as described above.

As yet, no substantial studies have been made on the class II-restricted T cell response to the polymerase and X proteins.

During chronic infection, the class II-restricted T cell response to all viral antigens is significantly lower compared to that in an acute hepatitis infection (Ferrari *et al.*, 1990; Jung *et al.*, 1991; Wakita *et al.*, 1992; Löhr *et al.*, 1995). This probably contributes to the antibody profiles observed during chronic infection: both sources of Th-stimulation of anti-HBs antibody production (surface protein epitope- and core protein epitope-specific Th cells) are reduced. However, core particles are still capable of eliciting anti-HBc antibody production independent of Th cells.

Analyses of hepatic infiltrates during chronic infections have shown the presence of

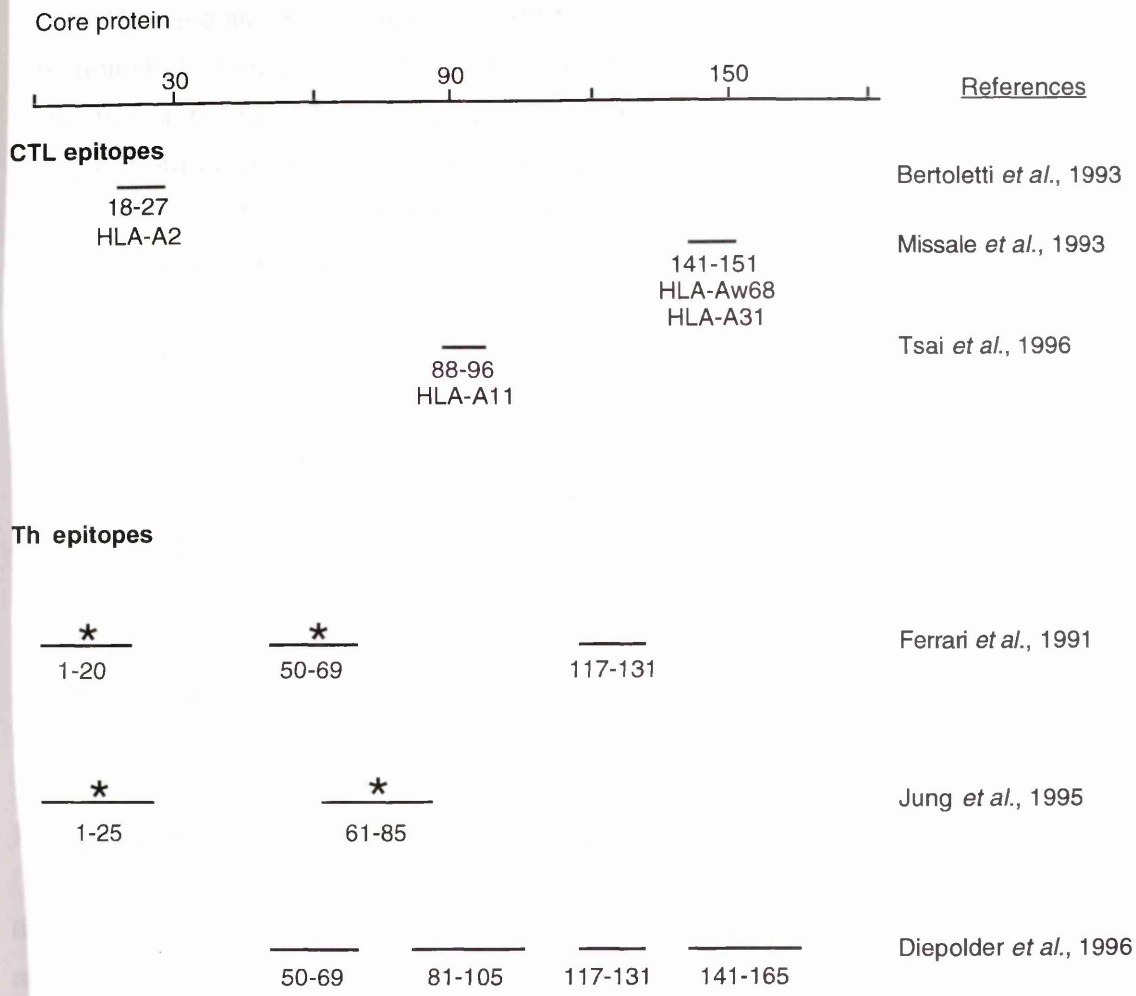


Figure 1C2: T cell epitopes of core protein
The positions of epitopes for cytotoxic T lymphocytes (CTL) and helper T lymphocytes (Th) in core protein are shown above. For CTL epitopes, the HLA type of the patient is indicated. Th epitopes which are indicated by an asterisk (*) were identified irrespective of the patient's HLA type.

$CD4^+$ T cells specific for both core and surface proteins, in contrast to peripheral blood (Ferrari *et al.*, 1987a, 1987b; Barnaba *et al.*, 1989, 1994; Löhr *et al.*, 1995). It may be that these cells are enriched in this compartment during chronic infections.

Core-specific Th cells are thought to play a key immunoregulatory role during infections, whereas surface-specific Th cells may be suppressed by the high antigen load. High core-specific Th cell activity is observed in the following incidences: (i) during acute hepatitis; (ii) in patients with chronic HBV infection compared to the asymptomatic carrier state (Wakita *et al.*, 1992; Maruyama *et al.*, 1993a); (iii) during acute exacerbations of chronic infections (Tsai *et al.*, 1992; Maruyama *et al.*, 1993b); (iv) during seroconversion

to anti-HBe and anti-HBs (Jung *et al.*, 1995); (v) in spontaneous remission from chronic infection (Rehermann *et al.*, 1996); and (vi) with transaminase normalisation, anti-HBe seroconversion and viral DNA clearance in chronic patients who respond to interferon therapy (Löhr *et al.*, 1995). The association between the class II-restricted DRB1*1302 allele and protection against persistent HBV infection in the Gambia also provides evidence for the importance of the Th response in the outcome of HBV infection (Thursz *et al.*, 1995).

Studies in a murine system have shown that a combination of the MHC restricting element, the core/e epitope recognised and the structure of the antigen can, along with other factors, determine whether the Th cell response is dominated by Th1 or Th2 subsets, which may also be a factor in determining the outcome of infection (Milich *et al.*, 1995a, 1997).

In summary, it is likely that the lack of a sufficient core-specific Th response is a key factor in the failure to eliminate virus and in the development of chronic hepatitis. The low levels of activity observed are probably responsible for the ongoing necroinflammatory disease.

1C3 HLA Class I-restricted T Cell Response

Studies on the cytotoxic T lymphocyte (CTL) response in acute hepatitis have shown that this is polyclonal and is directed against multiple epitopes in all viral proteins (Jin *et al.*, 1988; Nayersina *et al.*, 1993; Barnaba *et al.*, 1989; Ferrari *et al.*, 1992; Rehermann *et al.*, 1995; Missale *et al.*, 1993; Penna *et al.*, 1991; Bertoletti *et al.*, 1991, 1993; see Figure 1C2). The abundance of surface proteins produced during infection and the observation that exogenous surface protein can also enter the class I processing pathway (after initial endosomal processing) and induce CTL and Th cells as effectively as endogenously synthesised protein (Jin *et al.*, 1988; Penna *et al.*, 1992), may also contribute to the strength of the CTL response against surface proteins. However, uptake and presentation of exogenous surface protein also occurred in B cells and resulted in anti-HBs antibody-producing B cells being destroyed, leading to the decreased anti-HBs antibody levels which accompany the development of chronic hepatitis (Barnaba *et al.*, 1990). The core protein has only one HLA-A2 epitope, whereas surface and polymerase proteins have four and five, respectively. Therefore, the immune response to core protein does not seem to predominate in the $CD8^+$ CTL activity, as it does for the $CD4^+$ Th cells.

Studies of the acute liver disease induced in surface protein transgenic mice upon transfer of surface protein-specific CTL have revealed the important events in disease pathogenesis (Ando *et al.*, 1993). The first step is the triggering of target hepatocytes to undergo apoptosis. This occurs almost immediately as a direct consequence of CTL-target cell interaction. Within a few hours this is followed by a focal inflammatory response where the cytopathic effects of the CTL response are amplified in an antigen-nonspecific

manner by lymphocytes which have been recruited by CTL. Thus the liver disease observed is transient and relatively mild, where no more than 5% of hepatocytes are destroyed and with hepatic damage being mediated mainly by CTL-activated effector cells and antigen-nonspecific lymphokines.

A further series of experiments using transgenic models have shown that selected soluble lymphokines produced by antigen-specific CTL are themselves able to non-cytolytically downregulate HBV gene expression at the post-transcriptional level (Guidotti *et al.*, 1994b; Tsui *et al.*, 1995). Although the precise molecular pathways responsible remain unsolved, it appears that both interferon-γ (IFNγ) and tumour necrosis factor-α are principally responsible, although interleukin-2, IFNα and IFNβ are also involved (Gilles *et al.*, 1992; Guilhot *et al.*, 1993; Guidotti *et al.*, 1994c). Therefore, in addition to the direct cytolytic action of antigen-specific CTL on infected hepatocytes, CTL-derived signals can activate the infected cells to an antiviral state, thus providing an additional means of combating HBV infection.

In contrast to the vigorous CTL response detectable in the peripheral blood of patients with acute hepatitis who successfully clear the virus, a weak or undetectable CTL response is associated with chronic disease and viral persistence (Montaño *et al.*, 1983; Nouri-Aria *et al.*, 1988; Ferrari *et al.*, 1990; Jung *et al.*, 1991, 1995; Löhr *et al.*, 1993; Rehermann *et al.*, 1995). It is likely that this is linked to the low $CD4^+$ Th cell activity observed in these cases as described previously (see Section 1C2). The resultant inability to mount a vigorous CTL response is widely acknowledged as the primary reason for the failure to rapidly eliminate virus and instead allow the development of a chronic infection.

The basis for the variable CTL response is poorly understood, except perhaps in the case of infants born to e antigen-positive mothers. In these cases, where infection is thought to occur at or around the time of birth (Beasley and Hwang, 1983), approximately 95% of neonates develop chronic infections instead of clearing the virus. The probable explanation for this is that e antigen, being a secreted, non-particulate protein, is able to cross the placenta and is present in the fetal circulation (Hsu *et al.*, 1992). Consequently, thymic deletion of both e- and core-specific T cells occurs due to the cross-reactivity of both proteins at the T cell level (Milich *et al.*, 1990). This results in a depletion of the cells responsible for directing viral clearance. Thus, if infection of the fetus (Alexander and Eddleston, 1986) or of the neonate (Beasley and Hwang, 1983; Li *et al.*, 1986) occurs while it is tolerant to core/e protein at the T cell level, chronic infection will develop. Additional studies of incomplete tolerance in e antigen transgenic mice have shown that Th2 cells can evade tolerance induction more successfully than their Th1 counterparts (Milich *et al.*, 1995b). This is also likely to contribute to the abnormal T cell responses observed in chronically infected patients.

Suppression of cellular responses to interferons may also contribute to the inefficiency of the CTL response. Expression of the viral polymerase inhibited the response to IFNα

and γ, by preventing the activation of interferon-inducible genes (Foster *et al.*, 1991 (see also *erratum*, 1995), 1993). This phenomenon was observed both *in vitro* and in chronic patients who failed to respond to IFNα therapy. Core protein is also able to inhibit expression of IFNβ by acting in *trans* on the regulatory region of the gene (Twu *et al.*, 1988; Whitten *et al.*, 1991). Hepatocytes from patients with e antigen-positive chronic hepatitis showed no increased interferon-induced membrane expression of HLA class I molecules compared to anti-HBe chronic cases and uninfected controls (Montaño *et al.*, 1982; Ikeda *et al.*, 1986). In addition, levels of 2'5'-oligoadenylate synthetase were not increased in patients with CAH, in contrast to those with CPH or acute hepatitis (Poitrine *et al.*, 1985).

Whilst this section has primarily considered the host response to HBV infection and its possible defects which allow viral persistence, it is likely that variation within the infecting viral population is another important factor.

1D GENETIC VARIATION

With a mutation rate calculated to be comparable to the slowest evolving retroviral genes (Girones and Miller, 1989), a large number of HBV variants have been detected during HBV infection. However, both the replication competence of a particular variant and the resultant immune response determine to what extent it is selected for and whether or not it predominates in the viral population. Apart from the previous mention of surface gene variation in relation to subtype definition (Section 1B4.4), only pre-core and core gene variation will be discussed in this section as recent reviews have been published on both surface and X gene variation (Wallace and Carman, 1997; Carman, 1995a).

1D1 Pre-core Variants

Since the description of the original pre-core mutant, G1896A (substitution of a G nucleotide with an A at position 1896), in which the pre-core codon 28 was converted from a tryptophan to a stop codon, many other pre-core variants have been described and are summarised in Figure 1D1 (Carman *et al.*, 1989; Fiordalisi *et al.*, 1990; Okamoto *et al.*, 1990; Raimondo *et al.*, 1990; Tong *et al.*, 1990; Blum *et al.*, 1991; Santantonio *et al.*, 1991; Carman *et al.*, 1992; Günther *et al.*, 1992; Laskus *et al.*, 1994). Although not all variation leads to the formation of a stop codon, production of e antigen is precluded by many of the variations. Other variants have been reported which do not affect e antigen synthesis but these are not shown in this Figure 1D1 (Akarca *et al.*, 1994; Laskus *et al.*, 1994). The reason for the 200-fold higher prevalence of G1896A compared to other variations is unknown and it is simply termed a mutational 'hotspot' (Yuan *et al.*, 1995).

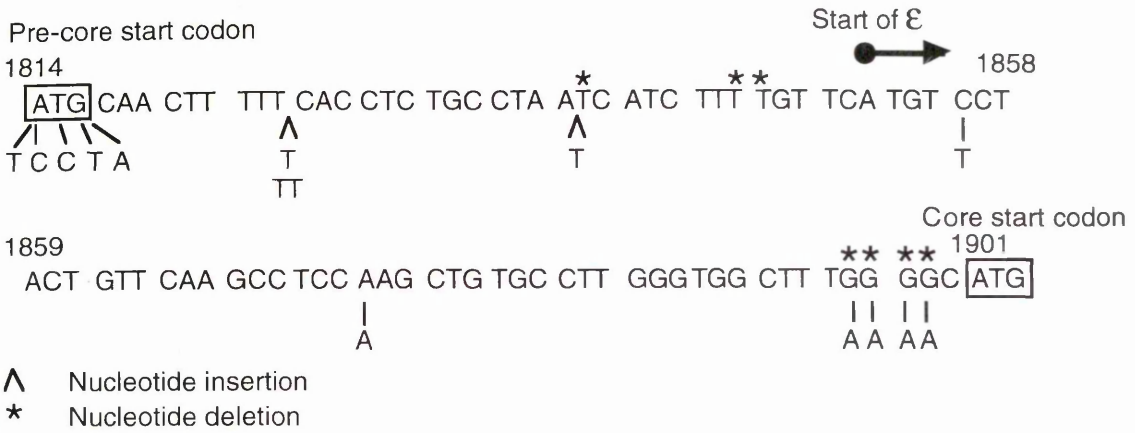


Figure 1D1: Variations observed within the pre-core region

The nucleotide sequence of the pre-core region is shown, along with reported mutations that abolish the production of e antigen, and some additional commonly observed variants. The key at the bottom of the figure depicts the type of mutation; all others shown are nucleotide substitutions. The start codons for both e antigen (pre-core start codon) and core protein are boxed. References identifying these mutations are cited in the main text.

Examination of the secondary structure of the ε RNA sequence (see Figure 1D2) shows that the G1896A and G1899A mutations replace the non-Watson-Crick U:G bond with a U-A bond. This may stabilise the ε signal (Lok *et al.*, 1994), especially as G1899A rarely occurs without G1896A. The secondary structure of the ε sequence may also explain the occurrence of other pre-core variations. For example, C1856T, which was always detected in association with the conservative variation T1858C, was found to be mutually exclusive with G1896A, possibly as the new combination of base pairs would lead to an unstable encapsidation signal (Carman *et al.*, 1992; Lok *et al.*, 1994). The C1856T variation does not seem to affect either the amount of e antigen produced, or the efficiency of its secretion from cells (Boner *et al.*, 1995). C1856T can, however, be detected in combination with T1856C and G1898A, as the variation at the two positions maintains the overall number of base pairs in the lower stem of ε. The positions of other naturally occurring pre-core variants have been mapped and it is interesting to note that no variation occurs in the loop or right-hand side of the upper stem of ε; these sequences are indispensable for encapsidation. Also, there is no variation in the UUCA sequence in the bulge, which is required for the initiation of minus-strand DNA synthesis (Laskus *et al.*, 1994). Another line of evidence for the importance of the structure of ε in permitting

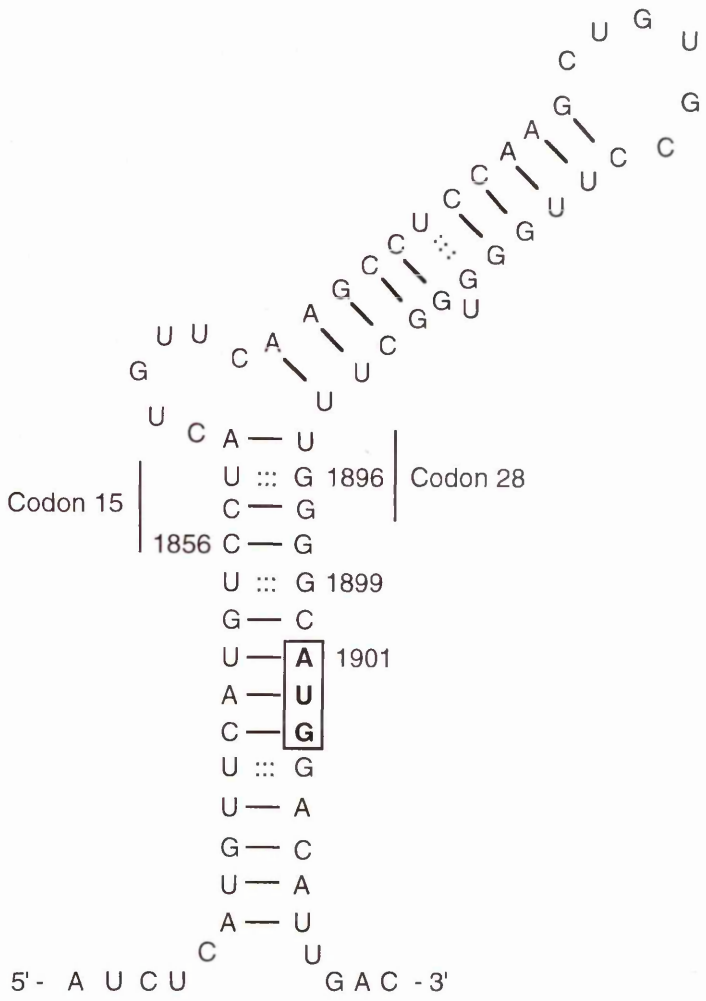


Figure 1D2: Secondary structure of the ε encapsidation signal
The pregenomic RNA encapsidation signal possesses a stem-loop structure formed by sequences in the pre-core region. This structure includes the six nucleotide bulge required for initiation of minus-strand DNA synthesis. The positions of some important variant nucleotides are marked and the start codon for the core gene is boxed.

variation at only certain positions is the observation that G1896A is very rarely found in viruses of genotype A (Li *et al.*, 1993), which have the sequence CCC at positions 1856-1858. The G1896A mutation would destroy the C-G base pair which, as shown in transfection experiments, decreases the RNA packaging efficiency.

The G1899A mutation is also commonly described, usually in association with G1896A; the combination has been linked with severe disease (Carman *et al.*, 1989; Ulrich *et al.*, 1990; Hasegawa *et al.*, 1994). However, they have no effect on the levels of viral RNA or replicative intermediates produced after transfection of genomic constructs

containing both mutations (Hasegawa *et al.*, 1994). Virion assembly and secretion were also similar to wild-type (Ulrich *et al.*, 1990), demonstrating that neither the variants nor the absence of e antigen affects any of these processes. Similarly, the G1896A variation alone did not affect the replication efficiency of transfected genomes (Hasegawa *et al.*, 1994; Tong *et al.*, 1992).

Many studies have been carried out in order to determine whether or not a link exists between selection of G1896A and the severity of the disease. During and after seroconversion to anti-HBe, pre-core stop mutants were detected both in patients who become asymptomatic and also in patients who continued to have ongoing liver disease (Lai *et al.*, 1994; Tur-Kaspa *et al.*, 1992; Okamoto *et al.*, 1990; Tong *et al.*, 1990). These same studies also showed that a high proportion of pre-core stop mutants after seroconversion was, in general, associated with a higher viremia and a more active hepatitis whilst mixed populations of G1896A strains and e antigen-producing strains seemed to be associated with a less severe form of hepatitis (Carman *et al.*, 1989; Naoumov *et al.*, 1992; Lai *et al.*, 1994). Predominance of pre-core stop mutants after seroconversion to anti-HBe was also associated with a course of hepatitis characterised by flare-ups of liver cell necrosis interspersed with asymptomatic periods (Brunetto *et al.*, 1991). That selection of a pre-core mutant virus does not necessarily lead to a quiescent state indicates that other factors contribute to disease progression. One possibility is variation within the core gene, although variation in other genes may be as important.

The most likely mechanism of selection for pre-core stop mutants is the pressure imposed by the immune system. The lack of a virus protein which is normally secreted and accessible to the immune system may result in less scope for sensitisation of immune effector cells and recognition of hepatocytes containing either e or core proteins.

1D2 Core Variation

As the core protein is thought to be a major target of the antiviral immune response, a number of groups have examined the variation occurring in the core gene during chronic infection, to determine whether or not a causal link exists. These studies attempted to identify whether amino acid changes were more prevalent in certain regions and whether these regions were functionally relevant.

1) Do core mutations cluster?

Unlike the pre-core region, only the extreme 5' end of the core gene contains elements which are functionally essential at the RNA level. Therefore, for the majority of the nucleotide sequence only missense mutations, affecting the amino acid sequence, are important. The high ratio of non-synonymous to synonymous mutations also implies that mutation at the amino acid level is of consequence.

There have been six major studies to date that have examined mutation clustering in the core gene (Ehata *et al.*, 1992, 1993; Chuang *et al.*, 1993; Boner *et al.*, 1995; Carman *et al.*, 1995b; Akarca and Lok, 1995a). The results of all studies showed that the core gene mutations affecting the core protein did cluster in specific regions, but the locations and sizes of these regions were disputed.

Ehata *et al.* (1992) showed that, of 39 core gene mutations leading to amino acid substitutions in virus isolates from 20 chronic carriers with fluctuating ALT levels, 21 (54%) clustered in the region encoding amino acids 84-101. In addition, 15 of these patients, suffering from CAH, all had at least one mutation in this region, whereas the 5 remaining patients had no mutations in this region and had only CPH. Upon closer examination of the 84-101 region, 14 of the 15 patients all had mutations leading to amino acid substitutions within the smaller region encoding amino acids 87-97. A second mutation clustering region was identified in the region encoding amino acids 130-156, with 14 of the 39 mutations (36%) located here and, in addition, sporadic mutations were identified at amino acids 27, 49 and 60.

In a second paper the same group identified 29 or 34 missense mutations in the core genes from patients infected with *adr* or *adw* subtype viruses respectively (Ehata *et al.*, 1993). These patients all had either acute exacerbations of chronic hepatitis or fulminant hepatitis, the most severe form of the disease. For the *adr* subtype viruses, mutations clustered mainly between amino acids 84-99 (12/29; 41%), 130-156 (6/29; 21%) or 20-35 (5/29; 17%) with sporadic mutations at amino acids 5, 13, 60, 74 and 105. Mutation in the *adr* subtype viruses, however, clustered between amino acids 48-60 (9/34; 26%) or 105-113 (5/34; 15%), with 7/8 patients having an additional substitution at amino acid 77. Sporadic mutations were observed at ten other positions.

A third paper by this group detected 62 missense mutations in the core genes from 17 CAH cases (Chuang *et al.*, 1993). Similar clustering regions to those described previously were examined and found to contain 19/62 (31%), 14/62 (23%) and 10/62 (16%) mutations in the regions encoding amino acids 84-101, 48-60 and 147-155, respectively. However, the positions of the mutational clusters did not seem to show the same subtype dependency as mentioned previously, as the same clusters were found in both *adr* and *adw* patients.

A much larger study of core gene mutations in chronically infected patients was carried out by Akarca and Lok (1995a), who identified regions encoding core protein amino acids 59-66, 87-100, 125-135 and 147-155 as mutational clustering regions and regions encoding amino acids 14-25, 101-124, 136-146 and 156-183 as completely devoid of mutations. The relative absence of mutations leading to amino acid substitutions in the C-terminus of the core protein is probably due to the overlap of core and polymerase genes in this region: mutations here could possibly have a deleterious effect on the polymerase protein. Also, clusters of arginine residues in the C-terminus of the core protein are

required for its nucleic acid binding properties (see Section 1E1). It was therefore of interest that, in this region, only Arg-151 was mutated, and in some cases an arginine residue was restored at position 153. All cysteine residues, important for core particle assembly, were also conserved. Of additional interest was the fact that the mutation clustering regions overlapped with known epitopes in the protein. This linkage of increased mutations and Th and B-cell epitope regions was also identified in a smaller study of serial samples from Chinese chronic carriers (Boner *et al.*, 1995). Another region which did not correspond to any known epitopes, but showed higher variability than expected was identified between amino acids 21-40. Carman *et al.* (1995b) have also shown that mutations occur predominantly in both Th and B-cell epitopes. Both these papers also reported a low incidence of cysteine mutations or mutations in the C-terminal region with only two cases of mutated arginine residues.

2) Functional relevance of core mutations

An increase in core gene mutations leading to amino acid substitutions was found in patients with CAH when compared to asymptomatic carriers (Ehata *et al.*, 1992; Chuang *et al.*, 1993); in patients with fulminant hepatitis or with severe exacerbations, when compared to acute hepatitis cases (Ehata *et al.*, 1993); and when CAH cases were compared to those with only CPH (Ehata *et al.*, 1992; Akarca and Lok, 1995a). These associations suggest a role for core gene variation in determining the severity of an infection.

Selection of the pre-core stop mutant G1896A was also associated with increased core mutation (Akarca and Lok, 1995a; Boner *et al.*, 1995; Carman *et al.*, 1995b). This may be linked to seroconversion to anti-HBe antibody, as greater numbers of mutations have been reported in some patients who are anti-HBe positive, compared to those who are e antigen positive (Carman *et al.*, 1995b; Akarca and Lok, 1995a). It has also been observed that mutations do not develop at a constant rate, but are more frequent around the time of seroconversion (Akarca and Lok, 1995a) or occur at the same time as, or after, the selection of the pre-core stop codon (Boner *et al.*, 1995; Carman *et al.*, 1995b).

The relevance of core gene mutations which lead to amino acid substitutions may be connected to the importance of the core protein as an immune target. Despite the importance of CTL in clearing the virus, relatively few studies have examined any possible associations between core gene mutations and CTL epitopes and very few patients have been HLA-typed in any of these studies. Much attention has been focused on the well-defined HLA-A2 core epitope between amino acids 18-27. Carman *et al.* (1995b) showed that some patients with ongoing active disease after seroconversion had mutations within this epitope (at Ser-21 and Val-27). However, mutations in the epitope were also found in a patient who went into remission after seroconversion; not all patients with active disease after seroconversion had mutations in this region; and not all patients

studied were HLA-A2 positive. Ehata *et al.* (1992) identified a potential HLA-A2 epitope between amino acids 84-91, positioned in the mutation clustering region, although none of the patients were actually HLA-typed and this epitope has never been recognised before. Work by Bertoletti *et al.* (1994a, b) showed that mutations observed in the epitope between core protein residues 18-27 in two HLA-A2 positive chronic patients could inhibit the core-specific CTL response. Peptides containing the observed mutations at Ser-21 (S21N, S21A or S21V) or double mutations at Ser-21 and Val-27 (V27A or V27I) acted as T cell receptor antagonists for >80% of the CTL clones, when presented to CTL by the same target cell. Val-27 mutations also reduced the HLA binding affinity of the peptide. The CTL inhibition was observed at physiological levels of peptide and was not due to competition with wild-type peptides for HLA binding or to anergy induction. As infected cells harbouring these viral variants would presumably escape cell-mediated elimination, the variant would be allowed to propagate. However this scenario is restricted in providing an explanation for the relevance of viral variation in the development of chronic disease. The CTL response against this epitope in the two patients studied was vigorous, in contrast to the weak CTL response normally observed in chronic cases. It was also focused on one epitope, in contrast to the multispecific nature of the CTL response against HBV.

Other studies have examined core protein epitopes recognised by B- or Th cells, the latter being recognised seemingly irrespective of HLA class II-restriction (Ferrari *et al.*, 1991). Akarca and Lok (1995a) claimed that the mutation cluster regions observed overlapped with both Th and B- cell epitopes. However, no functional analysis of the effects of these mutations has been attempted. The clusters of mutations observed by Carman *et al.* (1995b) also correlated with the positions of Th and B- cell epitopes, but again, these mutations were not analysed functionally. As the relevance of anti-HBc antibody in chronic infection is not fully understood, the importance of mutations observed within B-cell epitopes is hard to determine, but mutations with Th epitopes could weaken, or even eliminate, the $CD4^+$ T cell response against core protein, with chronic hepatitis ensuing.

Another problem is relating viral variation to the ongoing disease observed in some patients. If variation within an epitope leads to non-recognition by the immune system, why do such patients have continuing disease? In addition, is the variation observed a result of immune pressure or does it evoke the immune response? The observation of mutation in the clustering region suggested by Ehata *et al.* (1992), after the ALT level had risen, supports the former possibility. However, the existence of a minor viral population having the same mutation before the ALT increase is also possible. It may also be the case that variation creates new targets for the immune system, thus evoking an ongoing, cyclic response.

1E CORE PROTEIN

1E1 Characterisation of the core protein

The discovery that expression of the core gene alone was sufficient to produce nucleocapsids, or core particles, antigenically (Pasek *et al.*, 1979; Burrell *et al.*, 1979) and morphologically (Cohen and Richmond, 1982) identical to the virion-derived nucleocapsid, initiated the search for the mechanism of core particle assembly and the amino acid sequence requirements. The ability of the core protein to self-assemble into core particles has now been demonstrated in mammalian cells (Roossinck and Siddiqui, 1987), insect cells (Lanford and Notvall, 1990), yeast (Miyanohara *et al.*, 1986), *Xenopus* oocytes (Zhou and Standring, 1991) and *in vitro* via association with a cytosolic chaperonin (Lingappa *et al.*, 1994).

The predicted core protein amino acid sequence, derived from cloned HBV DNA, gave a 183 amino acid (185 for subtype *adw* (Ono *et al.*, 1983)) protein with a highly basic C-terminus (see Figure 1E1). In this region (amino acids 150-183), 16 out of the 34 residues are arginines, arranged in four clusters of three or four arginines, including three repeats of the motif SPR_4. The obvious prediction for the function of this region was nucleic acid binding. In addition, the repeated SPXX sequence is known to be a DNA-binding motif (Suzuki, 1989).

The C-terminal limit for core particle assembly was investigated by a number of groups using C-terminally truncated core proteins in a variety of systems (Birnbaum and Nassal, 1990; Hatton *et al.*, 1992; Nassal, 1992a; Beames and Lanford, 1993), as earlier experiments had indicated that the arginine-rich region was dispensable (Gallina *et al.*, 1989). All showed that residues C-terminal to Pro-144 were not required but truncation N-terminal to Leu-140 abolished particle production. Later studies showed that truncation C-terminal to Leu-140 allowed assembly (Zlotnick et al., 1996).

To test the predicted nucleic acid binding function of the C-terminal arginine-rich region, packaging of nucleic acids into particles formed from C-terminally truncated proteins was also assayed. Truncation after Arg-164 allowed the same amount of RNA packaging as full-length protein, showing the C-terminal nineteen amino acids to be dispensable (Birnbaum and Nassal, 1990; Nassal, 1992a). A different set of truncations were made by Hatton *et al.* (1992) in an *adw* subtype core protein which sequentially removed groups of arginine residues. Truncation after residues 172, 162 or 157 all decreased RNA packaging by only two-fold, whereas truncation after Val-149 (Δ149) resulted in a twenty-fold decrease, suggesting that the first group of three arginine residues (150-152) was important. This was confirmed when the addition of four arginine residues to the Δ149 truncation restored packaging levels to only two-fold less than for full-length protein. Examination of DNA-binding and subsequent DNA synthesis showed

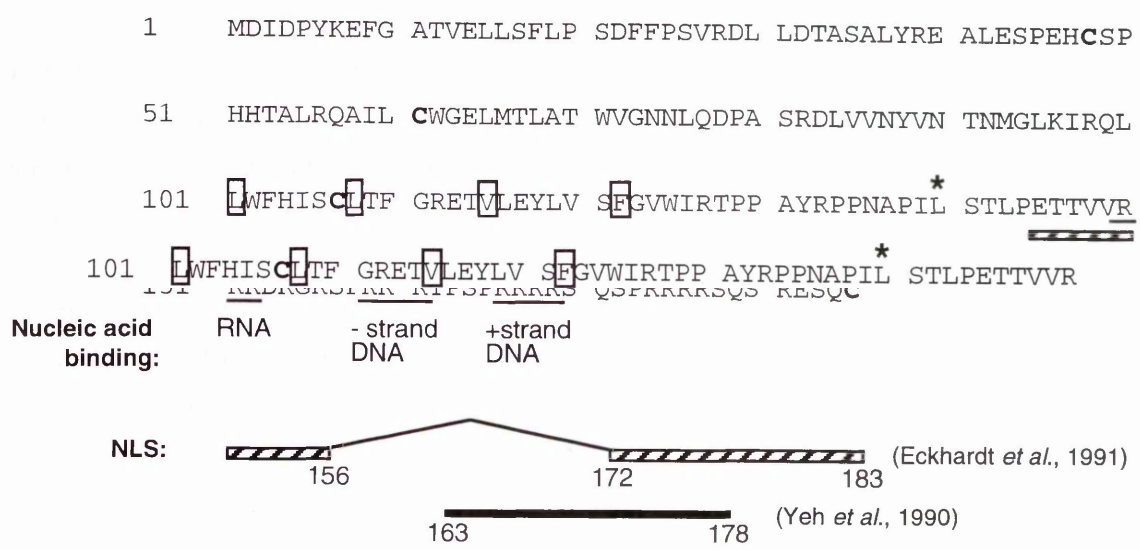


Figure 1E1: Functionally important regions of the core protein
The amino acid sequence of the core protein (subtype *adw*) is shown: other subtypes are missing the D and R residues at positions 153 and 154. The asterisk denotes the position of the maximal C-terminal truncations which still allow core particle assembly. Cysteine residues are shown in bold and residues comprising the hydrophobic heptad repeat element are boxed. Groups of arginine residues involved in nucleic acid binding are underlined and labelled, while the nuclear localisation signals (NLS) mapped by two groups are also indicated.

different requirements in the core protein: core proteins truncated beyond residue 157 did not bind DNA in South-western blots but the Δ162 truncation showed a high affinity for DNA, thus identifying residues 157-162 as the essential region for DNA binding. Interestingly this contains another group of arginine residues in an SPRR motif. As the Δ162 and Δ172 truncations showed decreased DNA binding compared to full-length protein, the remaining two groups of arginines must also contribute. A model was proposed whereby the groups of arginines at 150-152, 159-161 and 166-169 were required for binding RNA, minus-strand DNA and plus-strand DNA, respectively, with the final group being redundant. This model was supported in experiments where strand-specific probes were used to demonstrate that the Δ164 truncation showed a defect in plus-strand DNA synthesis, possibly at the oligoribonucleotide primer translocation stage (Nassal, 1992a). In this genomic context, amino acid sequences beyond residue 164 were also non-essential for the formation of enveloped virions. This has been recently supported by results showing that only the initiation of minus-strand DNA synthesis is necessary for envelopment to occur (Gerelsaikhan *et al.*, 1996).

The nucleic acid-binding properties of core protein fulfil two functions: the protein:nucleic acid interactions act to stabilise the particle structure and to condense the DNA into its most compact form. The latter function has been proposed due to the similarities between the core protein C-terminal amino acid sequence and cellular protamines, found in sperm.

How is this stable interaction eliminated to allow disassembly of the core particle during natural infection? As a virus-associated kinase (now known to be protein kinase C (Kann *et al.*, 1993)) had been discovered with specificity for serine residues (Gerlich *et al.*, 1982; Roossinck and Siddiqui, 1987), modification of the protein:nucleic acid interaction by phosphorylation of C-terminal serines seemed the most likely mechanism. Use of a phosphorylation state-dependent antibody against an epitope between amino acids 165-175 showed that Ser-168 and/or Ser-170 were phosphorylated *in vivo* (Machida *et al.*, 1991). In addition, core protein synthesised in *Xenopus* oocytes showed no detectable DNA binding, in contrast to bacterially-expressed protein, the only difference being phosphorylation of the C-terminal region of the former protein (Hatton *et al.*, 1992). In *in vitro* binding experiments, phosphorylated core protein showed a decreased affinity for DNA (Machida *et al.*, 1991).

1E2 Assembly of core particles

Much of the information regarding the mechanism of core particle assembly has come from studies in *Xenopus* oocytes by the group of Standring. A population of 'free' core protein is the direct precursor of core particles, which self-assemble once a critical threshold level of 'free' core protein is reached (Zhou *et al.*, 1992). Later work showed that the 'free' core protein population consisted mainly of disulphide-linked core protein dimers, but not of higher order structures (Zhou and Standring, 1992a). Particle assembly is a co-operative process which initiates once the concentration of core protein reaches an estimated 0.8μM and proceeds without the detection of any intermediates (Seifer *et al.*, 1993). Once formed, the particles showed an increased protease resistance compared to core protein dimers. As the C-terminal 44 amino acids are thought to be the major protease target it suggests that this region is located inside the particle, (Zhou *et al.*, 1992; Seifer and Standring, 1994).

In the *Xenopus* system, removal of the nucleus did not affect particle production, which is in keeping with the observation that the majority of free core protein is also located in the cytoplasmic compartment (Zhou and Standring, 1991). However opinions differ with regard to this subject (see Section 1E4).

Due to their conservation among all mammalian hepadnaviruses, the function of the four cysteine residues in the core protein was examined in site-directed mutagenesis experiments. Replacement of all cysteine residues with serine or alanine had no detectable effect on dimer formation, particle assembly or envelopment (Nassal, 1992b; Nassal *et*

al., 1992; Zhou and Standring, 1992b). In addition, such particles were antigenically identical to wild-type particles (Nassal *et al.*, 1992). However, dimers and particles formed from the Cys-minus core protein did show a decreased stability (Zhou and Standring, 1992b). Therefore, these residues probably act to lock dimers and particles into their conformations and positions during particle assembly. Detailed analyses of disulphide bonding patterns and free thiol groups have shown that Cys-61 participates in a disulphide bond to the corresponding residue in another core protein subunit, whereas Cys-107 does not, due to its buried position in the folded core protein (Zheng *et al.*, 1992; Nassal *et al.*, 1992). Cys-48 partially forms a disulphide bond, but also exists as a free sulphydryl group. Cys-183 does participate in disulphide bond with other Cys-183 residues, but it is not clear as to whether these bonds form between core protein dimers (Nassal *et al.*, 1992; Zheng *et al.*, 1992) or between core protein monomers in one dimer (Seifer and Standring, 1994). In the virion, the free Cys-48 may be important for binding to one of the cysteine residues in SHBs, acting to link the core particle to the envelope layer. Lack of this residue in particular may have accounted for the slightly decreased envelopment of Cys-minus core particles observed by Nassal (1992b).

Although experiments in HepG2 cells using a modified yeast two-hybrid system have shown that core proteins from WHV and HBV (which differ at 46 residues) can hetero-oligomerise, subsequent co-expression of full-length and C-terminally truncated HBV core proteins in *Xenopus* oocytes showed that dimer assembly was not a non-specific process (Chang *et al.*, 1994). In these co-expression experiments, homodimers formed preferentially, even although the smaller protein was missing only 26 residues. As core protein monomers were undetected, dimerisation must occur rapidly after synthesis, with the core protein most likely associating with the nearest monomer. This is probably produced from the same polysome, thus giving rise to the observed *cis*-preferential nature of dimerisation.

Relatively little attention has been paid to the sequence requirements for particle assembly in the N-terminal and central regions of the protein. Analysis of particle assembly using twelve mutants of the larger and less homologous DHBV core protein showed that three to four amino acid insertions in the N-terminus and deletions of C-terminal amino acids were still assembly-competent. However N-terminal or central deletions, or a four amino acid central insertion eliminated particle assembly (Yang *et al.*, 1994). Similar studies with two to four amino acid linker insertion mutants throughout the core gene showed regions which were essential for assembly (Beames and Lanford, 1995). Core proteins with insertions after positions 36, 44, 66 and 89 all formed particles when expressed in insect cells, as did a K96R point mutant. However, when expressed in Huh7 cells, core protein with an insertion after position 69 also formed particles. When supplied in *trans* to transfected core-defective genomes, only core proteins with an insertion after residue 44 and the K96R mutant were able to encapsidate HBV RNA and

show DNA synthesis in the endogenous polymerase assay, albeit at lower levels than wild-type core protein. These results showed that even though large insertions can be tolerated around residue 78 (See Section 1E5), much smaller insertions in both termini and in the central region of the protein can disrupt particle assembly, or other stages in the viral life cycle.

The central region of WHV core protein contains a hydrophobic heptad repeat element: L101, L108, V115 and F122. These residues are conserved among the mammalian and avian hepadnaviruses (Yu *et al.*, 1996). Mutagenesis experiments showed that changing the regular spacing of these residues abolished particle assembly, as did the introduction of a proline residue by mutagenesis of any amino acid in this region. This indicated that the secondary structure of this region was important, and probably α-helical, with the hydrophobic residues on the outer surface of the helix. This would provide a hydrophobic surface accessible to the equivalent region in another monomer, allowing the two proteins to interact in a manner analogous to leucine zipper-containing proteins. Although four residues are conserved in the mammalian protein (five for the larger DHBV protein), mutation of any one of these residues did not affect particle assembly. The conservation of four residues may be due to the increased stability this would give to the protein:protein interaction. Unfortunately the previous insertion mutants generated by Beames and Lanford did not cover this region, so correlations could not be made. Interestingly, Akarca and Lok (1995a) concluded that the region between amino acids 101-124 was devoid of mutations. In other studies, mutation of the heptad repeat residues was rare and tended to be conservative: L101I (Ehata *et al.*, 1992) and L108I (Carman *et al.*, 1995b) have been observed. Mutations producing proline residues in this region are also very rare.

1E3 Three-dimensional structural analyses of core particles

Until recently, very little information was available concerning both the tertiary structure of the core protein and the orientation of the protein, as a dimer, in the core particle. Most of this information had come from direct biochemical studies on core particles along with indirect information obtained though the use of core particles as carriers of foreign epitopes (see Section 1E5).

Surface-exposed regions of core protein that bound to monoclonal antibodies were mapped by Pushko *et al.* (1994) and are shown in Figure 1E2. Some antibodies recognised only core particles in solution, but not denatured protein or synthetic core peptides (in a competition ELISA). This confirmed the existence of conformational core epitopes. These antibodies did, however, recognise immobilised peptide corresponding to amino acids 74-90, which may be a linear component of a larger discontinuous epitope, as was previously suggested (Sallberg et al., 1991a). This region is therefore likely to be exposed at the surface of the particle. The region between amino acids 125-133 was also

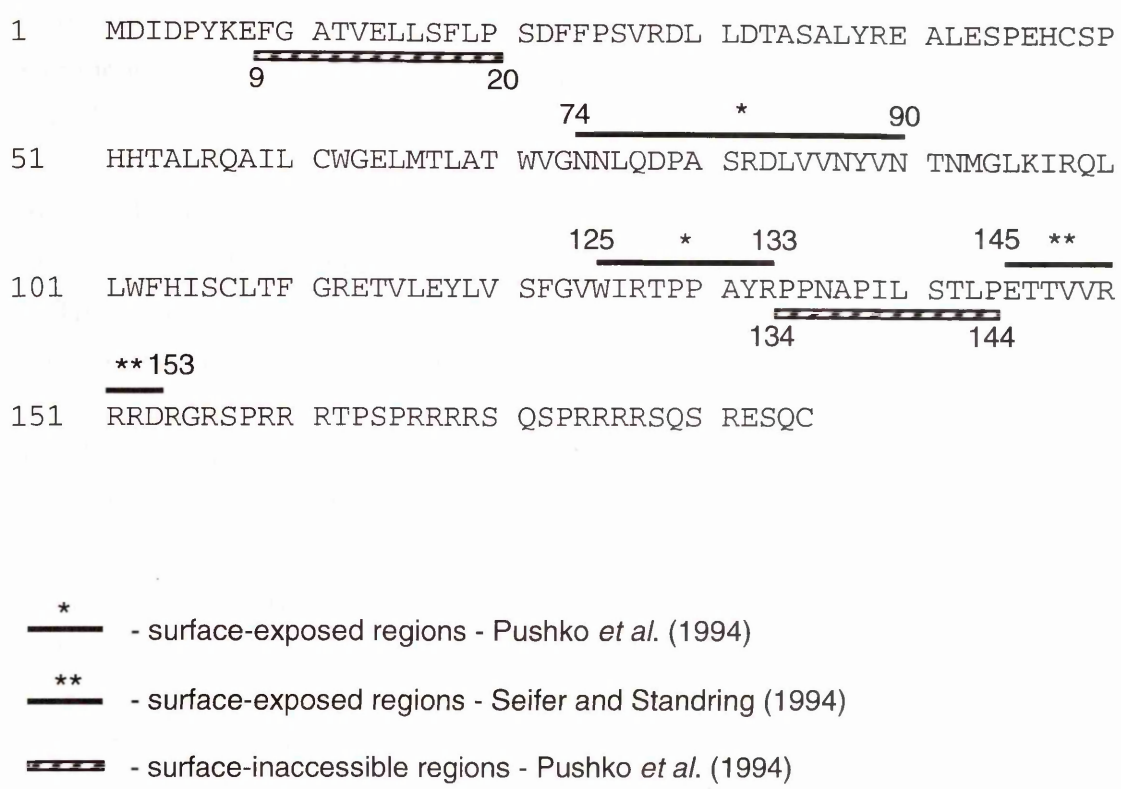


Figure 1E2: Surface-exposed regions of the core protein in core particles
The amino acid sequence of the core protein is shown above. Lines above sequences indicate that these regions were identified as being exposed on the particle surface by the workers indicated. Hatched lines below sequences indicate that these regions were not exposed at the surface.

surface-exposed, but the adjacent region from 134-144 had only low-level accessibility. Two proline residues at positions 134 and 135 possibly mediate this turn in the polypeptide backbone. Residues 9-20 were also surface-inaccessible and this N-terminal region may be important for assembly (Khudyakov *et al.*, 1991).

Protease studies of core particles by Seifer and Standring (1994) showed that the C-terminal region between amino acids 145-153 was the only accessible target (see Figure 1E2). This sequence was proposed to act as a hinge between the C-terminal domain and the N-terminal region essential for assembly. The basic C-terminal region was sensitive to proteolysis when in the form of dimers, but not when in particles, suggesting that it was tethered inside the core particle. This retention was shown to be mediated by the inter-dimer disulphide bond formed by Cys-185, as increased surface exposure of the C-terminal domain was observed when this residue was deleted or the protein truncated after amino acid 172. The latter truncated protein was also capable of packaging RNA (Hatton

et al., 1992), suggesting that a dynamic equilibrium exists between the interior and exterior localisation of the C-terminal region. Despite the seemingly conflicting results from these two studies, the polypeptide chain may loop around between amino acids 134 and 144 and reappear on the core particle surface again, before once again burying itself after residue 153. As proteases with targets between residues 134 and 144 were not used it was not possible to test this theory.

These biochemical studies have been dwarfed by the results generated by the use of electron cryomicroscopy and image reconstruction in the study of the core particle structure. Early electron microscopic studies of liver-derived core particles had shown that particles were of 28nm external diameter and 21nm internal diameter, with a 3.5nm thick shell (Onodera *et al.*, 1982). The particle was predicted to be icosahedral in shape with approximately 180 component subunits and a T=3 morphology. Electron cryomicroscopic studies on bacterially-expressed core particles by Crowther *et al.* (1994) showed two distinct sizes of particles with diameters of 30 and 34nm, with the larger particles forming 85% of the population. The two different particle sizes were also observed in particles assembled from core protein C-terminally truncated after amino acid 144 (Δ144 truncation). Protein spikes were also visible projecting from the surface of the shells, giving overall diameters of 32 and 36nm for small and large particles, respectively. Particles formed from the expression of full-length protein were either empty or had internal contents, presumably RNA molecules. The Δ144 truncation particles appeared empty and contained <1% RNA by weight, compared to approximately 10% for the full-length protein particles.

Image reconstructions showed that the small particles consisted of 180 subunits, arranged with T=3 quasiequivalence and the large particles contained 240 subunits with T=4 quasiequivalence. In both types of particle, the subunits were dimer clustered and thought to have an inverted 'T' shape, with the 'handle' forming the spiked projections on the particle surface and the 'head' forming the shell of the particle. The mechanism of selection of the T=3 or T=4 morphology was not due to protein:nucleic acid interactions, as the Δ144 truncation protein also formed both types of particle. The C-terminal domain of full-length protein did seem to contact packaged RNA, as the RNA was ordered in an icosahedral orientation by this interaction. This allowed it to be detected and mapped. In empty particles no internal structure was observed, due to the absence of RNA and the disordered nature of the C-terminal protein domain. In both sizes of particles the shell was penetrated by 1.3 and 2nm holes (110 in the T=4 particles and 80 in the T=3), presumably allowing the influx of nucleotides for reverse transcription and DNA replication and the efflux of digested RNA. The two sizes were also present in the liver, as detected by electron microscopy (Cohen and Richmond, 1982). Although T=4 particles were more prevalent in bacterially expressed samples (Crowther *et al.*, 1994), only recently has it been demonstrated that the same is true for liver-derived particles, with an

even higher prevalence of the larger particles (Kenney *et al.*, 1995). These liver-derived particles also contained more RNA, probably due to the increased efficiency of the encapsidation process in infected cells.

Further studies provided evidence for the importance of the C-terminus of the core protein and possibly the disulphide bonds in determining the particle morphology (Zlotnick *et al.*, 1996). Progressive truncation of the region between amino acids 149 and 140 decreased the proportion of T=4 particles produced from >90% to only 15% and, in addition, more T=3 particles were formed when a Cys61Ala mutation was created in the Δ149 truncation. The relevance of these results to the *in vivo* situation is not immediately clear.

The most recent structural studies have increased the core particle image resolution from the original limit of approximately 30Å to 9Å (Conway *et al.*, 1997) or 7.4Å (Böttcher *et al.*, 1997; see Figure 1E3). The information made available by these studies, not only as regarding the overall structure of the core particle, but even concerning the exact conformation of the polypeptide backbone in the assembled particle, is invaluable. Prior to these electron cryomicroscopy studies, the only model had been theoretical (Argos and Fuller, 1988), based on the homology of WHV core protein with the Mengo virus vp3 capsid protein, which has an eight-stranded anti-parallel β-barrel conformation, common to capsid proteins. The resulting model showed a protein with nine β-sheets, comprising 48% of the polypeptide backbone, which concurred with the accessibility of the HBe1 epitope and the inaccessibility of the HBe2 epitope; the requirement of the N-terminal 140 amino acids for particle assembly; and the internal orientation of the phosphorylation sites within core protein. However, later studies on bacterially expressed particles using several biophysical methods showed that, in contrast to the Argos and Fuller model, core protein had a high (43-60%) α-helical content, whether in the form of dimers or particles (Wingfield *et al.*, 1995). The β-sheet content of the protein was very low (5-35%), suggesting that core protein may not adopt the canonical eight-stranded anti-parallel β-barrel conformation which all, but one, icosahedral viral capsids exhibit.

The new structural information did show that the HBV core protein conformation is structurally different to the standard isocahedral capsid protein: a new four-helix bundle viral capsid fold existed in the core particle, with neighbouring polypeptide chains contributing an α-helix pair to each bundle (Böttcher *et al.*, 1997; Conway *et al.*, 1997; see Figure 1E4). Core protein itself also contains four helices: helices 2 and 3 project out from the shell, while the other two contribute to the shell structure. In the model of Böttcher *et al.*, helix 2 was predicted to run from residue 50-78, halted by a conserved proline at position 79, with residues 78-82 constituting the exposed tip of the spike (see Figure 1E5). However, Conway *et al.* placed these residues at the base of helix 2. According to Böttcher *et al.*, helix 3 then runs from residue 82-110, with a kink approximately halfway down corresponding to the conserved Gly-94 residue. The

Figure 1E3: Surface representation of the T=4 core particle
The 310Å diameter particle is viewed down a 2-fold axis of symmetry (Scale bar = 50Å). The structure was determined at 9Å resolution by electron cryomicroscopy. (This image was kindly provided by Dr. J. Conway, NIH, Bethesda)

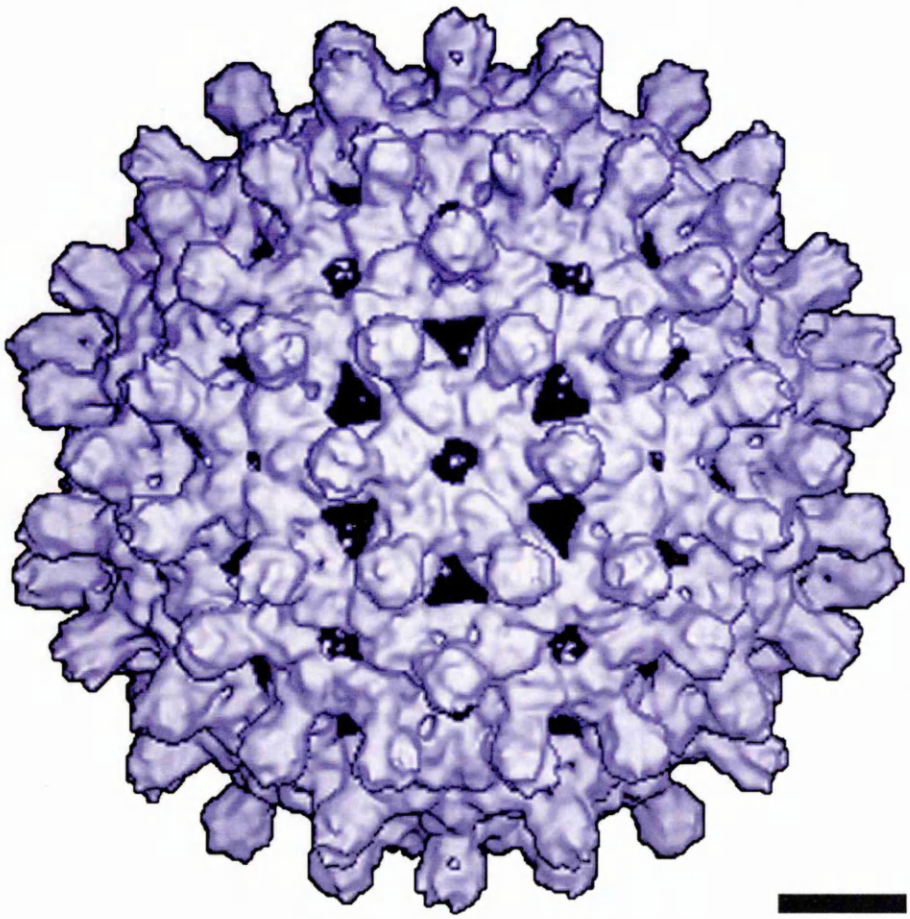
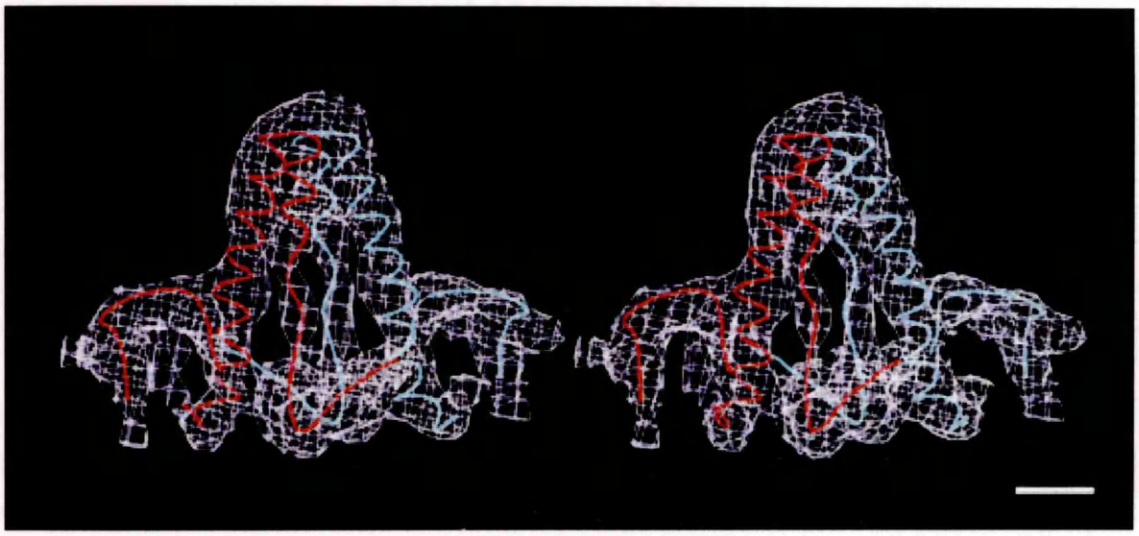


Figure 1E4: Stereo pair of density maps with 'wire' model of the core protein dimer
The core protein dimer is viewed tangentially to the particle surface. Subunits are shown in red and blue. Scale bar represents 10Å. (This image was kindly provided by Dr. J. Conway, NIH, Bethesda)



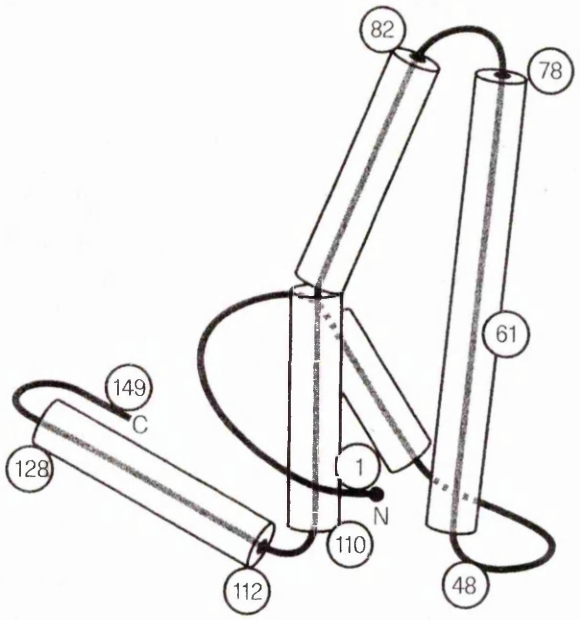


Figure 1E5: Polypeptide fold of the core protein
The four α-helices of core protein are shown as cylinders and the putative N and C termini are marked. In the text, helices are numbered 1 to 4, from the N- to the C-terminus. An approximate numbering scheme for amino acids is also indicated. (Reproduced from Böttcher *et al.* (1997) by kind permission of Dr. R. Crowther)

C-terminal helix runs from 112-128 and is also terminated by a proline residue at position 129. The proposed exposure of the region between 125-133 also concurs with the model as this region lies at the end of the C-terminal helix, protruding on the surface of the particle. The models also show Cys-48 and Cys-61 to be present on the outer surface of helix 2, making them available for the disulphide bonding necessary for dimerisation. Cys-107 was shown to be buried within the protein. As these models were determined using C-terminally truncated proteins, no position was determined for Cys-183. However, as the C-terminus appears to have no defined structure, this residue is likely to be available for the formation of intra-dimer or inter-dimer disulphide bonds. The previously described hydrophobic heptad repeat element would be partitioned between helix 3 and the C-terminal helix with an intervening short looped region. It may be, therefore, that this element does not consist of a single continuous helix, as was originally thought, but contributes to two separate helices. Hopefully the increase in resolution that should come from further electron cryomicroscopic studies, possibly even to atomic level, will help in answering these questions.

1E4 Subcellular Localisation of Core Protein

Immunohistochemical and immunofluorescence studies of liver biopsy samples from patients suffering from chronic hepatitis B have shown that core protein can be detected in the nucleus, cytoplasm or both (Gudat and Bianchi, 1977; Gowans *et al.*, 1985; Ulrich *et al.*, 1985; Chu and Liaw, 1987; Hsu *et al.*, 1987; Chu *et al.*, 1995). Similar results have been obtained for cell lines such as NIH 3T3 (Eckhardt *et al.*, 1991; Yeh *et al.*, 1993), Vero-C3 (Yeh *et al.*, 1993), COS 7 (Yeh *et al.*, 1990) and in HBV genome (Farza *et al.*, 1988) or core gene (Guidotti *et al.*, 1994a) transgenic mice.

Studies using both COS 7 (Yeh *et al.*, 1990) and NIH 3T3 (Eckhardt *et al.*, 1991) cells have identified nuclear localisation signals (NLSs) in the C-terminal region of the core protein (see Figure 1E1). The clusters of arginine residues in this region resemble the stretches of consecutive basic amino acids observed to function as NLSs in other proteins such as the influenza virus NS1 protein (Greenspan *et al.*, 1988), adenovirus E1a protein (Lyons *et al.*, 1987), yeast histone 2B (Moreland *et al.*, 1987) and nucleoplasmin (Dingwall *et al.*, 1988).

Deletion analysis of the C-terminal arginine-rich clusters demonstrated the importance of residues 145-156 for nuclear localisation, as removal of the region C-terminal to residue 144 caused this protein to accumulate in the cytoplasm (Eckhardt *et al.*, 1991), The possibility that this localisation was mediated by arginine residues between 150-154 and the particular importance of the second residue in this cluster, as has been shown for the corresponding residue in the NLS of SV40 large T antigen (Kalderon *et al.*, 1984), were examined by mutagenesis of this Arg-151 residue to a threonine. In the context of the full-length protein, cellular localisation was not affected by this mutation. However, when between 12-39 C-terminal residues were deleted, Thr-151 core protein localised to the cytoplasm, whereas the Arg-151 protein was located in the nucleus. These results also suggested the presence of a second NLS in the C-terminal 12 residues of the core protein (172-183), which also included an arginine cluster (amino acids 172-175). As yet it is not known whether these signals can function independently, although the presence of two NLSs in the same protein has been reported previously (Greenspan *et al.*, 1988).

Yeh *et al.* (1990) identified a C-terminal core protein NLS by examining the cellular localisation of human α-globin:HBV core fusion proteins in transfected COS 7 cells. The location of the NLS was finely mapped by demonstrating that core protein residues from Pro-163 to Ser-178 were necessary and sufficient to translocate the α-globin protein to the nucleus. This region contains two repeats of the sequence PRRRRSQS and is completely conserved in both WHV and GSHV.

Most studies agree that core protein detected in both the nucleus and the cytoplasm is in the form of core particles, as confirmed by electron microscopic examination (Yamada and Nakahane, 1977; Michalak and Nowoslawski, 1982; Sakamoto *et al.*, 1983; Guidotti *et al.*, 1994a). However, the results of Chu and Liaw (1992), obtained by the use of

monoclonal antibodies specific for either the c or e antigenic determinants on core protein or particles, suggested that whereas particles may predominate in the nucleus, non-particulate core protein was more prevalent in the cytoplasm. This was contradicted by Bock *et al.* (1996). who could not detect particles within the nuclei or at the nuclear membrane.

The function of the core protein NLS in the viral lifecycle is unclear as yet, although several possibilities exist: 1) transport of the uncoated core particle to the nucleus after viral entry, to deliver the genome; 2) directing newly synthesised core protein dimers or particles to the nucleus in order to allow pgRNA encapsidation into core particles; 3) directing newly assembled, genome-containing core particles to the nucleus to allow intracellular genomic amplification. As the NLS overlaps with the core protein nucleic acid binding region, it may not be accessible on the particle surface and therefore not detected by the cytoplasmic NLS-recognising proteins. The same problem exists if the NLS is involved in the third possibility listed above. It may be that some core protein C-terminal tails are not tethered in the particle interior and are surface-accessible, either because they are not required for nucleic acid binding or due to the lack of a disulphide bond between Cys-183 residues in the reducing environment of the cell. Immunoelectron microscopic examination of infected hepatocytes (Yamada and Nakahane, 1977) and an HBV-producing cell line (Bock *et al.*, 1996) showed core protein localised around the nuclear membrane and especially near nuclear pores, suggesting that the transfer of contents between core particles and the nucleus does occur.

An additional focus on the investigation of core protein localisation has attempted to relate this to the disease severity. Several studies have concluded that purely nuclear localisation of core protein, as detected mainly by immunostaining or immunofluorescence of liver biopsy samples, is more prevalent in CPH and occurs rarely in cases of CAH (Hsu *et al.*, 1987; Kojima *et al.*, 1987; Yoo *et al.*, 1987; Chu and Liaw, 1987, 1992; Chu *et al.*, 1995). Conversely, the localisation of core protein to the cytoplasm or plasma membrane was increased in CAH cases from these studies. *In situ* hybridisation in infected hepatocytes demonstrated that intracellular levels of cytoplasmic HBV replicative DNA correlated with the level of cytoplasmic core protein, but not with the presence or absence of nuclear core protein (Gowans *et al.*, 1985). This suggests that cytoplasmic core protein is indicative of active replication In addition it was thought that the shift in localisation of core protein, especially if it led to increased plasma membrane expression, could lead to an increased immune attack on these cells. Both of these scenarios could explain the more aggressive disease connected with the cytoplasmic location of the protein.

However, more recent experiments in both transfected cell lines and liver biopsy samples have shown that, rather than the shift in localisation being responsible for the worsened disease state, this subcellular shift may be a secondary effect, caused by the

regeneration of surviving hepatocytes after immune-mediated killing of infected cells.

Yeh *et al.* (1993) showed that the nuclear localisation of core protein was subject to regulation by the cell cycle in both NIH 3T3 and Vero-C3 cells. Cells were synchronised in either the G_0 phase (by using serum-free medium) or at the G_1/S phase boundary (by using the DNA polymerase inhibitor, aphidicolin) and core protein localisation determined at various time points after cells were allowed to re-enter the cycle. The results from both subcellular fractionation and immunofluorescence experiments showed an increase in the amount of core protein in the nucleus during the G_1 phase. Nuclear core protein decreased during S phase to undetectable levels, with a concomitant increase in the cytoplasmic levels. Levels of nuclear core increased again during G_2 phase. Double immunostaining for proliferating cell nuclear antigen (PCNA) and core protein in biopsy samples showed the same pattern (Chu *et al.*, 1995). Approximately 50% of hepatocytes containing cytoplasmic core protein also stained positive for PCNA, a reliable marker for proliferating cells, whereas this antigen was rarely detected in cells with nuclear core protein. Although both sets of results are consistent with disease severity being linked to altered core protein localisation, a mechanism for the cell cycle dependency is still unknown. Yeh *et al.* (1993) observed that cytoplasmic but not nuclear core protein was phosphorylated. The C-terminus contains three S/TPXK/R recognition motifs for $p34^{cdc2}$ kinase, which is activated during the G_1 to S transition of the cell cycle. As this kinase can phosphorylate core protein *in vitro* (unpublished observations quoted by Yeh *et al.*, 1993), this may be relevant.

1E5 Structural Flexibility of Core Particle Assembly

The ability of the core particle to be correctly self-assembled when expressed in a variety of systems, coupled with the presence of CTL, Th and B-cell epitopes on the protein and particle, make it suitable as a carrier for various foreign epitopes. Many chimeric core proteins, containing foreign epitopes, have been assayed for antigenicity. These studies are detailed in a recent review by Pumpens *et al.* (1995). These studies indicate a tremendous 'structural flexibility' of the core protein, shown by the range of alterations that can be made without abrogating particle assembly.

Foreign sequences from a variety of sources have been inserted at three major regions of the protein: the N-terminus (including the pre-core region), the C-terminus of full-length or truncated core protein and centrally in the region around amino acid 78.

N-terminal insertions of 50 (human chorionic gonadotrophin), 41 and 27 (HBV pre-S1) amino acids have been made after pre-core residues Leu-24, Trp-26 and the core start codon, respectively (Beesley *et al.*, 1990; Schödel *et al.*, 1992).

Internal insertions have been made in the central region containing the main B-cell epitope, as this was predicted to be exposed (Argos and Fuller, 1988). Insertions of up to 27 amino acids of the SIV envelope protein were possible in full-length core protein,

between Ser-81 and Arg-82 (Yon *et al.*, 1992) and of up to 39 amino acids from HBV SHBs in a C-terminally truncated protein, replacing core residues 79-81, although replacement of residues 76-82 has also been successful (Schödel *et al.*, 1992).

Insertion of up to 53 amino acids of HBV pre-S2 sequence was possible between Pro-144 and Glu-145, at the C-terminus (Borisova *et al.*, 1989), with smaller insertions between Arg-179/Glu-180 (del Val *et al.*, 1991) and after the C-terminal cysteine residue (Yon *et al.*, 1992). However perhaps the most impressive successful chimeric proteins were produced when multiple copies of the 180 amino acid long HCV core protein were added to the C-terminus of core protein truncated at residue 149 (Yoshikawa *et al.*, 1993). Fusion of linker sequences plus either three or four copies of the HCV core sequence, to produce insertions of 559 and 741 amino acids respectively, maintained the ability to correctly assemble into core particles. Such large C-terminal insertions are possible due to the probable location of the foreign sequences in the interior of the particle. They may even be exposed on the particle surface as they are detectable by ELISA (Yoshikawa *et al.*, 1993). The same mechanism may apply to the N-terminal insertion of foreign sequences. However, the proven ability of core protein to tolerate deletions of approximately 25% of its length, or to accommodate internal insertions of approximately 20% of its size without protein folding or particle assembly being affected is quite remarkable. The structural flexibility of this protein may have a bearing on the relevance of naturally-occurring HBV variants with deletions of core gene regions.

1E6 HBV Deletion Variants

In addition to variation in the amino acid sequence of the HBV proteins resulting from nucleotide substitutions caused by the pol/RTase infidelity, variation on a larger scale, in the form of insertion and deletion of sequences, has also been reported. This section will describe some of the deletion variants reported, along with any significance that has been attributed to them, and then focus on core gene deletion variants.

1) Deletion variants of the polymerase, surface and X genes

Although naturally occurring HBV isolates with amino acid variation in the polymerase gene have been reported, as yet no studies have shown deletions in genomic regions containing only polymerase sequences. Studies of either core (Okamoto *et al.*, 1987b), X (Feitelson *et al.*, 1995) or surface (Yamamoto *et al.*, 1994; Xu and Yen, 1996) genes have all reported deletions which would affect the N- or C-termini or internal regions of the polymerase protein, respectively. These deletions will be discussed under the headings of the specific gene being studied at the time of their discovery. *In vitro* mutational analysis of the polymerase gene did show that the spacer region, from amino acids 178-336, could be largely deleted (Δ201-292) with little effect on function (Radziwill *et al.*, 1990).

Many of the reported deletions in the surface gene occur in the pre-S regions, which overlap with the polymerase spacer region and so would probably not affect the function of the polymerase protein (Santantonio *et al.*, 1992; Yamamoto *et al.*, 1994). However the viability of these mutants has not been formally evaluated. This is especially important for deletions in the pre-S1 region, as the corresponding region in the LHBs protein is implicated in binding to hepatocyte plasma membrane receptors. This also applies to large variations in the surface gene sequence, such as the mutations leading to the insertion of three or eight amino acids between Thr-123 and Cys-124 (Yamamoto *et al.*, 1994; Hou *et al.*, 1995), as this protein is also important for hepatocyte binding.

Functional analysis has been performed on an in-frame deletion of 129bp in the 3' pre-S1 region, which removed the CCAAT element from the SpII promoter (Xu and Yen, 1996). When Huh-7 cells were transfected with a corresponding genomic construct, only very small amounts of 2.1kb surface transcripts were produced, with resulting low levels of M and SHBs proteins. As this deletion also included the negative regulatory region for the SpI promoter, increased levels of 2.4kb pre-S1 transcripts and LHBs protein were observed. As LHBs is incompetent for secretion, its overexpression caused the accumulation of 20nm particles within dilated perinuclear vesicles, giving the transfected cells the appearance of ground-glass cells which are commonly observed in HBV carriers (Hadziyannis *et al.*, 1973). Virion production from the transfected cells was also greater than that observed for a wild-type control, suggesting that not only does this type of deletion have a pathological relevance, but may also confer the viral variant with a growth advantage.

Analysis of large variations in the X gene is complicated by the fact that the precise role of this protein in infection is unknown, making determination of their functional relevance difficult. This gene also overlaps with important transcriptional regulatory elements, such as enhancer II, core upstream regulatory sequence and negative regulatory element, core promoter and DR1/2. In renal dialysis patients, X gene deletions resulted in removal of the pre-core start codon, DR2, core gene regulatory elements and also either removed sequences at the 3' end of the polymerase gene or produced possible X-core fusion proteins (Feitelson *et al.*, 1995). Analysis of the same region, but concentrating on the relevance of variation to the enhancer II/core promoter, showed that major viral populations had deletions, insertions and/or substitutions (Günther *et al.*, 1996a). All but two variations created an HNF1 site or potential HNF3 site. All variations decreased steady-state pre-core mRNA levels and increased pgRNA levels, with a consequent decrease in e antigen production and increased intracellular core and polymerase proteins, replicative intermediates and virion production.

2) Core gene deletions

Many groups have now reported the detection of deletions within the core gene, which are detailed in Figure 1E6. In general, the deletions cluster in the central region of the gene, around nucleotide 2200, but some groups have reported deletions which localise to the 5' end of the gene, either exclusively (Valliammai *et al.*, 1995; Zoulim *et al.*, 1996) or as one of two deletions in the gene (Fiordalisi *et al.*, 1994; Akarca and Lok, 1995b; Günther *et al.*, 1996b; Marinos *et al.*, 1996; Zoulim *et al.*, 1996). Some deletions extended far enough in the 3' direction to remove the polymerase ATG codon and presumably do not allow production of a functional polymerase protein (Okamoto *et al.*, 1987b; Akarca and Lok, 1995b; Marinos *et al.*, 1996; Günther *et al.*, 1996b). In contrast to these cases with mixtures of polymerase defective and non-defective genomes, Uchida *et al.* (1994) reported a single polymerase defective deletion, which presumably represented a non-viable virus. However viral DNA containing no evidence of core gene deletions was also isolated from this patient, thus providing the requisite polymerase, in a manner analogous to the *trans*-complementation demonstrated *in vitro* (Okamoto *et al.*, 1993). This scenario is common to all but two reported cases of core gene deletions: full-length virus can also be detected, either by PCR or sequencing of cloned PCR products. The only exceptions to this are of a 33 nucleotide deletion observed in clones sequenced from a patient with no anti-HBc antibodies (Valliammai *et al.*, 1995) and an unspecified deletion present in all six genomic clones analysed from an immunosuppressed renal transplant patient (Günther *et al.*, 1995). The former deletion removed nucleotides 2021-2053, corresponding to core protein amino acids 41-51. Either the HBV viruses from these two cases were able to produce functionally normal, but structurally altered, core particles, or the number of clones analysed was too small to detect what may be a minority viral population with full-length core genes.

Despite the large number of deletion mutants described, their mechanism of generation has not been elucidated. Mapping of a 2.2kb HBV RNA isolated from HCC cases showed two deletions: an in-frame deletion in the core gene spanning nucleotides 2068-2349 inclusively (core amino acids 56-150) and a 1016 nucleotide deletion from the 3' end of the core gene (nucleotide 2454) into the 5' end of the surface gene (Chen *et al.*, 1989). Analysis of the boundaries of these deletions showed sequences which were conserved in all subtypes subsequently examined and the deleted sequences showed the consensus 5' GC/T and 3' AG boundary sequences for introns, suggesting that these deletions were the products of RNA splicing. However, as the termini of all other deletions reported cover a large region of the core gene (nucleotides 1950-2334), with differing sequences at the boundaries, it is unlikely that RNA splicing is the usual mechanism responsible for their generation.

The contribution of core gene deletion variants to the disease states observed in

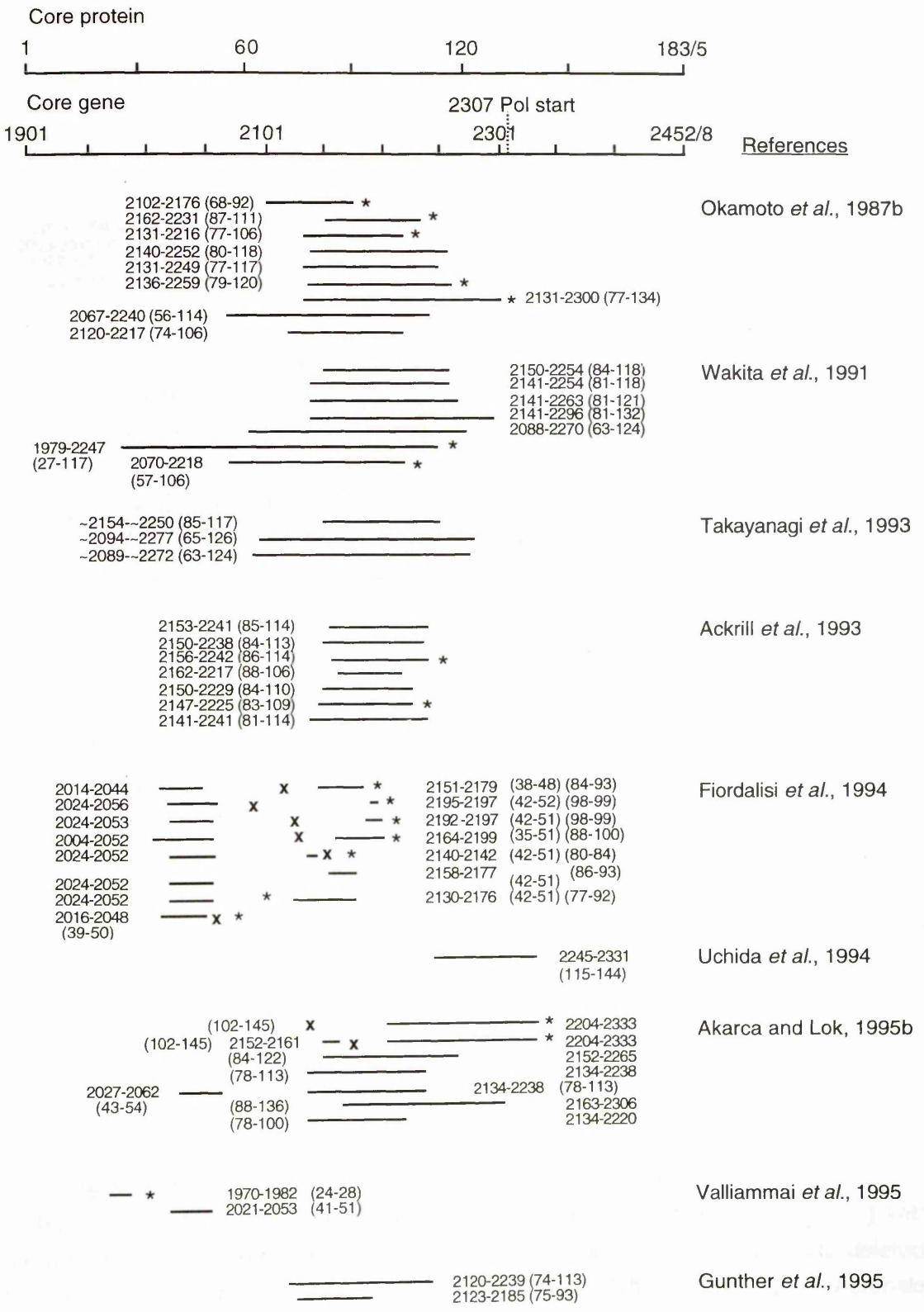


Figure 1E6: Summary diagram of core gene deletions (continued with figure legend on following page)

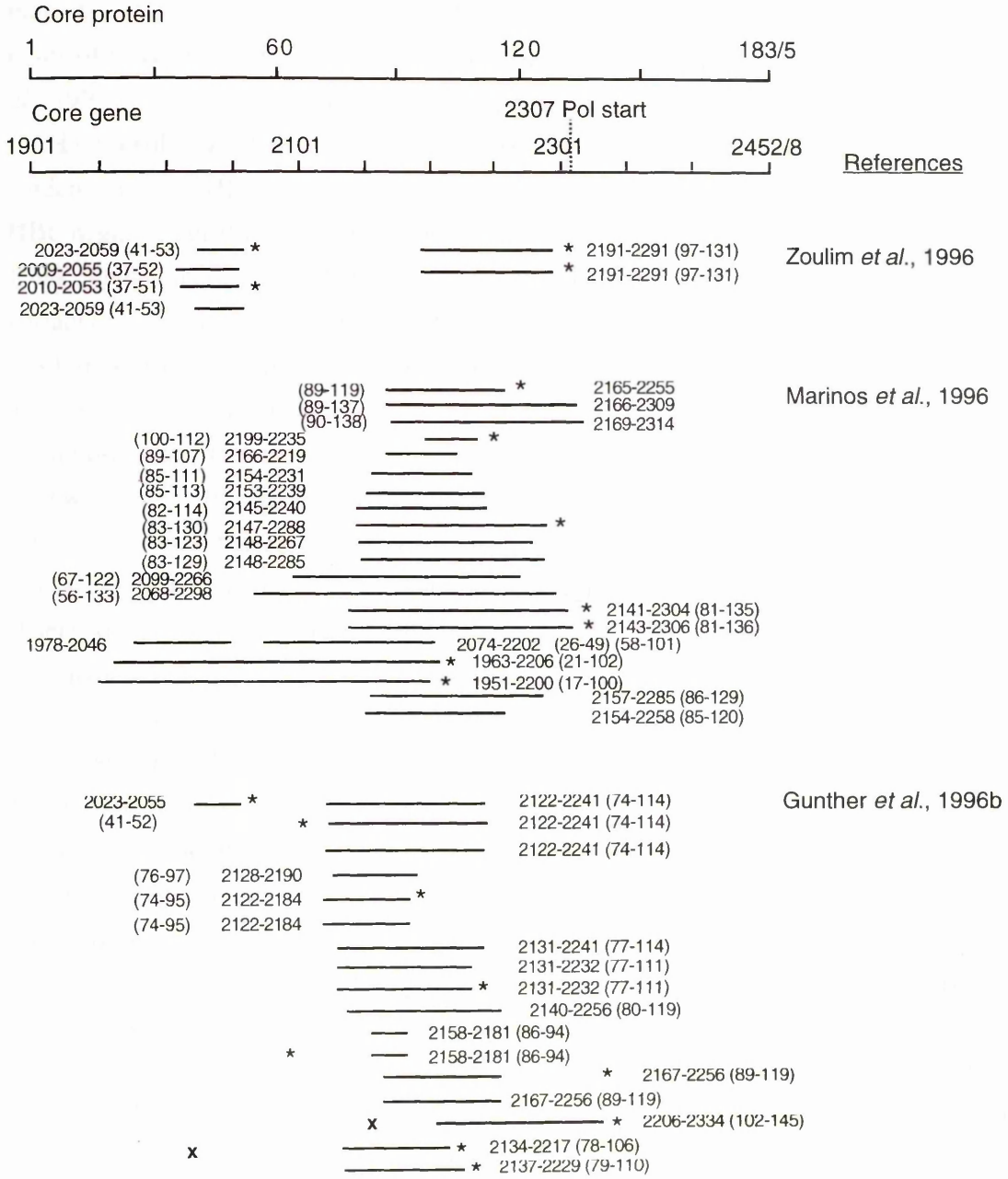


Figure 1E6: Summary diagram of reported core gene deletions (continued)

All reported cases of core gene deletions are summarised diagrammatically, along with the positions of the deletion/s. Numbers in brackets indicate the amino acids deleted. Deletions marked with an asterisk (*) resulted in a frame-shift of the core gene. Asterisks at additional positions in the gene show nucleotide deletions resulting in frame-shifts. Cases marked with a cross (**x**) show the presence of a novel stop codon either in addition to the deletion, or resulting from the frame-shift. The position of the polymerase start codon is also shown.

patients is also unknown at present. The presence of deletions has been demonstrated in cases of CAH, but not in ASCs (Wakita *et al.*, 1991; Takanayagi *et al.*, 1993; Uchida *et al.*, 1994); in e antigen positive CAH cases, but not those with CPH or anti-HBe positive CAH (Ackrill *et al.*, 1993); in e antigen positive patients with severe liver damage but no evidence of anti-HBc antibody (Zoulim *et al.*, 1996); and in patients with a similar anti-HBc negative serology, but for whom no clinical details were given (Fiordalisi *et al.*, 1994; Valliammai *et al.*, 1995). These results all point towards a role for core deletion variants in severe chronic hepatitis. However, other similar studies have not supported this hypothesis, finding deletions in e antigen positive patients with either CAH, CPH or non-specific reactive hepatitis (Akarca and Lok, 1995b) and in symptom-free carriers (Okamoto *et al.*, 1987b).

Two studies of chronic HBV-infected immunosuppressed renal transplant patients have shown that if deletions persist or increase, cirrhosis or death from end-stage liver disease occurs, whereas if the deletions disappear, no further liver deterioration is observed (Günther *et al.*, 1995, 1996b). In patients with an absence of deletions, only a few cirrhosis cases occurred and no mortalities.

As the general positions of the deleted regions correspond well to the locations of core protein epitopes, the deletion variants have been proposed to represent viral escape from the immune response. Even though the core-specific $CD4^+$ Th cell response is thought to be the most important determining factor in whether or not chronic or acute disease develops, epitopes related to all branches of the host immune response are affected by the overall range of deletions observed. In the absence of HLA-typing studies of patients with deletions, the relevance of the CTL epitopes cannot be determined. However as certain Th epitopes are HLA type-independent, disruption of these epitopes may contribute to the chronic disease. Additionally, although B-cell epitopes were also affected by the deletions, the significance of this is difficult to determine without further studies on the role of the anti-HBc or anti-HBe responses.

1E7 Aims of the work presented in this thesis

The main aim of the work presented in this thesis was to determine whether or not core genes with naturally-occurring central deletions were competent for core particle assembly. This would allow an assessment of their relevance in HBV infection. If core particle-like structures were able to form when these core genes were expressed, their ability to react with monoclonal and polyclonal antibodies would be assayed. This would determine whether or not this class of variant virus could represent escape from the host immune response.

As very little information was available on the three-dimensional structure of the HBV core particle when these studies were initiated, we also planned to provide samples of purified wild-type core particles to collaborating groups in order to assist this work. The

three-dimensional structures of any core particle-like structures formed from core deletion proteins would also be examined.

The core gene deletions examined in this thesis were all isolated by the group of Dr. N. Naoumov (Kings College, London). Expression of the core protein with amino acids 84-109 deleted had been initiated by this group using a recombinant baculovirus and both this and a similar virus expressing full-length core protein were kindly given to us in order that we could initiate our studies.

CHAPTER 2 MATERIALS AND METHODS

2A Materials

2A1 Plasmids

Plasmid pR1-11 was a kind gift from Prof. K. Murray (University of Edinburgh). This pBR322-based plasmid contained a fusion of β-galactosidase coding sequence and a fragment of the HBV genome containing the core gene (Stahl *et al.*, 1982). The fusion protein expressed was under the control of the *lacUV5* promoter and the cloning strategy resulted in the protein expressed containing 10 N-terminal amino acids of β-galactosidase with the core start codon deleted and the aspartate residue at position 2 mutated to a histidine (Figure 2A1). For the purposes of this project, the HBV fragment was excised from the plasmid with an *Eco* RI/*Bam* HI digest and replaced by PCR-amplified core gene sequence.

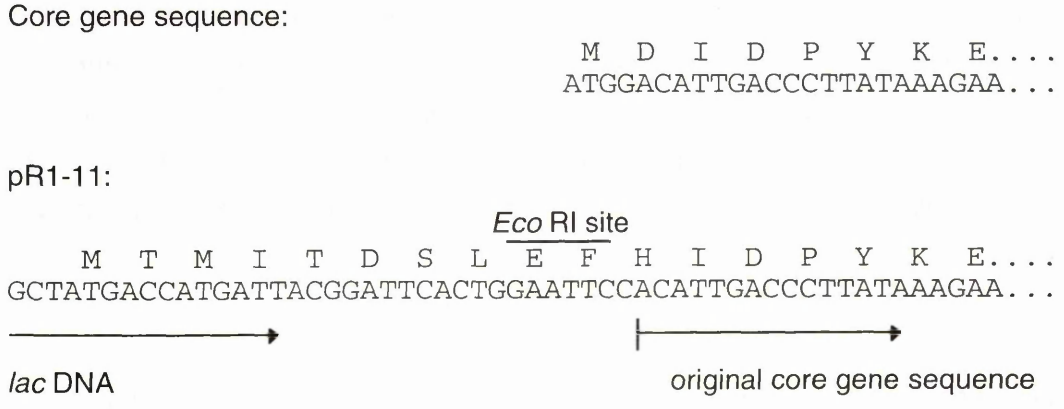


Figure 2A1: Core protein expression from plasmid pR1-11
Plasmid pR1-11 was used for expression of full-length and deletion core proteins in *E.coli* as described in the text. The core protein produced from this plasmid is a fusion protein with 10 N-terminal β-galactosidase residues, including the start codon. The original core protein methionine is missing and the aspartate residue at position 2 of core is mutated to a histidine.

Plasmids pET3a and pET9a were obtained commercially from Novagen (see Figure 2A2). Core gene sequences amplified by PCR with primer pair cp3 and Bamtag were cloned using the *Nde*I/*Bam*HI sites, placing them under the control of the

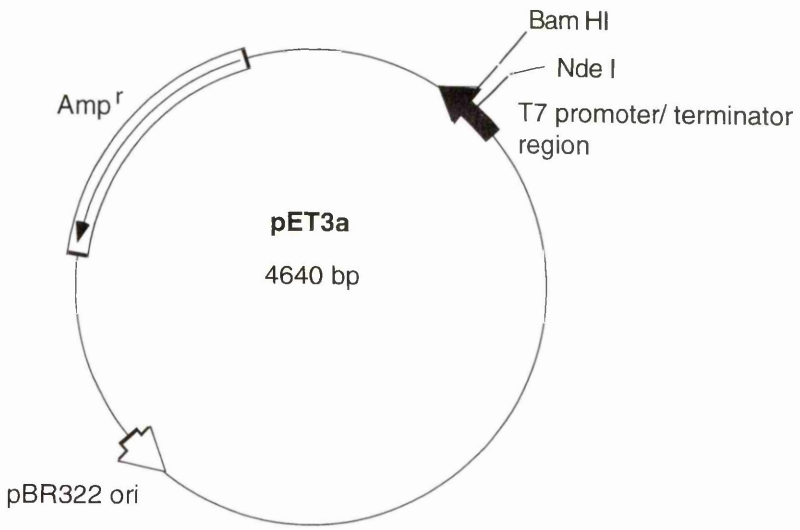


Figure 2A2: Map of plasmid pET3a

Plasmids pET3a and pET9a were used for the expression of full-length core genes and core gene deletions in *E.coli*. Core genes were amplified by PCR using primer pair cp3/Bamtag or cp3/cd2 (see Section 2A3) for cloning into pET3a or pET9a, respectively. Genes were cloned between the *Nde* I and *Bam* HI sites and were expressed under the control of the T7 promoter. Plasmid pET9a is identical to the plasmid shown, but encoded resistance to kanamycin instead of ampicillin.

bacteriophage T7 transcription and translation signals. Plasmid pET3a also encoded ampicillin resistance and pET9a encoded kanamycin resistance.

Plasmid pRK5 was obtained from Prof. H. Will (Heinrich Pette Institut, Hamburg). This pUC118-based plasmid contains a CMV IE promoter and SV40 polyadenylation site to allow expression of cloned genes in mammalian cells, as well as the gene for ampicillin resistance. This plasmid is shown in Figure 2A3.

Plasmid pTZ18R was purchased commercially from Pharmacia. This plasmid contains a gene conferring ampicillin resistance and also the pUC18 multiple cloning site, which was located in the middle of the *lac Z'* gene, allowing blue/white colour selection of recombinant plasmids on LB agar plates supplemented with IPTG and X-gal. This plasmid is shown in Figure 2A4.

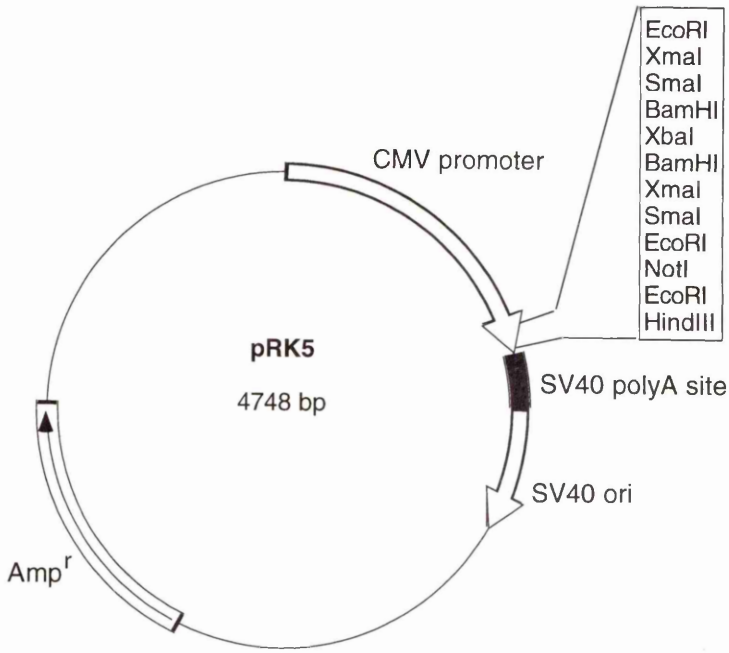


Figure 2A3: Map of plasmid pRK5

Plasmid pRK5 was used to express full-length core genes and core gene deletions in mammalian cells. Full-length core genes and core gene deletions were amplified by PCR using primer pair cpk/Hintag (Section 2A3) and cloned into *Eco* RI/*Hin* dIII-digested pRK5.

2A2 Enzymes

Restriction enzymes were obtained from Boehringer Mannheim, with the exception of *Nde*I, which was obtained from New England Biolabs. RNase and lysozyme were purchased from Sigma. DNase I, calf intestinal phosphatase, T4 DNA ligase and proteinase K were all purchased from Boehringer Mannheim. *E. coli* DNA polymerase I Klenow fragment and T4 polynucleotide kinase were purchased from New England Biolabs. PCR was carried out using *Thermus aquaticus* DNA polymerase obtained from Gibco. DNA polymerase I/DNase I was also obtained from Gibco.

2A3 Synthetic oligonucleotides

Oligonucleotides were synthesised on site using a Cruachem PS250 automated synthesiser by Dr. J. McLauchlan, Miss S. Fitzpatrick, Mr. J. McGeehan or Mr. R. Reid. Oligonucleotides used during this project are listed in Table 2A5.

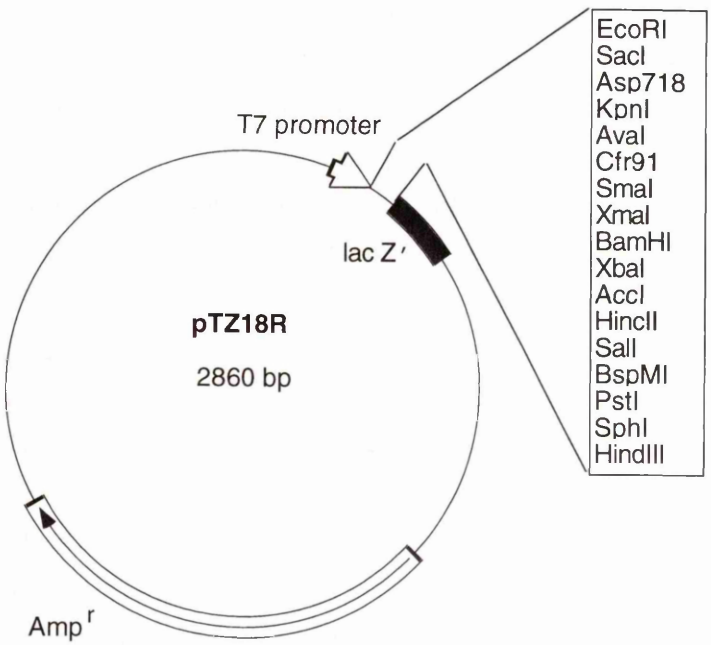


Figure 2A4: Map of plasmid pTZ18R

Core genes were amplified by PCR amplification of HBV DNA, extracted from human serum, using primer pair C1/C2 (see Section 2A3). Amplified core genes were made blunt-ended and cloned into *Sma* I-digested pTZ18R. The *lac Z'* gene allowed identification of recombinant plasmids by blue/white colour selection.

2A4 Bacteria (*E. coli* strains)

The *E.coli* strain DH5α (Φ80*dlacZΔM15 recA1 endA1 gyrA96 thi-1 hsdR17*(r_k^- m_k^-) *supE44 relA1 deoR Δ(lacZYA-argF)U169*) was used for maintenance and propagation of plasmid DNA and for expression of core protein from plasmid pR1-11.

Strain BL21(DE3)pLysS (F$^-$ *ompT hsdS$_B$* (r_B^- m_B^-) *dcm gal* λ(DE3) pLysS(Cm^r) (Studier *et al.*, 1990) was used for expression of proteins using pET vectors: the integrated lysogenic λ bacteriophage DE3 has an IPTG inducible *lacUV5* promoter from which T7 RNA polymerase is expressed and the plasmid pLysS encodes T7 lysozyme and also a gene conferring chloramphenicol resistance.

2A5 Bacterial culture media

The DH5α strain was grown in L-broth (10g NaCl, 10g BactoTryptone, 5g yeast extract in 1l water) and the BL21(DE3)pLysS strain in 2YT broth (5g NaCl, 16g BactoTryptone, 10g yeast extract in 1l water). Agar plates were made with 1% (w/v) agar in L-broth.

Where necessary, media and LBagar plates were supplemented with antibiotics: 100mg/ml ampicillin for bacteria harbouring plasmids pR1-11, pET3a and pRK5, 30mg/ml kanamycin for strains harbouring the pET9a plasmid and 25mg/ml chloramphenicol for strains harbouring the pLysS plasmid. For blue/white colour selection, bacteria transformed with plasmid pTZ18R were plated onto LB agar plates to which 10mg/ml X-gal (Section 2A9) and 50mM IPTG had been added.

2A6 Cell culture media

All cell culture media were obtained from Gibco.

*Sf*21 cells, derived from the pupal ovarian tissue of the fall army worm (*Spodoptera frugiperda*), were grown in TC100 Medium supplemented with 5% bovine calf serum (Applied Protein Products), 100 units/ml penicillin and 100µg/ml streptomycin (Gibco).

COS 7 cells (Gluzman, 1981) and HepG2 cells (Knowles *et al.*, 1980) were grown in Dulbecco's modified Eagles Medium (DMEM) supplemented with 10% bovine calf serum, 2mM glutamine (Gibco) and antibiotics as before.

2A7 Antibodies

The anti-core mouse monoclonal antibodies αHBcBF1-5 were a kind gift from Dr. B. Ferns (University College Hospital, London) and were used at a recommended dilution of 1:10.

Anti-core mouse monoclonal antibody 42B12 was a kind gift from Prof. W.H. Gerlich (Justus-Liebig Universität, Giessen). The antibody was thought to recognise an epitope between residues 115-135, and was used at a dilution of 1:750.

The anti-pp65 mouse monoclonal antibody 9220 was purchased commercially from Capricorn Products Inc. and was used at a 1:1000 dilution on Western blots and at a 1:300 dilution for immunofluorescence assays.

Anti-core rabbit polyclonal antisera were purchased commercially from Zymed Laboratories Inc., Dako Corporation and Life Sciences. These were used at dilutions of 1:2, 1:1000 or 1:200, respectively on Western blots. The Zymed polyclonal antiserum was used undiluted in immunofluorescence assays.

Goat anti-mouse IgG whole molecule peroxidase conjugate, goat anti-rabbit IgG whole molecule peroxidase conjugate, goat anti-mouse IgG whole molecule FITC conjugate and anti-rabbit IgG whole molecule FITC conjugated immunoglobulin were all supplied by Sigma.

Sense primers

Primer	Position	Sequence (5'>3')
cd1	1905-1932 (core gene)	ggaattccACATTGACCCGTATAAAGAATTTGGAGC
cp3	1901-1922	ggaattccatATGGACATTGACCCGTATAAAG
s1	2112-2140 (core)/458-479 (surface)	GGGTGGGAAGTAATTGGAAGACCCAGCA**GGTATGTTGCCCGTTTGTCCTC**
s4	458-479 (surface)	**GGTATGTTGCCCGTTTGTCCTC**
cd4	2264-2289	TTT/CGGAGTGTGGATTCGCACTCCTCC
5f10A	2145-2170/2264-2286	CCCTTGATCATCAGTCGATACAGTTA/TTTGGAGTGTGGATTCGCACTCC
5f20A	2179-2200/2264-2286	CCCGGATTTTTAGTCCGTTGAT/TTTGGAGTGTGGATTCGCACTCC
5f30A	2205-2230/2264-2286	CCAAAGTGTAAAGGACGGAATGAAAA/TTTGGAGTGTGGATTCGCACTCC
3f10	2234-2258	AGAGAGACTGTACTTGAATATTTGGTCTCTTTCGG
3f20	2204-2233	TGGTTTCATATATCTTGCCTTACTTTTGGAAGAGAGACTG
3f30	2174-2206	AACATGGGTTTAAAGATCAGGCAACTATTGTGG
cpk	1901-1914	ggaattcATGGACATTGACCC
C1	1737-1754	GAGTTGGGGGAGGAGATT

Table 2A5: Oligonucleotide primer positions and sequences (continued on following page). Numbering is according to Ono *et al.* (1983).

Antisense primers

Primer	Position (core gene)	Sequence (3'>5')
cd2	2441-2458	GCCCTTAGAGTTACAATCgccctagggc
Bamtag	2439-2455	GAGCCCTTAGAGTTACA**ctcgcgttctgcggggcgc agtggccgccgatc**cctagggc
Thrtag	2267-2284	GGAGTGTGGATTCGCACT**ctcgcgttctgcgggg**c **gcagtggccgccgatc**cctagggc
s2	557-580 (surface)/2264-2291 (core)	**GGGAGAACAACGACATGTTTTGGA-**AAACCTCACACCTAAGCGTGAGGAGGGC
s3	557-580 (surface)	**GGGAGAACAACGACATGTTTTGGA**
cd3	2112-2140	CCCACCCACCATTAAACCTTCTGGGTCGT
5f10	2145-2170	CCCTTGATCATCAGTCGATACAGTTA
5f20	2179-2200	CCCGGATTTTTAGTCCGTTGAT
5f30	2205-2230	CCAAAGTGTAAAGGACGGAATGAAAA
3f10A	210-238/332-356	CCCACCCATTATTAAACGTTCTAGGTCGT-TCTCTCTGACATGAACTTATAAACC
3f20A	210-238/302-331	CCCACCCATTATTAAACGTTCTAGGTCGT-ACCAAAGTATATAGAACGGAATGAAAACC
3f30A	210-238/272-293	CCCACCCATTATTAAACGTTCTAGGTCGT-TTTGTACCCAAATTTCTAGTCCG
Hintag	2439-2455	GAGCCCTTAGAGTTACA**ctcgcgttctgcggggcgc agtggccgccgatc**ttcgaagggg
C2	2475-2495	TTCCACCCTTTGAAATGACCC

Primer sequences for amplification of surface gene regions are shown in bold
Restriction enzyme sites and non-coding sequences are shown in lower case
pp65 epitope tag sequences are shown in bold lower case

Table 2A5: Oligonucleotide primer positions and sequences (continued). Numbering is according to Ono *et al.* (1983).

2A8 Radiochemicals

All radiochemicals were purchased from Amersham at the following specific activities:

α^{32}P-dATP	3000Ci/mmol (10μCi/μl)
α^{32}P-dTTP	3000Ci/mmol (10μCi/μl)
^{35}S-L-Methionine	~800Ci/mmol (15μCi/μl)
α^{35}S-dATP	1000Ci/mmol (10μCi/μl)

2A9 Solutions

6xagarose gel loading buffer:	1xTE, 40% sucrose, 1mg/ml bromophenol blue
Blocking buffer (Western blots):	3% gelatin in TBS
Coomassie stain:	0.2% Coomassie brilliant blue, 50% methanol, 5% acetic acid in water
40xdNTPs	10mM of each of dATP, dCTP, dGTP, dTTP
Denaturation buffer:	0.6M NaCl, 0.2M NaOH
50xDenhardt's solution:	1% polyvinylpyrrolidine, 1% ficoll, 1% BSA
2xHBS:	280mM NaCl, 50mM HEPES, 1.5mM Na_2HPO_4, pH to 7.05 with NaOH
Hybridisation buffer:	5xSSC, 10xDenhardt's solution, 0.5% SDS, 20μg/μl denatured salmon sperm DNA
5xkinase buffer:	350mM Tris.HCl, 50 mM $MgCl_2$, 25mM DTT, 5mM ATP, pH7.6
5xligase buffer:	250mM Tris.HCl, 50mM $MgCl_2$, 5mM ATP, 5mM DTT, 25% PEG 8000, pH7.6
Lysozyme extraction buffer:	50mM Tris.HCl, 100mM NaCl, 0.2mM EDTA, 1mM lysozyme, 1.5mM $MgCl_2$, O.5mM PMSF, pH8.0
Miniprep solution I:	25mM Tris.HCl, 50mM glucose, 10mM EDTA,
Miniprep solution II:	0.2N NaOH, 1% SDS
Miniprep solution III:	3M potassium acetate, 11.5% acetic acid
Neutralisation buffer:	1M Tris.HCl, 0.6M NaCl, pH8.0
Oligonucleotide elution buffer:	0.5M ammonium acetate, 10mM $MgCl_2$, 0.1% SDS, 1mM EDTA
Oligonucleotide sample buffer:	90% deionised formamide in 1xTBE
PBS(A):	170mM NaCl, 3.4mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.2
PBS:	PBS(A) plus $CaCl_2.2H_2O$ and $MgCl_2.6H_2O$ at 1g/l

Pre-hybridisation buffer:	5xSSC, 5xDenhardt's solution, 0.1% SDS, 20µg/µl denatured salmon sperm DNA
Protein gel destain:	50% methanol, 5% acetic acid in water
10xSDE:	250mM sodium acetate, 25mM EDTA
SDS PAGE buffers:	
(i) 3xSDS PAGE sample buffer:	10% SGB, 2% SDS, 5% β-ME, 10% glycerol, 1mg/ml bromophenol blue
(ii) 3xSDS PAGE sample buffer: (non-reducing)	10% SGB, 2% SDS, 10% glycerol, 1mg/ml bromophenol blue
(iii) RGB (resolving gel buffer):	1.5M Tris.HCl, pH8.8
(iv) SGB (stacking gel buffer):	0.5M Tris.HCl, 0.4% SDS, pH6.9
(v) Tank buffer:	40mM Tris, 185mM glycine, 0.1% SDS
Serum extraction buffer:	1xSDE, 1mg/ml proteinase K, 0.45% SDS, 10µg/ml tRNA
Sonication buffer:	50mM Tris.HCl, 1% Triton X-100, pH8.0
20xSSC:	3M NaCl, 0.3M trisodium citrate
50xTAE	2M Tris, 50mM EDTA, pH to 8.0 with acetic acid
10xTBE:	1M Tris, 1M boric acid, 20mM EDTA
TBS:	20mM Tris.HCl, 500mM NaCl, pH7.5
TE:	10mM Tris.HCl, 1mM EDTA, pH8.0
TE/RNase:	1xTE, 20µg/ml RNase
TN buffer:	50mM Tris.HCl, 100mM NaCl, pH8.0
TNE I:	10mM Tris.HCl, 50mM NaCl, 0.1mM EDTA, pH7.4
TNE II:	50mM Tris.HCl, 100mM NaCl, 0.2mM EDTA, 0.1% Triton X-100, 1.5mM $MgCl_2$, 0.5mM PMSF, pH8.0
Towbin blotting buffer:	25mM Tris.HCl, 192mM glycine, 20% methanol, pH8.3
TTBS:	TBS plus 0.05% Tween-20
Versene:	0.6mM EDTA in PBS(A), 0.002% phenol red
X-gal	20mg/ml X-gal in dimethylformamide

2A10 Chemicals and reagents

All chemicals and reagents were purchased from BDH Chemicals UK or from Sigma Chemical Co. unless otherwise stated in this section or in the Methods section:

Aldrich:	4-Chloro-1-naphthol
Beecham Research:	Ampicillin

Bio-Rad:	TEMED, ammonium persulphate, Coomassie brilliant blue, gelatin
Boehringer Mannheim:	Tris, agarose
Camlab:	Phenol:chloroform (25:24:1 with isoamyl alcohol)
DuPont:	En^{3}Hance
Gibco:	IPTG, tRNA
Melford Laboratories:	CsCl
Pharmacia:	ATP
Prolabo:	Boric acid, butanol, chloroform, ethanol, glycerol, hydrochloric acid, isopropanol, methanol

2B Methods

2B1 Nucleic acid manipulation and cloning procedures

2B1.1 Large scale plasmid DNA preparations

(i) Continuous CsCl:EtBr gradients

An aliquot of *E.coli* harbouring the plasmid was spread on an agar plate containing the appropriate antibiotics (see Section 2A5) and incubated overnight at 37°C. Colonies were inoculated into 300ml L-broth containing the appropriate antibiotics and incubated overnight at 37°C with shaking.

Plasmid DNA was prepared by a modification of the alkaline lysis method of Birnboim and Doly (1979) and supercoiled plasmid DNA was purified on CsCl gradients. Bacterial cultures were centrifuged at 6k rpm for 15min at 4°C (Sorvall GS3 rotor) and the pellets resuspended thoroughly in 5ml miniprep solution I. Fresh miniprep solution II was prepared and 10ml was added to the resuspended cell pellet, which was then mixed by several inversions and incubated on ice for 5min. Ice-cold miniprep solution III (7.5ml) was added and the mixture shaken vigorously to mix, followed by a 10min incubation on ice. Cell debris was then pelleted by centrifugation at 3k rpm for 15min at 4°C (Sorvall SS34 rotor). The supernatant was carefully removed and mixed with an equal volume of phenol:chloroform solution before being centrifuged again at 3k rpm (Sorvall RT6000B benchtop centrifuge) for 10min at 4°C and the aqueous phase retained. DNA was precipitated by the addition of 0.6vol isopropanol and collected by centrifugation at 3k rpm (Sorvall RT6000B) for 20min at 4°C. The pellet was air dried, resuspended in CsCl solution (10ml water plus 10g CsCl) and 0.4ml 10mg/ml EtBr solution added. Protein:EtBr complexes were removed by centrifugation at 3k rpm (Sorvall RT6000B) for 15min at 18°C and the supernatant transferred to a Beckman Quick-Seal centrifuge tube. Tubes were balanced with CsCl solution and heat sealed using metal former caps. Supercoiled plasmid DNA was purified by centrifugation at 45k rpm (Sorvall T1270 angled rotor) for 64hr at 18°C. The plasmid DNA band was removed using a syringe, taking care to avoid contamination from the upper chromosomal DNA band. The EtBr was removed from the plasmid solution by 2-3 extractions with TE-saturated butanol and diluted by the addition of 3vol of water. DNA was precipitated by the addition of 3.5vol of ethanol and incubation at -20°C. The DNA was pelleted by centrifugation at 3k rpm (Sorvall RT6000B) for 10min at 4°C, resuspended in 2ml water and reprecipitated by the addition of 0.5vol 7.5M ammonium acetate and 2vol ethanol. After centrifugation at 3k rpm (Sorvall RT6000B) for 30min at 4°C the DNA pellet was air dried and resuspended on 200µl TE.

(ii) Plasmid purification using Qiagen columns

Qiagen columns were occasionally used to purify smaller quantities of plasmid DNA for use in sequencing reactions. Bacterial lysates are applied, under defined salt conditions, to the column resin and the DNA is selectively bound and purified from RNA, proteins and other cellular contaminants.

Smaller 50ml cultures were grown as above and the cells pelleted by centrifugation at 6k rpm for 15min at 4°C (Sorvall GSA rotor). The cell pellet was resuspended thoroughly in 4ml of buffer P1, 4ml of buffer P2 was added and mixed by inversion 5-6 times. After a 3min incubation at room temperature 4ml of buffer P3 was added, the solution mixed immediately by several inversions and incubated on ice for 15min. After centrifugation at 16k rpm (Sorvall SS34 rotor) for 30min at 4°C, the supernatant was removed and added to a Qiagen 100 column, equilibrated with 4ml equilibration buffer QBT. After the supernatant had dripped through, the resin was washed twice with 10ml wash buffer QC before elution of the DNA by the addition of 6ml elution buffer QF. Plasmid DNA was precipitated by the addition of 0.7vol isopropanol and centrifugation at 10k rpm (Sorvall SM24 rotor) for 30min at 4°C. The pellet was washed with 70% ethanol, air dried and resuspended in 100µl TE.

Plasmid concentrations were determined by measuring the absorbance at 260nm, assuming $1A_{260}$=50µg/ml.

2B1.2 Miniprep plasmid DNA preparation

Single colonies of transformed bacteria were inoculated into 3ml of L-broth containing the appropriate antibiotics and incubated overnight at 37°C with shaking. Aliquots (1.5ml) of the cultures were centrifuged at 13k rpm for 30sec in a microfuge and the pellet resuspended in 100µl of miniprep solution I. A 250µl aliquot of fresh miniprep solution II was added and mixed by several inversions before addition of 150µl of ice-cold miniprep solution III, which was mixed by vortexing. A 100µl aliquot of phenol:chloroform was then added to the mixture, which was vortexed once more before centrifugation for 10min in a microfuge. Plasmid DNA was precipitated from the aqueous phase by the addition of 400µl ethanol and pelleted by centrifugation for 20min in a microfuge. The pellet was washed with 70% ethanol and air-dried before resuspension in 30µl TE/RNase and incubation at 37°C for 30min.

2B1.3 Restriction enzyme digestion of DNA

Restriction enzyme digestion of DNA was carried out at 37°C (or the temperature specified by the supplier) in 20µl volumes using 1 unit of enzyme per 0.5µg DNA per hour in the buffer supplied with the enzyme. The buffer system used most often was the Boehringer Mannheim A, B, M, L, H system and less frequently the New England

Biolabs NEBuffer 1, 2, 3, 4 system. Typically 0.5µg DNA was digested to identify recombinant plasmids and 1-5µg for isolation of specific restriction fragments.

2B1.4 Electrophoretic separation and purification of DNA fragments

(i) Non-denaturing agarose gels

DNA fragments produced by restriction enzyme digestion or PCR were resolved by non-denaturing agarose gel electrophoresis. Horizontal slab gels approximately 0.5cm thick were used and samples loaded following the addition of agarose gel loading buffer. For separation of fragments over 200bp, 1% agarose gels in 1xTBE were made up to a final concentration of 1µg/ml EtBr before casting and run at 7V/cm in 1xTBE buffer, containing EtBr. For separation of small PCR products, 2% agarose gels were used. Gibco 100bp ladder or 1kb ladder standard size markers were used where appropriate. Following electrophoresis, DNA was visualised under short-wave UV light (long-wave for preparative gels). Photography was carried out using The Imager (Appligene).

(ii) Purification of DNA from agarose gels

Agarose blocks containing appropriate DNA fragments were excised from gels under long-wave UV transillumination and the DNA recovered using a commercial kit, Geneclean II (Bio 101 Inc., La Jolla, CA). The kit contains a silica matrix which binds DNA in the presence of high concentrations of sodium iodide. Agarose blocks were mixed with 4.5 volumes of NaI solution and 0.5 volumes TBE modifier and incubated at 55°C until the gel slice had melted. The silica matrix was then added (5µl for up to 5µg DNA) and the mixture vortexed and incubated on ice for 5min, with further occasional mixing. Following a 30sec centrifugation, the silica matrix was washed 3 times in 500ml 'NEW' wash and the DNA eluted into an appropriate volume of water by mixing the washed pellet with the water and incubating at 55°C for 5min. The matrix was again pelleted and the supernatant containing the DNA removed for subsequent use. The Geneclean II kit was also used in situations where electrophoresis was not required, particularly for separating DNA from nucleotides and PCR primers. In these cases, 3 volumes NaI were used and the TBE modifier omitted.

(iii) Non-denaturing polyacrylamide gels

Vertical non-denaturing polyacrylamide (Protogel (acrylamide:bisacrylamide 37.5:1), National Diagnostics) gels, containing 1xTBE, were used to purify synthetic oligonucleotides. Polymerisation was instigated by adding 100µl TEMED and 100 µl 25% APS. Prior to loading, oligonucleotides were boiled with sample buffer for 2min. Electrophoresis was carried out using 15% polyacrylamide gels (0.3cm thick x 25cm long) in 1xTBE running buffer at 200V for 3-5hr.

(iv) Denaturing polyacrylamide gels

Vertical denaturing polyacrylamide gels, containing 8M urea and 1xTBE (Sequagel 6 (5.7% acrylamide:0.3% bisacrylamide) National Diagnostics) were used to resolve the products of DNA sequencing reactions. Polymerisation was instigated by adding 600µl 10% APS. Sequencing gels were 0.4mm thick x 39cm long. The samples were heated to 72°C for 2min prior to loading. Electrophoresis was in 1xTBE running buffer at 70W for 1.5-5hr.

2B1.5 End repair and DNA ligation

Modifications to the ends of purified DNA fragments were carried out prior to ligation, as detailed below, and the modified DNA purified by extraction with phenol:chloroform, followed by extraction by chloroform and an ethanol precipitation. The DNA pellets were washed with 70% ethanol and resuspended in water for use in ligation reactions.

(i) Removal of 3' overhangs

It was necessary to remove adenosine nucleotides added to the 3' termini of DNA fragments amplified by PCR using *Taq* DNA polymerase (Clark, 1988), to enable efficient ligation. DNA fragments were made blunt-ended in Klenow buffer by incubation with 3 units of Klenow fragment. The reaction was carried out for 40min at 37°C.

(ii) Phosphate removal from 5' ends

Removal of the 5' phosphates from vector fragments with blunt ends was done by incubation with 1 unit of calf intestinal phosphatase in the appropriate buffer for 30min at 37°C. A further 1 unit of enzyme was added and the incubation repeated before purification of the vector fragment.

(iii) Phosphate addition to 5' ends

It was necessary to add phosphate to the 5' ends of blunt-ended PCR products to facilitate blunt-end ligation reactions with a plasmid vector. The reactions were carried out in 30µl of 1xkinase buffer with 5 units T4 polynucleotide kinase for 45mins at 37°C.

(iv) Ligation reactions

Purified vector and insert fragments prepared by restriction enzyme digestion and with ends suitably repaired, were ligated in a 1:3 or 2:1 ratio for sticky- or blunt-ended ligations, respectively. Reactions were carried out in a 10µl volume of 1xligase buffer with 1 or 3 units of T4 DNA ligase for sticky- or blunt-ended ligations respectively.

Reactions were carried out at 16°C for 4-15hr. Ligated DNA was stored at -20°C prior to transformation of competent *E.coli.*

2B1.6 Preparation and transformation of competent *E.coli* cells for plasmid growth and maintenance

Plasmids were grown and maintained in *E.coli* strain DH5α. A 0.5ml aliquot of overnight culture grown in L-broth from a single colony was inoculated into 50ml L-broth and grown for 2.5-3hr at 37°C in a shaking incubator, until the OD_{630} was approximately 0.3. The culture was cooled on ice and the cells pelleted by centrifugation at 3k rpm for 15min at 4°C. The cell pellet was resuspended in 15ml cold 50mM $CaCl_2$. After a 1hr incubation on ice, the cells were pelleted as before and resuspended in 2.5ml cold 50mM $CaCl_2$. For storage, 500µl aliquots of resuspended cells were added to an equal volume of 30% glycerol, flash frozen and stored at -70°C. For more immediate use resuspended cells were stored at 4°C overnight to enhance their competency.

For transformation of *E. coli*, 100ng plasmid DNA (or half of a ligation reaction) was added to a 100µl aliquot of competent DH5α bacteria, which had been thawed on ice if frozen. The mixture was incubated on ice for 20-45min prior to heat shock in a 42°C water bath for 1min, followed by a 2min incubation on ice. A pre-warmed 1ml aliquot of 2YT broth was added to the cells which were then shaken for 1hr at 37°C before plating onto LB agar plates containing the appropriate antibiotics. The plates were incubated overnight at 37°C.

2B1.7 Preparation and transformation of competent *E.coli* cells for protein expression

E.coli strain BL21(DE3)pLysS was transformed with plasmids from the pET expression system immediately before use. Bacteria were streaked onto LBagar plates containing chloramphenicol and fresh colonies inoculated into 10ml 2YT containing chloramphenicol and grown in a shaking incubator for 2-3hr at 37°C, until the OD_{630} reached approximately 0.3. The cells were pelleted by centrifugation at 3k rpm for 5min at 4°C, resuspended in 0.5ml cold 0.1mM $MgCl_2$ and repelleted by centrifugation under the same conditions for 2min. The pellet was resuspended in 0.5ml cold 0.1M $CaCl_2$ and incubated on ice for 2hr prior to use. Approximately 0.5µg of the relevant plasmid was used to transform a 100µl aliquot of the competent cells as described in Section 2B1.6.

2B1.8 Purification of synthetic oligonucleotides

Synthetic oligonucleotides were produced by the phosphoramidite method using 200µl synthesis columns. To remove the oligonucleotide from the column, 5ml

disposable syringes were attached to both ends of the column with one containing 1.5ml ammonium hydroxide (0.88 specific gravity (Fisher Scientific)). The solution was pushed through the column in 200µl aliquots with a 20min incubation between each fresh addition. After incubation with the final aliquot, the 1.5ml solution was pushed backwards and forwards through the column 4-5 times to mix the aliquot thoroughly. The oligonucleotide solution was incubated in a tightly sealed tube for 5hr at 55°C to remove protecting groups. Following lyophilisation, the oligonucleotide was dissolved in 200µl water and boiled in an equal amount of oligonucleotide sample buffer prior to electrophoresis on a 15% non-denaturing polyacrylamide gel (Section 2B1.4(iii)). Oligonucleotides were purified by passive elution from gel slices excised from the gel following visualisation of the DNA by long wave UV transillumination: bands were visible as dark shadows against a fluorescing TLC plate. Gel slices were diced, added to 1ml oligonucleotide elution buffer and the DNA eluted overnight at 37°C in a shaking incubator. The liquid phase, containing the DNA, was separated from the gel fragments and any residual acrylamide removed by centrifugation for 5min (microfuge). The DNA solution was extracted with phenol:chloroform, followed by chloroform extraction and DNA purified by ethanol precipitation followed by centrifugation for 10min (microfuge). The pellet was lyophilised and resuspended in TE. The DNA concentration was determined by measuring the absorbance at 260nm, assuming that for short oligonucleotides $1A_{260}$=20µg/ml.

2B1.9 Extraction of HBV DNA from serum

All manipulations involving serum samples were carried out in a Microflow class 2 biological safety cabinet.

Viral DNA was extracted from a 25µl aliquot of serum by overnight incubation at 37°C in serum extraction buffer. The mixture was then phenol:chloroform extracted, followed by an extraction with chloroform, and the DNA precipitated by the addition of 0.1vol 3M sodium acetate and 2.5vol ethanol. Following incubation at -70°C for at least 30min, the DNA was pelleted by centrifugation for 15min (microfuge) and washed twice with 70% ethanol before lyophilisation and resuspension in 15µl water. Extracted viral DNA was stored at -20°C and 5µl used in subsequent PCR reactions.

2B1.10 Preparation of baculovirus DNA

Baculovirus DNA was extracted from infected *Sf*21 cells (see Section 2B2.2 for infection of *Sf*21 cells) for subsequent analysis by Southern blotting.

Infected *Sf*21 cells in 35mm plates were scraped into 0.5ml PBS and pelleted by centrifugation at 6k rpm (microfuge) for 2min. The pellet was resuspended in 250µl TE and 250µl buffer containing 50mM Tris.HCl, 5% β-ME, 0.4% SDS, 10mM EDTA, 10mg/ml proteinase K and 10mg/ml RNase and incubated at 37°C overnight. The mix

was extracted twice with phenol:chloroform and DNA precipitated from the aqueous phase by the addition of 0.1vol 3M sodium acetate, 2vol ethanol and incubation at -70°C. The DNA was pelleted by centrifugation for 10min at room temperature (microfuge), washed with 70% ethanol and air dried before the addition of 100µl TE. DNA was allowed to resuspend slowly by incubation for 1hr at 37°C and overnight incubation at 4°C and was stored at 4°C.

2B1.11 Polymerase chain reaction (PCR) amplification of DNA

PCR amplification was carried out to amplify core gene sequences from extracted DNA for sequence analysis or, using primers including restriction enzyme sites, to amplify core gene sequences for subsequent cloning steps. PCR reactions were set up in a Microflow class 2 biological safety cabinet, to avoid contamination, and carried out using a Biometra TRIO-Thermoblock machine.

Reactions were carried out in 50µl volumes containing 1x reaction buffer (supplied with the enzyme), 250µM dNTPs, 25pmol of each primer, 1.5mM $MgCl_2$, 1.2 units *Taq* DNA polymerase and template DNA (purified DNA fragments or extracted viral DNA). The mixtures were overlaid with mineral oil to prevent evaporation during cycling. Cycling conditions were typically: 5 cycles at 95°C for 4min (denaturing), 55°C for 1min (annealing), 72°C for 1min (elongation), followed by 35 cycles at 90°C for 1min, 55°C for 1min 15sec, 72°C for 1min 30sec. A final elongation step was then performed at 72°C for 5min. Cycling conditions were varied in attempts to allow efficient priming to occur or to minimise non-specific priming events, by lowering the annealing temperature to 42°C or raising it to 60°C, respectively. PCR products were purified (Section 2B1.4(i)) before restriction enzyme digestion, end repair or use in further PCR reactions.

Core deletion protein Δ81-121 was modified by exact replacement of deleted core gene sequence with surface gene sequence, to create the Δcore-surface hybrid gene (Section 3C4) and also by gradually replacing the deleted core gene sequence from either the 5' or 3' end of the deletion, creating a series of 'fill-in' core genes (Section 3E3). Details of the PCR-based protocols used are shown in Figure 2B1.11a and Figure 2B1.11b, respectively and Table 2B1.11c shows all proteins produced using this method.

2B1.12 Dideoxy sequencing of DNA

Plasmid DNA or DNA amplified by PCR was sequenced by the dideoxynucleotide chain termination method based on that of Sanger *et al.* (1977). Sequencing was carried out using the Amersham Sequenase version 2.0 kit, which utilises a modified bacteriophage T7 DNA polymerase.

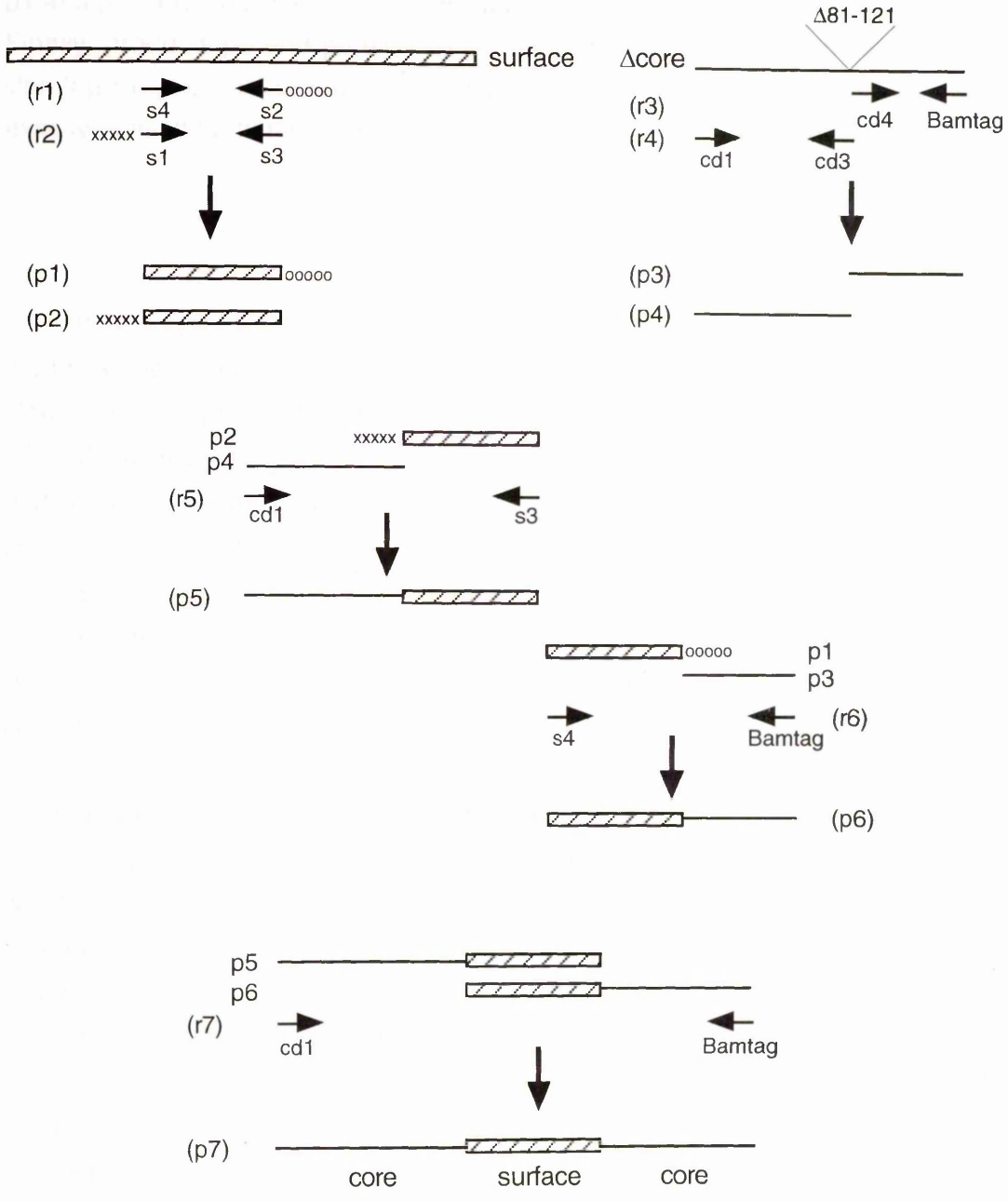


Figure 2B1.11a: Construction of the Δcore-surface hybrid gene by PCR
Primers used are detailed in Table 2A5 and are indicated by arrows (←). A 123bp fragment of the HBV surface gene (nucleotides 458-580) was amplified twice by PCR (r1 and r2) using two separate pairs of primers (s4/s2 and s1/s3) to give products p1 and p2. From the core gene with nucleotides 2141-2263 deleted (Δcore) two fragments adjacent to the deleted region were amplified separately (r3 and r4) using primer pairs cd4/Bamtag and cd1/cd3, resulting in products p3 and p4, respectively. Primers s2 and s1 include sequences complementary to core gene sequence adjacent to the 3'('oooo') and 5'('xxxx') limits, respectively, of the deleted region. Products p2 and p4 were annealed and amplified (r5) with primers cd1 and s3 to give product p5 and products

p1 and p3 were amplified similarly (r6) with primers s4 and Bamtag to give product p6. Finally, products p5 and p6 were annealed and amplified (r7) with primers cd1 and Bamtag to give final product p7, which has the deleted core sequence 2141-2263 exactly replaced with foreign sequence.

Purified plasmid DNA (2-5µg) was denatured at room temperature for 10min in a 20µl reaction volume containing NaOH at a final concentration of 0.4M. The denatured DNA was precipitated by addition of 6µl 3M sodium acetate, 14µl water and 120µl ethanol and incubated in a dry-ice:ethanol bath for 10min. Denatured DNA was then pelleted by centrifugation for 10min (microfuge), washed with 70% ethanol, air-dried and dissolved in 7µl water.

The appropriate primer (2-5pmol in 1µl volume) was added to the DNA with 2µl 10xreaction buffer (supplied with the kit) and annealing carried out for 20min at 37°C. Annealed DNA was added to 3µl labelling mix (7.5µM each of dGTP, dCTP and dTTP) containing 5mCi α^{35}S-dATP. The four chain termination mixes were aliquotted in 2.5µl volumes into four wells of a round bottomed 96-well microtitre plate (Costar). The chain termination mix for each nucleotide contained the appropriate dideoxynucleotide at a concentration of 8µM, 80µM of the three remaining deoxynucleotides and 50mM NaCl. Sequenase T7 DNA polymerase was diluted 1:8 in enzyme dilution buffer (supplied with the kit) and 2µl (approximately 3units) added to the labelling mix, followed by incubation for 2min at 20°C. For the chain termination reactions, 3.5µl aliquots of the labelling reaction were transferred to each termination mix and the reactions incubated for 15min before being stopped by the addition of 4µl of stop solution, supplied with the kit.

For sequencing PCR products, the initial alkaline denaturation step was omitted and the primer annealed to the template using the following protocol: 95°C for 2min (followed by a brief centrifugation (microfuge)), 70°C for 3min, 95°C for 45sec and 37°C for 10min. Labelling and termination reactions were carried out as above.

Prior to electrophoresis, the reaction mixes were denatured at 72°C for 2min using a Biometra UNO-Thermoblock. Electrophoresis was carried out on a denaturing polyacrylamide gel (Section 2B1.4(iv)) and gels vacuum dried and exposed to Kodak X-OMAT film. Sequences were analysed using the Wisconsin Genetics Computer Group Package, Version 8.

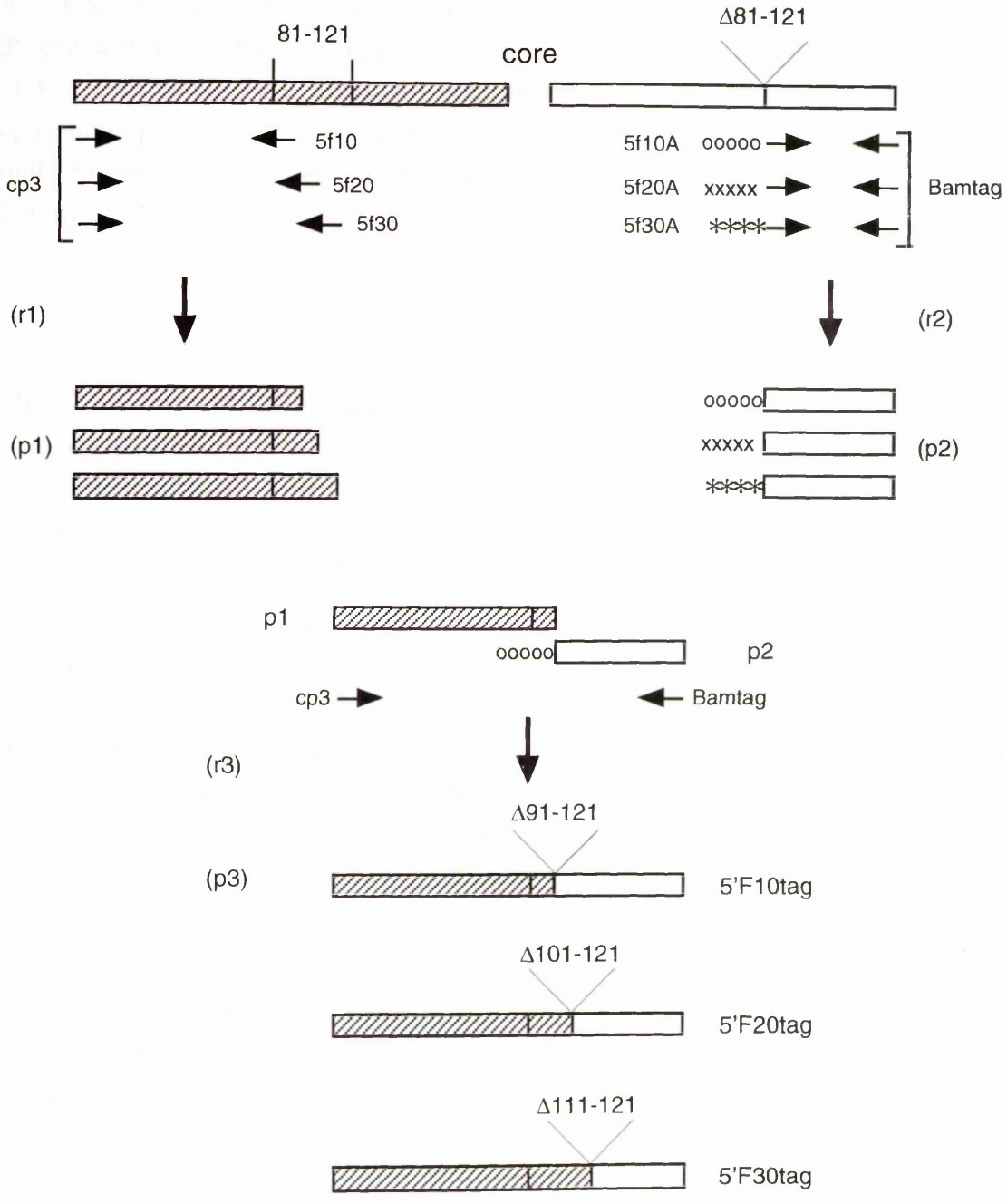


Figure 2B1.11b: Construction of the Δ81-121 5'fill-in genes

Primers used are detailed in Table 2A5 and are indicated by arrows (←—). Fragments of the 5' region of a full-length core gene were amplified by PCR (r1) using primer pair cp3 and 5f10, 5f20 or 5f30. The products of these reactions (p1) consisted of the original core gene sequence 5' to the region deleted in Δ81-121 with increasing sizes of additional core gene sequence, deleted in the Δ81-121 gene. Sequences 3' to the deleted region were amplified by PCR (r2) using primer pair Bamtag and 5f10A, 5f20A

or 5f30A. Primers 5f10A, 5f20A and 5f30A included sequences complementary to the 3' terminii of the corresponding product from r1 (oooo, xxxx and ****). Corresponding 5' and 3' fragments were annealed and amplified by PCR (r3) using primer pair cp3/Bamtag to give the final products (p3). These products consist of the Δ81-121 deletion filled in, from the 5' terminus of the deleted region, with increasing sized fragments of original core gene sequence.

2B1.13 Southern blotting (transfer of DNA to nylon) and DNA-DNA hybridisation

Restriction enzyme digestions of miniprep DNA and baculovirus DNA were probed with core gene fragments to determine the presence and size of the cloned core genes.

(i) Transfer of DNA to nylon

Restriction enzyme digestion reactions were electrophoresed on an agarose gel in 1xTBE. The gel was incubated in denaturation buffer for 45min, followed by neutralisation buffer for 45min, both with shaking at room temperature. The DNA was transferred overnight onto Hybond-N membranes (pre-soaked in water before use) by placing the gel on a wick of Whatman 3MM paper (soaked in 20xSSC and descending into a 20xSSC reservoir), then adding sequentially to the top of the gel, Hybond-N membrane, three pre-soaked sheets of Whatman 3MM paper and a stack of absorbent paper towels. Care was taken to eliminate air bubbles between layers and a light weight was added to the top of the stack to aid buffer transfer. After transfer, the filter was air dried and DNA cross-linked to the filter by exposing the filter, DNA side down, to short-wave UV light for 4min.

(ii) DNA-DNA hybridisation

Hybond-N membranes were pre-hybridised for 4hr at 65°C in pre-hybridisation buffer, followed by overnight incubation at 65°C in hybridisation buffer containing probe DNA (see below). The filter was washed once in 2xSSC, 0.1% SDS for 15min and twice in 0.1xSSC, 0.1% SDS for 10min, at room temperature. The filter was air dried and exposed to Kodak X-OMAT film at -70°C.

(iii) Probe preparation

PCR amplified DNA fragment probes were prepared by nick translation. Purified core gene PCR product was added to a reaction mix containing 20μM dGTP and dCTP, 25μCi α^{32}P-dATP, 25μCi α^{32}P-dTTP and 2units of DNA pol I/DNase I in buffer containing 50mM Tris.HCl pH7.8, 5mM $MgCl_2$, 10mM β-ME, 10μg/ml BSA and

incubated at 16°C for 1hr. The reaction was stopped by the addition of 10µl of 100mM EDTA, 50% sucrose. Bromophenol blue was added to colour the mix which was then loaded on a Sephadex G-50 (Pharmacia) column that had been equilibrated with TE. Fractions (approximately 0.25ml) were collected from the column and those containing the highest counts in the initial peak were pooled and stored as probe. Before use, the probe was denatured by boiling for 5min, followed by flash cooling on ice.

Protein	Residues deleted	Predicted molecular weight (kDa)
Full-length	-	24.3
Δ84-109	84-109	20.9
Δ81-121	81-121	18.8
Δ79-125	79-125	18.2
Δ60-117	60-117	16.7
T-128	129-183	16.9
Δcore-surface hybrid	81-121 replaced	23.6
5'F10	91-121	20.2
5'F20	101-121	21.5
5'F30	111-121	22.9
3'F10	81-111	20.2
3'F20	81-101	21.6
3'F30	81-91	22.7

Table 2B1.11c: Summary table of core deletion proteins and core deletion protein modifications

The names of all core deletion proteins are shown, along with a description of the deletion or other modification in the protein. The predicted molecular weights of all proteins are also shown. All molecular weights were calculated without the inclusion of the pp65 epitope tag (predicted molecular weight=1.3kDa).

2B2 Cell culture

2B2.1 Growth of cells

Mammalian-derived cells were passaged in sterile, disposable 175cm^2 plastic flasks (Nunclon) in the appropriate media (Section 2A6) and were grown at 37°C in a humidified incubator under 5% CO_2. Confluent monolayers were harvested by washing the monolayer with 20ml versene, followed by washing in 25ml trypsin:versene (1:4) (supplied by the Institute of Virology Media Services) and resuspending the cells in 10ml medium.

*Sf*21 cells were grown at 28°C in the absence of CO_2 and trypsinisation was unnecessary, as cells could be removed from flasks by tapping.

For continual passage, all cells were split in a 1:10 ratio every 3-4 days.

2B2.2 Infection of *Sf*21 cells with recombinant baculovirus and plaque assays

For preparation of viral DNA, and ^{35}S-methionine labelling experiments, *Sf*21 cells were seeded in 35mm plates at 1x10^6 cells per plate and infected at a moi of 10. For the production of high titre viral stocks, 2x10^7 cells were seeded in 175cm^2 flasks and infected at a moi of 0.1. For protein purification experiments, cells were grown in roller flasks and infected at a moi of 10 when the cell density was 1-2x10^6 cells/ml.

Medium was removed from cell monolayers and the virus suspension added and allowed to adsorb to the cells for 1hr at room temperature. The virus suspension was then removed from the monolayer, an appropriate volume of TC100 Medium added and the cells incubated as normal. For infection of *Sf*21 cells in roller flasks, virus was added to the cell suspension and incubation continued as normal.

Plaque assays to determine titres of virus stocks were carried out on *Sf*21 cells seeded in 35mm diameter plates, as above. Seaplaque low gelling temperature agarose (3% in water (Flowgen)) was melted, cooled and added to an equal volume of pre-warmed TC100 Medium. After infection of the cells with a range of dilutions of virus suspension, and removal of the medium, 2ml of the agarose mix was added to the plates and allowed to set. TC100 Medium (2ml) was added and the plates incubated for 5 days. Plaques were visualised by adding 1ml 0.4% neutral red solution (Institute of Virology Media Services) diluted 1:50 with TC100, to the plates, followed by incubation in the dark for 2hrs. The stain was removed and the plates inverted and allowed to dry before plaques were counted.

2B2.3 ^{35}S-methionine labelling of infected *Sf*21 cells

*Sf*21 cells were seeded in 35mm diameter plates, as above, and infected at a moi of 10 with recombinant baculovirus. At the required time post-infection, medium was removed and replaced with 1ml TC100-methionine (Gibco) and the cells incubated for

30min at 28°C. A further 1ml of TC100-methionine containing 30μCi ^{35}S-methionine was added to the plate and the incubation continued for up to 18hr. Cells were scraped into the medium, pelleted by centrifugation at 6k rpm (microfuge) for 3min at room temperature and the cell pellet washed once with cold PBS. The pellet was resuspended in 0.5ml TE, mixed with SDS PAGE sample buffer and heat denatured by boiling for 10min. Duplicate infections were carried out without radiolabelling and the extracts analysed by Western blotting. Cellular extracts were stored at -70°C.

2B2.4 Calcium phosphate-mediated transfection

Cells were seeded on 13mm glass coverslips in 35mm (HepG2 cells) or 60mm (COS 7 cells) diameter plates at $2x10^5$ cells per plate for HepG2 cells and $1x10^6$ cells per plate for COS 7 cells, in 2ml of the appropriate medium and incubated overnight at 37°C prior to transfection. Plasmid DNA (5μg) was made up to a volume of 125μl in a solution of 250mM $CaCl_2$ and to this 125μl of 2xHBS was added dropwise, before vortexing for 5sec. The solution was incubated for at least 20min at room temperature before addition to the cells. After mixing by gently swirling the plates, the cells were incubated for 16-20hr at 37°C. The medium was removed and replaced by a 2ml aliquot of fresh medium and the incubation continued for a further 24hr.

HepG2 cells were blocked at the G_1/S boundary of the cell cycle prior to transfection and immunofluorescence assays. Approximately 8hr before transfection, cells were incubated in fresh medium containing 2μg/ml aphidicolin. Cells were then maintained in aphidicolin-containing medium until assayed by immunofluorescence.

Transfected COS7 cells were assayed for expression of core proteins by Western blot analysis (Section 2B3.4). Cells were scraped into the medium and pelleted by centrifugation at 6k rpm for 3min (microfuge) at room temperature. The pellet was washed once with cold PBS, resuspended in 0.5ml PBS and mixed with SDS PAGE sample buffer, before heat denaturation for 5min in a boiling water bath.

2B2.5 Immunofluorescence

Transfected cell monolayers were washed twice with PBS for 5min and fixed at room temperature for 10min with cold methanol. Cells were washed three times for 5min with PBS and permeabilised for 15min at room temperature with 0.5% Triton X-100 in PBS. The cells were again washed three times for 5min at room temperature in PBS and then washed briefly in PBS containing 0.05% Tween 20. Aliquots (20μl) of the relevant primary antibody diluted in PBS were then added to each coverslip for 60min. The anti-pp65 mouse monoclonal antibody was used at a 1:300 dilution and the Zymed anti-core rabbit polyclonal antiserum was used undiluted. The cells were washed three times for 5min at room temperature with PBS containing 0.05% Tween 20 and then incubated with secondary antibody (20μl) diluted in PBS. The goat anti-mouse FITC

was used at a 1:100 dilution and the goat anti-rabbit FITC conjugate was used at a 1:60 dilution. Incubations with secondary antibodies were carried out for 30min, after which the cells were washed three times for 5min at room temperature in PBS containing 0.05% Tween 20. The cells were air dried and mounted on glass slides with 10ml Citifluor, a glycerol/PBS solution (UKC Chemical Laboratories). Cells were examined using a Nikon Microphot-SA fluorescence microscope and photographed using Kodak ASA400 black and white film.

2B3 Detection and purification of core particles

2B3.1 Core protein expression and purification using recombinant baculovirus

The procedure for core protein expression and purification was based on that of Takehara *et al.* (1988). High titre stocks of recombinant baculoviruses were produced by continual rounds of low multiplicity infection of *Sf*21 cells, until the titre was at least $1x10^8$ pfu/ml, as determined by plaque assay. Typically, *Sf*21 cells were grown in 300ml volumes in roller bottles until the cell density had reached $1\text{-}2x10^6$ cells/ml, when cells were infected with the relevant virus at a moi of 10 and the incubation continued for a further 4 days. Cells were collected by centrifugation at 3k rpm (Sorvall RT6000B centrifuge), the pellet resuspended in 3ml TNE I buffer and disrupted by sonication for 15sec with a Branson 450 soni-probe. Soluble protein was removed after centrifugation at 10k rpm (Sorvall SS34 rotor) for 10min at 4°C to remove cell debris, and stored at 4°C until required.

Attempts were made to purify core particles using two approaches, isopycnic CsCl centrifugation and sucrose density gradient centrifugation, as described below:

(i) Isopycnic CsCl centrifugation

Centrifugation was performed according to the method of Takehara *et al.* (1988). Soluble extract from infected *Sf*21 cells was centrifuged through a 30% sucrose cushion (30% sucrose in TNE I) at 24k rpm (Sorvall TST41 rotor) for 14hr at 4°C. The resulting pellet was resuspended in 200µl TNE I and layered onto CsCl solution (1.25g/ml in TNE I buffer). The gradient was centrifuged at 35k rpm (Sorvall TST41 rotor) for 35hr at 18°C and 1ml fractions removed from the top of the gradient. Any pelleted material was resuspended in 1ml TNE I and all fractions analysed on Coomassie stained polyacrylamide gels (Section 2B3.3)

(ii) Sucrose density gradient centrifugation

The protocol for purification of core particles using sucrose gradients was adapted from that supplied by the group of Prof. K. Murray (University of Edinburgh). Linear 10-50% sucrose gradients (10ml volume) were prepared in TNE I buffer, soluble

extracts from infected *Sf*21 cells layered onto the gradient and centrifuged at 35k rpm (Sorvall TST41 rotor) for 5hr at 4°C. Fractions were removed from the top of the gradient and analysed by Western blotting.

2B3.2 Core protein expression and purification from *E.coli*

Core proteins were expressed in *E.coli* using two plasmid systems, pR1-11 and pET plasmids, as described below:

(i) Core protein purification using plasmid pR1-11

Plasmid pR1-11 was kindly supplied by the group of Prof. K. Murray (University of Edinburgh). A single fresh bacterial colony transformed with the relevant pR1-11 plasmid was inoculated into 10ml L-broth containing ampicillin and grown up overnight at 37°C with shaking. An aliquot (6ml) of this culture was added to 300ml pre-warmed (37°C) L-broth containing ampicillin and the constitutive protein expression carried out at 37°C with shaking for 4hr. Bacteria were harvested by centrifugation at 6k rpm (Sorvall GS3 rotor) for 15min at 4°C and the cell pellet resuspended in 10ml sonication buffer prior to sonication in 20sec bursts with a soni-probe. Cell debris was removed by centrifugation at 10k rpm (Sorvall SS34 rotor) for 10min at 4°C. Ammonium sulphate was added to the supernatant to 30% volume and the mixture stirred on ice for 1hr. Precipitated proteins were pelleted by centrifugation at 10k rpm (Sorvall SS34 rotor) for 10min at 4°C and the pellet resuspended in 3ml TN buffer before overnight dialysis against TN buffer at 4°C. Precipitated protein was removed by centrifugation at 10k rpm (Sorvall SS34 rotor) for 10min at 4°C and the supernatant layered on 10-50% sucrose gradients.

Modifications to this protocol included omission of the overnight dialysis step, with the resuspended ammonium sulphate pellet layered onto the gradient, or precipitation of nucleic acids by the addition of PEI, prior to ammonium sulphate precipitation of protein. The cellular extract was made up to 0.5M NaCl to disrupt any protein:DNA interactions and nucleic acid precipitated by the addition of PEI to 0.2%. After incubation for 5min on ice, precipitated nucleic acid was removed by centrifugation at 15k rpm (Sorvall SS34 rotor) for 15min at 4°C. Protein was then precipitated from the supernatant, as above.

(ii) Core protein purification using pET plasmids

A single fresh *E.coli* colony (strain BL21(DE3)pLysS), transformed with either pET3a or pET9a plasmids containing core gene sequences, was inoculated into 10ml 2YT broth containing appropriate antibiotics and grown shaking overnight at 37°C. A sample (200µl) was added to 50ml pre-warmed (37°C) 2YT broth with antibiotics and the culture grown at 37°C with shaking until the OD_{630} reached 0.5. Protein expression

was induced by adding IPTG to 300μM and incubating for a further 4hr. The bacteria were harvested by centrifugation at 3k rpm (Sorvall RT6000B centrifuge) for 20min at 4°C and the pellet stored frozen overnight at -20°C. After thawing, 3ml of TNE II, 8μl of 10mg/ml DNase I and 30μl of 10mg/ml RNase were added and the mixture incubated on ice for 30min to allow cell lysis. Complete lysis of the cells was ensured by sonication in 20sec bursts with a soni-probe. Cell debris was pelleted by centrifugation either at 13k rpm (microfuge) or at 35k rpm (Beckman TLA 100.3 rotor) for 20min at 4°C and soluble protein removed. Aliquots of protein removed before sonication, as total protein samples, were stored at 4°C, as was soluble protein prior to purification on sucrose gradients.

Centrifugation and fraction analysis were as described in Section 2B3.1(ii).

2B3.3 SDS polyacrylamide gel electrophoresis (SDS PAGE) of proteins

Proteins were resolved by electrophoresis through SDS polyacrylamide minigels (Laemmli, 1970) using the Bio-Rad Miniprotean II apparatus. Gel mixes were made up as follows:

Solution	12% resolving gel	17.5% resolving gel	Stacking gel
30% acrylamide (2.7% cross-linker)	8ml	11.7ml	1.3ml
RGB	5ml	5ml	-
SGB	-	-	2.5ml
10% SDS	0.2ml	0.2ml	-
water	6.7ml	3ml	6.2ml

A 10μl aliquot of TEMED and 100μl of 10% APS were added to each resolving gel mix and 10μl TEMED and 50μl 10% APS to stacking gel mix.

Resolving gel mixes were poured into glass plate sandwiches in the gel-former apparatus, overlaid with 80% isopropanol and allowed to set. The isopropanol was washed away with water, the stacking gel mixture overlaid and a comb inserted. Protein samples were mixed with SDS PAGE sample buffer and boiled for 10min prior to loading. Gels were run at 200V until the bromophenol blue dye reached the bottom. The gels were stained with Coomassie blue for 15min and destained in protein gel destain. Gels with radiolabelled protein samples were fixed in protein gel destain for 15min, incubated in En^{3}Hance for 30min, washed with water for 10min and vacuum dried prior to exposure to Kodak X-OMAT S film.

2B3.4 Western blot analysis of proteins

(i) Electroblotting to nitrocellulose

Proteins resolved on SDS PAGE minigels were transferred to nitrocellulose by the method of Towbin *et al.* (1979) in a Bio-Rad mini transblot cell. A blotting sandwich was set up with the gel in contact with the nitrocellulose (Schleicher and Schuell Inc.) and both were sandwiched between Whatman 3MM paper. This was in turn sandwiched between sponges provided by Bio-Rad. All materials were equilibrated in blotting buffer for at least 20min prior to assembly of the blotting sandwich. Blotting was carried out at 250mA for 3hr.

(ii) Immunodetection

Nitrocellulose membranes were blocked by incubation at 37°C with gentle shaking for a minimum of 60min in blocking buffer, with one change of buffer. Membranes were then washed twice in TTBS for 5min at room temperature before incubation with the primary antibody, diluted appropriately in 1% gelatin in TTBS, at room temperature with shaking. Incubation with primary antibody was for a minimum of 4hr. Membranes were washed twice for 10min in TTBS and anti-mouse IgG whole molecule or anti-rabbit IgG whole molecule peroxidase conjugate, diluted in 1% gelatin in TTBS, added. After a 2hr incubation at room temperature, again with shaking, membranes were washed four times for 15min in TTBS and proteins detected using the Amersham enhanced chemiluminescence (ECL) system. The two reagents were mixed in equal volumes and a total of 3ml poured onto the membrane, which was then agitated for 1min. The membranes were wrapped in cling film and exposed to Kodak X-OMAT film for 5-45sec.

On blots from the initial experiments, a colour development reaction was used as to detect proteins on the membrane. The substrate for the peroxidase conjugate was prepared by mixing 20ml cold methanol, containing 60mg 4-chloro-1-naphthol, with 100ml TBS, containing 60µl of 30% hydrogen peroxide. After incubation with secondary antibody, blots were washed twice for 5min in TTBS and twice for 5min in TBS, both at room temperature and the freshly prepared substrate added immediately.

Blots were stripped before being probed with different primary antibody by incubation at 55°C for 45min in 60mM Tris, 0.7% β-ME, 2% SDS, with one change of solution after 20min. Membranes were then washed four times for 5min in TTBS at room temperature with shaking, before being blocked as before and incubated with the second primary antibody.

2B3.5 Detection of core particles by electron microscopy

Core protein containing fractions from sucrose gradients were pooled and dialysed overnight against the appropriate buffer at 4°C to remove the sucrose, followed by concentration by centrifugation at 6k rpm (Sorvall SS34 rotor) in Centricon 50 centrifugation units (Amicon) at 4°C.

Samples were examined by negative stain electron microscopy, performed by Mr. J. Aitken or Dr. F. Rixon. The concentrated sample was mixed with an equal volume of either 1% (w/v) phosphotungstic acid, pH 7.0, or sodium silicotungstate, before being placed onto a 200 mesh copper grid coated with parlodion and allowed to stand for 2-3min at room temperature. Excess liquid was removed using Whatman filter paper and the sample examined at a magnification of 40 000 on a Jeol 1200EX2 electron microscope, operating at 80kV.

2B3.6 Agarose gel assay for core particles

Core particles can be detected on agarose gels by Coomassie staining of the protein and detection of the encapsidated nucleic acid under UV transillumination, in the presence of EtBr (Birnbaum and Nassal, 1990). Concentrated core protein containing samples were mixed with agarose gel loading buffer and loaded onto a 1% agarose gel in TAE, containing 0.5mg/ml EtBr. Electrophoresis was carried out at 70V and the nucleic acid visualised under UV transillumination. The gel was then stained with Coomassie blue and destained for protein visualisation. Gels were photographed using The Imager (Appligene).

CHAPTER 3 RESULTS

3A Expression and purification of full-length core protein and core deletion protein Δ84-109 using recombinant baculoviruses

3A1 Introduction

The purpose of this project was to characterise possible functions of naturally occurring core proteins from which central regions were deleted. The initial aim was to determine whether or not these core deletion proteins retained the ability to spontaneously assemble into core particles, or more properly, core particle-like structures. The method chosen and initiated by the group of Dr. N. Naoumov (Kings College, London) was expression of these proteins in *Sf*21 cells using recombinant baculoviruses. The baculovirus *Autographica californica* multiple nuclear polyhedrosis virus (AcMNPV) has been widely used as an expression vector (see review by Bishop, 1992) and has been used previously by the group of Bishop to produce and analyse wild-type core particles (Takehara *et al.*, 1988 and Hilditch *et al.*, 1990). Before the work presented in this thesis was started, the group at Kings College had constructed two recombinant baculoviruses: one containing the full-length core gene and the other containing the gene for the core deletion protein Δ84-109. These viruses were constructed using the Invitrogen transfer vector pVL1393, placing the cloned genes under the control of the polyhedrin promoter. Initial experiments by this group had shown expression of the full-length core protein on Coomassie stained polyacrylamide gels, but were unable to detect the core deletion protein using this method.

This section describes the experiments carried out to authenticate these recombinant baculoviruses and to detect expression of both proteins. Initial attempts at purification of core particles or core particle-like structures from infected *Sf*21 cells are also described and finally, details of problems encountered using this particular system for the expression and purification of core particles.

3A2 Authentication of recombinant baculoviruses by Southern blotting

Before any further expression experiments were carried out it was necessary to characterise the recombinant baculoviruses to ensure that the material received from the group at Kings College had not been damaged or degraded in any way. It was decided to extract viral DNA from infected *Sf*21 cells and confirm the presence of the core genes by Southern blotting (Figure 3A2). The transfer vector pVL1393, containing either the core gene or the core deletion gene, used to construct these viruses had also been supplied and were included as size controls for the genes digested from the baculovirus DNA. The blot

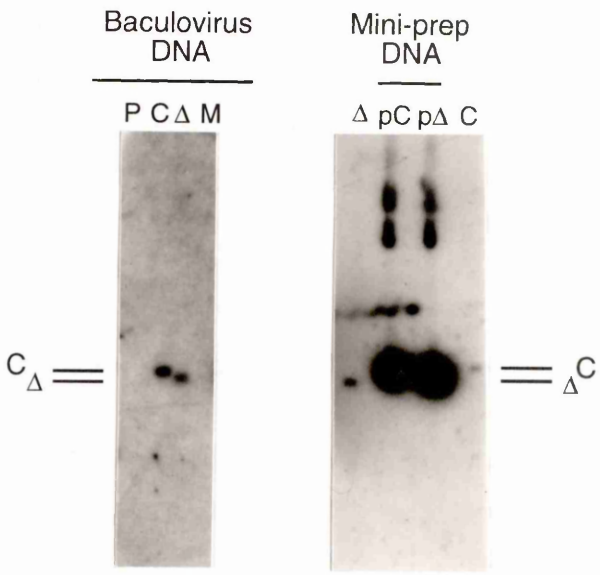


Figure 3A2: Southern blot of recombinant baculovirus DNA

DNA was extracted from *Sf*21 cells infected, at a moi of 10, with recombinant baculovirus containing the full-length core gene (C), core gene with nucleotides 2152-2229 deleted (Δ), AcPAK6 baculovirus (P) or mock-infected cells (M). *Xba* I/*Pst* I digests of baculovirus transfer vector pVL1393 mini-prep DNA containing the full-length core gene (pC) or core gene with nucleotides 2150-2227 deleted (pΔ) were used as controls. Nucleic acid was transferred to nylon membrane by Southern blotting, and the membrane probed with a ^{32}P-labelled core gene PCR product, amplified using primer pair C1/C2.

showed single bands, indicating fragments which had hybridised to the labelled core gene probe, in the lanes containing digested baculovirus DNA. The sizes of these bands were equal to those of their corresponding controls, indicating that the construction of the recombinant viruses was correct. No fragments hybridised in the lanes containing digested DNA from cells infected with AcPAK6 virus, which contains the gene for β-galactosidase under the control of the polyhedrin promoter, or from mock infected cells, indicating that the hybridisation observed was specific for the core gene. The recombinant viruses were further characterised by fully sequencing the inserts in their transfer vectors (results not shown), which confirmed the size and position of the deletion in the core gene as identical to that published (Ackrill *et al*, 1993).

3A3 Detection of expression of full-length and Δ84-109 core proteins by ^{35}S-methionine labelling

As stated previously, initial results from the Kings college group showed that expression of the full-length core protein by the recombinant baculovirus could be detected on a Coomassie stained gel, but the core deletion protein was not visible by this method. This may have been due to the fact that its predicted size would result in co-migration with other Coomassie stained bands or that expression levels for this protein were not as high, when compared to those of the full-length protein. The possibility of the core deletion protein being toxic to the cells was only briefly considered, as expression from the late polyhedrin promoter occurs in insect cells after the production of mature, infectious virus particles. This prevents synthesis of cytotoxic proteins from adversely affecting virus replication and spread to other cells. It was decided to investigate expression of both proteins by radiolabelling virus-infected cell proteins with ^{35}S-methionine. High titre stocks of recombinant viruses containing full-length core and core deletion genes and AcPAK6 virus had been produced with titres of $2x10^8$, $1x10^8$ and $2x10^8$ pfu/ml respectively, on *Sf*21 cells. These stocks were used to synchronously infect cells at a high moi and at 24hr and 48hr pi cells were incubated in labelling medium for an 18hr period, before total cellular extracts were prepared.

As was seen from the gel (Figure 3A3), at 24hr pi proteins of the approximate predicted size were being synthesised (see Table 2B1.11c), but the core deletion protein was present at lower levels compared to the full-length protein. This may have reflected the differences in efficiency of expression for different proteins, as can occur with any expression system. By 48hr pi the full-length core protein represented one of the major proteins present in the cells, confirming the initial results from Coomassie stained gels. The core deletion protein seemed to be present at moderate levels, however this was difficult to determine precisely, due to the presence of several co-migrating proteins which had also been labelled. For both time-points examined the levels of β-galactosidase

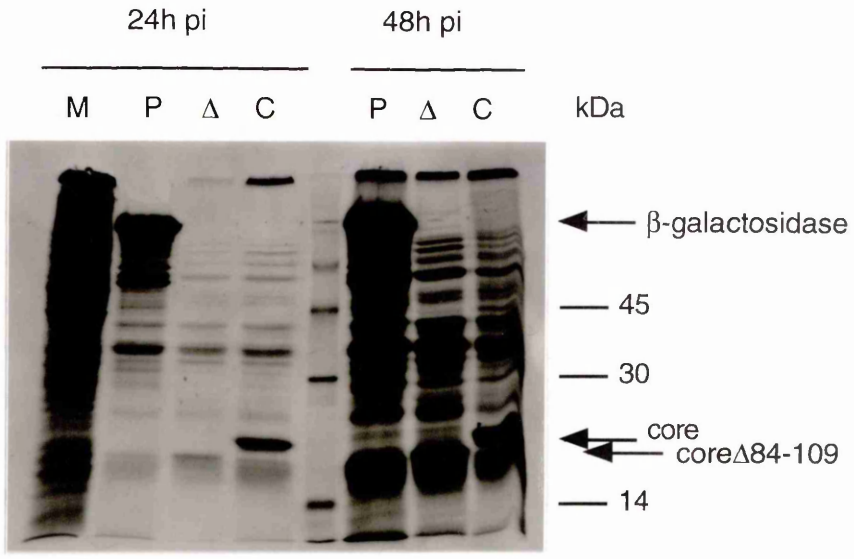


Figure 3A3: ^{35}S-methionine labelling of infected *Sf*21 cells

*Sf*21 cells in 35mm dishes were infected, at a moi of 10, with recombinant baculoviruses expressing full-length core protein (C), core protein with amino acids 84-109 deleted (Δ), AcPAK6 baculovirus (P) or mock-infected (M). Cells were labelled with ^{35}S-methionine for 18h at 24h and 48h pi. Cellular extracts were separated on a 12% polyacrylamide gel, which was dried down and exposed to film. The sizes of the radiolabelled molecular weight markers used are shown (kDa).

protein produced by the AcPAK6 virus were higher than the levels of core proteins produced by the two other viruses. This was probably because the expression of β-galactosidase by this virus has been thoroughly optimised for commercial purposes. Although it was not possible to obtain the original wild-type AcMNPV DNA from which the recombinant viruses were constructed, detection of novel proteins of the correct molecular weight in cells infected by these viruses was good evidence that these were the core proteins, and not cellular or other viral proteins.

3A4 Detection of expression of full-length and Δ84-109 core proteins by Western blotting

In order to confirm that the proteins of the correct predicted molecular weights detected in the previous section were full-length and core deletion proteins it was decided to use Western blotting and detection with antibodies raised to core protein. A number of monoclonal antibodies raised against either core protein or e antigen which had been given as gifts from other groups were available, however the precise backgrounds as to which protein these antibodies had been raised against, or which epitopes they identified, was unknown. In an initial Western blot experiment (results not shown), using cellular extracts from infected *Sf*21 cells at different times pi, one of these monoclonal antibodies reacted with a protein of 21kDa present in the extract from cells infected by the full-length core gene-containing baculovirus, and not in extracts from cells infected by baculovirus containing the core deletion gene or AcPAK6 or in mock-infected cells. This confirmed the identity of the labelled band from Section 3A3 as core protein, and showed that the protein was synthesised at least up to 120hr pi. However, none of the extracts from cells infected with core deletion-containing baculovirus showed any reactive bands. As core protein has one immunodominant B-cell epitope which has been mapped between amino acids 74-89 (Salfeld *et al.*, 1989), the likely cause of non-reactivity in these extracts was that this epitope was partially within the deletion. In view of this result it was decided that a variety of different monoclonal and, more importantly, a polyclonal antibody raised against core protein should be assayed for the ability to detect the core deletion protein on Western blots. At 48hr pi, extracts were prepared from cells infected with core and core deletion baculoviruses and Western blotted (Figure 3A4). The membranes were cut into strips, with each strip containing a lane each of extract from cells infected with full-length core and core deletion baculovirus. Antibodies αHBe 141/152 and 87-141-03 were used at an experimental dilution of 1:1000 and all other antibodies used at the recommended dilution, as indicated. All of the monoclonal antibodies raised against core protein also detected the core deletion protein, which was unexpected, although mAb BFαHBc2 and BFαHBc1 were less effective. It may have been that these antibodies, despite the immunodominancy of the 74-89 epitope, had been raised against other epitopes of the protein, which were unaffected by the Δ84-109 deletion. Monoclonal antibodies

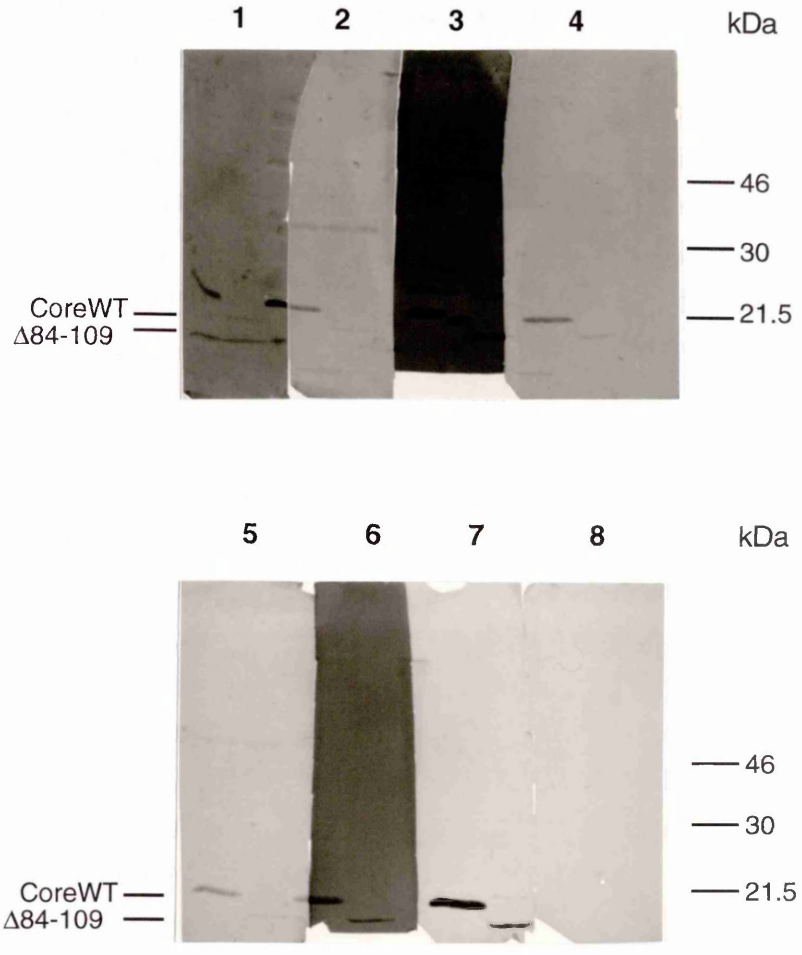


Figure 3A4: Western blots of *Sf*21 infected cell extracts

*Sf*21 cells were infected, at a moi of 10, with recombinant baculoviruses expressing full-length core protein (CoreWT) or core protein with amino acids 84-109 deleted (Δ84-109). Cellular extracts were separated on 12% polyacrylamide gels and transferred to nitrocellulose membranes by Western blotting. After blocking, the membranes were incubated with the primary antibodies listed below and specific binding detected with HRP-conjugated secondary antibody. Full-length core protein is on the left and core deletion protein Δ84-109 on the right of each nitrocellulose strip.

Antibodies used:

1 mAb BFαHBc2 - 1:10
2 mAb BFαHBc3 - 1:10
3 Zymed pAb - undiluted
4 mAb BFαHBc4 - 1:10
5 mAb BFαHBc1 - 1:10
6 mAb 42B12 - 1:750
7 mAb αHBe 141/152 - 1:1000
8 mAb αHBe 87-141-03 - 1:1000

BFαHBc3 and BFαHBc1 also showed reactivity with proteins of approximately 34kDa and 50kDa, respectively. The core deletion protein seemed to be detected less efficiently than the full-length protein by all monoclonal antibodies used, but this may have simply been a consequence of the lower levels of expression of this protein already observed in the infected cells. The monoclonal antibody 42B12 was chosen for further use, due to the fact it was capable of detecting both proteins and the epitope it recognised had been identified as being approximately amino acids 115-135 (W.H.Gerlich - personal communication). In addition, we had been kindly supplied with a reasonable quantity to use with this project.

As hypothesised, the Zymed polyclonal antibody detected both the full-length protein and the core deletion protein. This was probably due to detection of a number of different epitopes throughout the protein, so increasing the chance of core deletion protein detection. A number of other commercial polyclonal antibodies were also tested in this manner and were shown to have the ability to detect both proteins (results not shown). The polyclonal antibody used, however, reacted with many other proteins on the blot. It may have been of value to repeat the experiment with an increased period of blocking the membrane to try and reduce the non-specific nature of this binding, but as the other bands were larger in size than the core proteins and did not interfere with their detection, this was not thought to be essential.

Only one of the antibodies raised against e antigen (αHBe 141/152) detected both the proteins. As core protein and e antigen are known to have epitopes in common, antibodies raised against one have been known to react with the other, as the result here has shown. The 87-141-03 antibody, again raised against e antigen, did not react with either protein. It may be that the experimental dilution used for this antibody was too high, compromising its detection ability. As all monoclonal antibodies gave a fainter band for the core deletion protein, nothing can be stated conclusively about the ability of this antibody to detect this protein, at the dilution used.

3A5 Purification of core particles using recombinant baculoviruses

The production of hepatitis B virus core particles has been achieved previously by expression of the core gene in recombinant baculoviruses, under the control of the polyhedrin promoter (Takehara *et al.*, 1988 and Hilditch *et al.*, 1990). The method of Takehara *et al.* consisted of infection of *Sf*21 cells with recombinant baculovirus, containing the full-length core gene, at a high moi for 4 days. Cellular extracts were prepared by sonication and clarified by centrifugation before being subjected to centrifugation through a 30% sucrose solution. After the resulting pellet was resuspended, core particles were purified to homogeneity by CsCl isopycnic centrifugation. The method of Hilditch *et al.* consisted of the same CsCl centrifugation step followed by a shallow glycerol gradient to remove two contaminating proteins. It

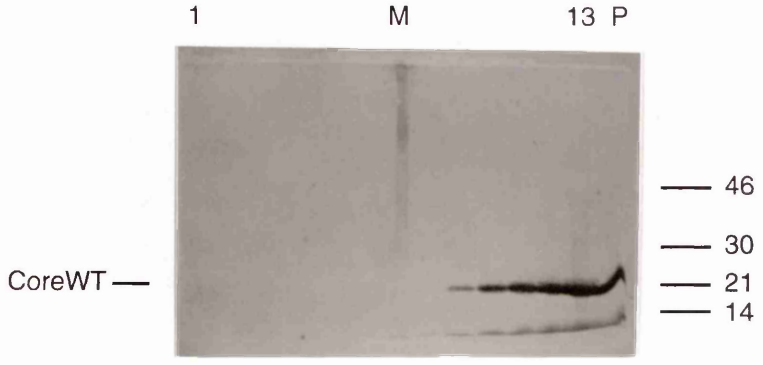


Figure 3A5a: Purification of core protein expressed by recombinant baculovirus

*Sf*21 cells in roller flasks were infected, at a moi of 10, with recombinant baculovirus expressing full-length core gene. Infected cells were collected by centrifugation 4 days pi and extracts subjected to sucrose gradient centrifugation on a 10-70% gradient. Fractions were collected from the top (fraction 1) to the bottom (fraction 13) of the gradient and any pelleted material was resuspended (P).Core protein was detected by Western blot analysis using mAb 42B12. Molecular weights of the markers (M) used are shown.

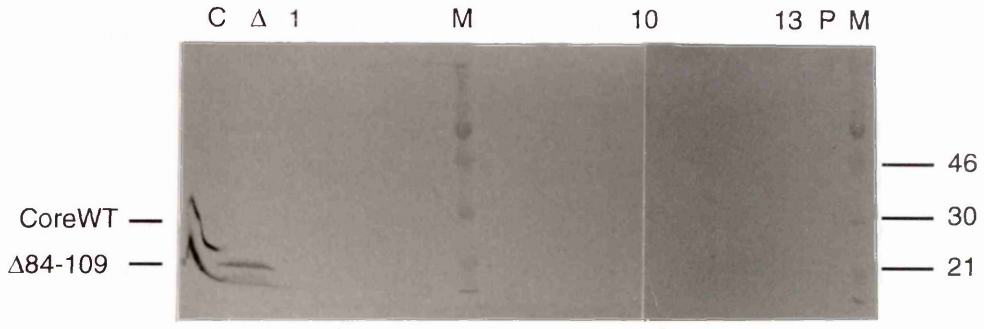


Figure 3A5b: Purification of core deletion protein expressed by recombinant baculovirus *Sf*21 cells in roller flasks were infected, at a moi of 10, with recombinant baculovirus expressing full-length core gene. Infected cells were collected by centrifugation 4 days pi and extracts subjected to sucrose gradient centrifugation on a 10-70% gradient. Fractions were collected from the top (fraction 1) to the bottom of the gradient (fraction 13) and any pelleted material was resuspended (P). Fractions were electrophoresed through 12% polyacrylamide gels alongside extracts from cells infected with recombinant baculoviruses expressing full-length core protein (C) or core deletion protein Δ84-109 (Δ). Proteins were detected by Western blot analysis using a Zymed polyclonal antibody. Molecular weights of the markers (M) used are shown.

was decided to follow the former method initially, using *Sf*21 cells grown in roller flasks. Cells were infected at a moi of 10 with recombinant baculovirus containing the full-length core gene and collected by centrifugation 4 days pi. However, after CsCl centrifugation, no visible band of core particles was observed, as was recorded in the original paper. Analysis of the fractions resulting from this step on Coomassie stained gels showed no evidence for core protein at any point in the gradient (results not shown). Gradient conditions, such as the initial density of CsCl and the rotor used for centrifugation, were altered but no core particle purification was achieved. After these attempts, contact was made with the group of Prof. K. Murray (University of Edinburgh). This group purified core particles from a bacterial expression system, using 30% ammonium sulphate precipitation of bacterial extracts, followed by centrifugation of the resuspended pellet through an 8-45% sucrose gradient. However, when this method was applied to infected *Sf*21 cell extracts prepared as before, core protein was found to precipitate out of solution during an overnight dialysis step used to remove ammonium sulphate from the resuspended pellet. The method was consequently altered to remove the ammonium sulphate precipitation step and instead the infected *Sf*21 extract was centrifuged through the sucrose gradient. Upon analysis of the fractions from this gradient on Coomassie stained gels, the only position in the gradient where core protein was observed was in the pellet which formed at the bottom of the centrifugation tube. Only when the concentration of sucrose was increased and the extracts centrifuged through a 10-70% gradient, was core protein actually detected in the lower gradient fractions (fractions 9-13, Figure 3A5a). However, when the protein-containing fractions were dialysed against TNE1 buffer and concentrated approximately five-fold for examination by electron microscopy, no 28nm particles were detected. The only particulate objects present were approximately four times as large, and few in number. When an identical purification procedure was carried out on extracts of cells infected by recombinant virus containing the core deletion gene, no protein was detected in any of the gradient fractions (Figure 3A5b). These results are very inconclusive as, although core particles appeared to have been purified using the sucrose gradient, particles of the correct size were not observed by electron microscopy. The observation of the larger particulate structures does correlate with the fact that a higher concentration of sucrose was required (70% compared to 45% used by the group at the University of Edinburgh) in the purification. From these results, no valid conclusions could be drawn regarding the purification of the core deletion protein. Although protein was detected in the extract loaded onto the gradient, none was detected in the fractions after centrifugation had taken place. This may have been because the protein would have been diluted in the larger fraction volume and, consequently, the smaller absolute amount of protein loaded onto the gel may have fallen below the sensitivity limit for the antibody used. The larger particulate structures observed from the full-length core protein purification seemed to suggest that the spontaneous assembly of

core particles had been incorrect or that at some stage during the procedure, perhaps by the sonication stage, the core particles had been damaged or aggregated together. It was decided that in view of these results and the difficulties that still existed in setting up a satisfactory purification procedure for wild-type core particles from infected *Sf*21 cell extracts, that changing to a bacterial expression system may aid in solving these problems.

3A6 Discussion

The results in this section have shown the expression and detection of both full-length core protein and core deletion protein Δ84-109 from *Sf*21 cells infected by recombinant baculoviruses. Detection of both proteins was initially by ^{35}S-methionine labelling of infected cell proteins but was further enabled by a variety of monoclonal and polyclonal antibodies, raised against either core protein or e antigen, which were used on Western blots of infected cell extracts. Attempts at purification of wild-type core particles by either CsCl or sucrose gradient centrifugation were unsuccessful for undetermined reasons.

Takehara *et al.* (1988) showed that recombinant baculovirus-expressed core protein constituted 40% of the Coomassie-stained cellular proteins, corresponding to a yield of 5mg/l cell culture. Hilditch *et al.* (1990) achieved a yield of 35mg/l of cell culture. Although no quantitation was performed, visual comparisons of ^{35}S-methionine labelled infected cell extracts showed that these expression levels were much higher than the levels achieved by the recombinant baculoviruses used in this section. This may be due, in part, to the different transfer vectors used to produce the recombinant viruses and also due to a decreased efficiency of core particle expression or recovery from larger volume cultures. Growth and infection of *Sf*21 cells in multiple smaller cell culture flasks, as opposed to roller flasks, may have increased the efficiency of infection as the virus suspension would be in closer contact with the cell monolayer. Consequently, the levels of core protein produced would be higher.

Particles of approximately 30nm were also observed by Takehara *et al.* (1988) and Hilditch *et al.* (1990) in electron microscopic examination of peak core protein-containing gradient fractions. Although it appeared that core particles had been purified in the results presented here (Figure 3A5a), upon examination of peak core protein-containing fractions the only particulate structures observed were too large to represent wild-type core particles. It seems unlikely that these structures arose from aggregation of incorrectly folded core protein as *Sf*21 cells possess a normal eukaryotic translational pathway, allowing correct folding and disulphide bond assembly in proteins. Insect cells have also been used to produce other complex structures, such as polyoma and parvovirus capsids (Montross *et al.*, 1991 and Brown *et al.*, 1991), as well as the previous mentions of core particle production by the group of Bishop. Electron microscopic examination of sections

of infected *Sf*21 cells would have shown whether particles of the expected size and morphology were initially assembled. These may then have been damaged or become insoluble during the subsequent purification steps. Similar examination of *Sf*21 cells infected with the recombinant virus expressing the core deletion protein Δ84-109 would also have revealed whether particles were formed by this protein. These particles may have been less stable than wild-type particles and not remained intact during purification. This may explain the presence of the Δ84-109 protein in infected cell extracts, but not in the sucrose gradient (Figure 3A5b).

Expression of protein in *E.coli* has many advantages over the baculovirus system. Bacterial plasmids are simpler to produce and manipulate, in contrast to the production of recombinant baculovirus genomes. The production of high titre virus stocks is not required and both cell growth and the induction of protein expression are simpler for bacterial systems. Although the prokaryotic expression pathway differs from that of eukaryotes, the previous expression and characterisation of core particles expressed in bacteria have shown that this system is able to reproduce the core particle assembly process.

3B Expression and purification of full-length core protein and core deletion proteins in a bacterial system

3B1 Introduction

In view of the problems encountered with purifying wild-type core particles from infected insect cells in our hands, it was decided to change to a bacterial expression system. Since it was first used by Cohen and Richmond (1982) and Stahl *et al.*(1982) to produce core particles from cloned HBV DNA that were morphologically and immunologically identical to those isolated from patients serum, *E.coli* has been successfully used by a large number of groups for the production of wild-type core particles. We had received a gift of plasmid pR1-11 from the group of Prof. K. Murray, which was the original plasmid used to produce core particles. Details of this plasmid are described in Section 2A1. The group of Dr. N. Naoumov had sent a further three core genes containing in-frame central deletions, Δ81-121, Δ79-125 and Δ60-117, and it was decided initially to attempt to purify these using this system, along with a full-length core protein control. In order to achieve this, both full-length and core deletion genes were amplified by PCR using primer pair cd1/cd2, which contained *Eco* RI and *Bam* HI sites respectively and would allow reconstruction of the correct 5' coding sequences in pR1-11 after removal of the excised *Eco* RI/*Bam* HI fragment. Throughout this section the full-length core protein with the additional sequence from pR1-11 is simply referred to as 'full-length protein' to distinguish it from the core deletion proteins, and particles produced from this protein are referred to as 'wild-type' core particles.

3B2 Expression of full-length core protein and core deletion proteins using pR1-11 and detection by Western blotting

The initial experiments carried out with pR1-11 were to determine whether the core proteins produced could be detected on Western blots, which would be necessary for monitoring of any purification procedure. Proteins were expressed routinely at 37°C for 4hr before cellular extracts were prepared by sonication and samples of soluble protein prepared by centrifugation of these extracts. Western blots of samples were performed using both anti-core mAb 42B12 and a number of commercial anti-core polyclonal antibodies (see Figures 3B2a and 3B2b).

On Western blots with anti-core mAb 42B12, although the full-length protein was expressed and detected, neither of the two core deletion proteins examined could be detected. This may confirm the approximate epitope position for this antibody (residues 115-135) as both of the deletions encroach upon this region. As the sizes of these deletions were larger than that of the core deletion protein expressed in baculovirus it was decided that use of a polyclonal antibody would maximise the probability of their

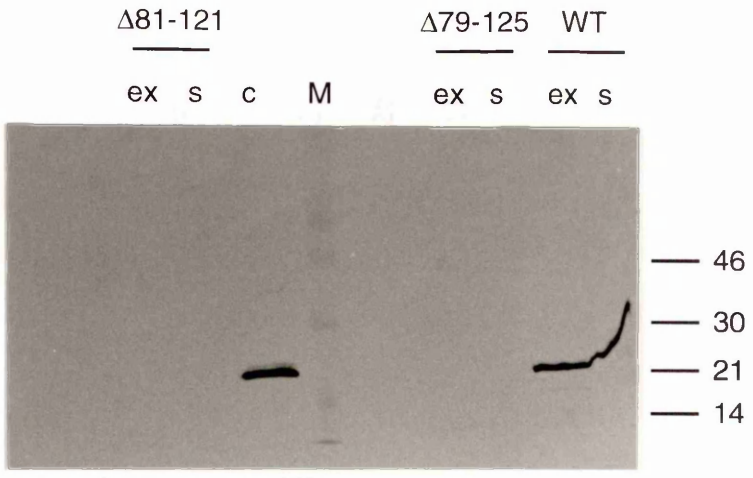
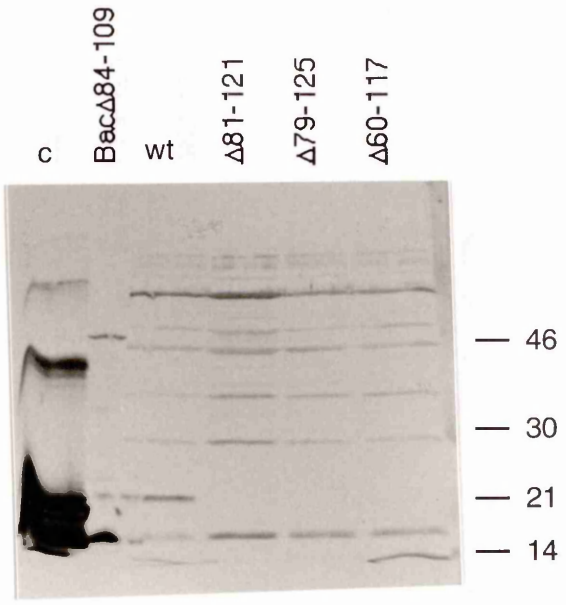


Figure 3B2a: Expression of full-length core protein and deletion variants
Full-length core protein (WT) and two core deletion proteins (Δ81-121 and Δ79-125) were expressed using plasmid pR1-11. Expression was carried out for 4 hrs at 37°C. Aliquots of total cellular extracts (ex) and soluble protein (s) were subjected to electrophoresis through 12% polyacrylamide gels and analysed by Western blotting. The blot was incubated with anti-core mAb 42B12. An aliquot of purified HBcAg (c) was included as a control and molecular weights of the markers used (M) are shown.

detection. Three anti-core polyclonal antibodies were obtained and all tested on three core deletion proteins (Figure 3B2b). All three antibodies were able to detect the full-length core protein expressed in this system. This appeared as a number of bands slightly differing in size from 14-22kDa. The lower molecular weight bands probably represent degradation products of the full-length protein. Proteins of approximately 44 and 64kDa were also observed. These are the correct sizes for higher order multimers of full-length core protein (see Section 3D3). Analysis of the extracts from cells expressing core deletion proteins Δ81-121 and Δ79-125 (predicted molecular weights of 18.8 and 18.2kDa, respectively) did show a protein of approximately the correct molecular weight. This protein was detected by all three polyclonal antibodies used. However, it was also present, albeit at lower levels, in the extracts from cells expressing full-length core protein. It is therefore possible that this protein was non-specifically detected by the polyclonal antibodies. Another protein of lower molecular weight (approximately 14kDa) was also detectable by all three antibodies. However, both the Life Sciences and Zymed

A



B

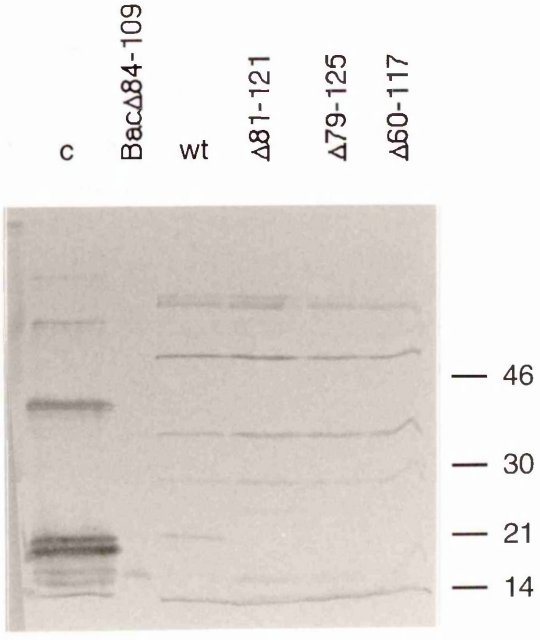


Figure 3B2b: Western blots of expressed core protein deletion variants (continued with figure legend on the following page)

C

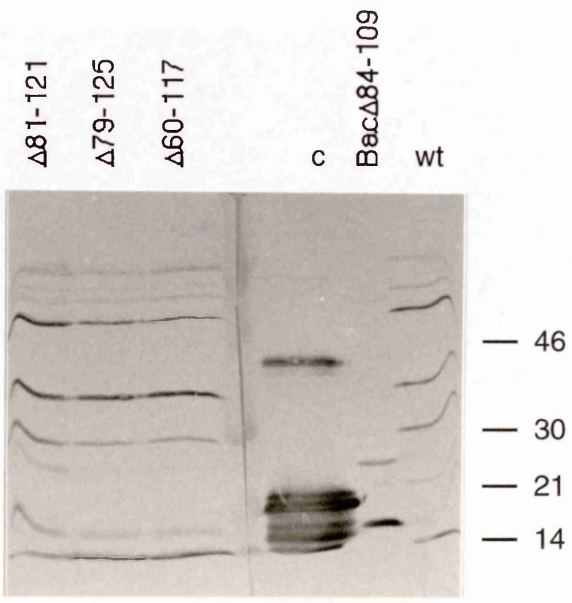


Figure 3B2b: Western blots of expressed core protein deletion variants (continued) Full-length core protein (wt) and deletion variants Δ81-121, Δ79-125 and Δ60-117 were expressed using plasmid pR1-11. Soluble protein produced from cellular extracts was analysed by Western blot using (A) Dako anti-core polyclonal antibody (B) Life Sciences anti-core polyclonal antibody and (C) Zymed anti-core polyclonal antibody. Aliquots of purified HBcAg (c) and baculovirus-expressed deletion protein Δ84-109 were included as controls and the molecular weights of markers used are shown.

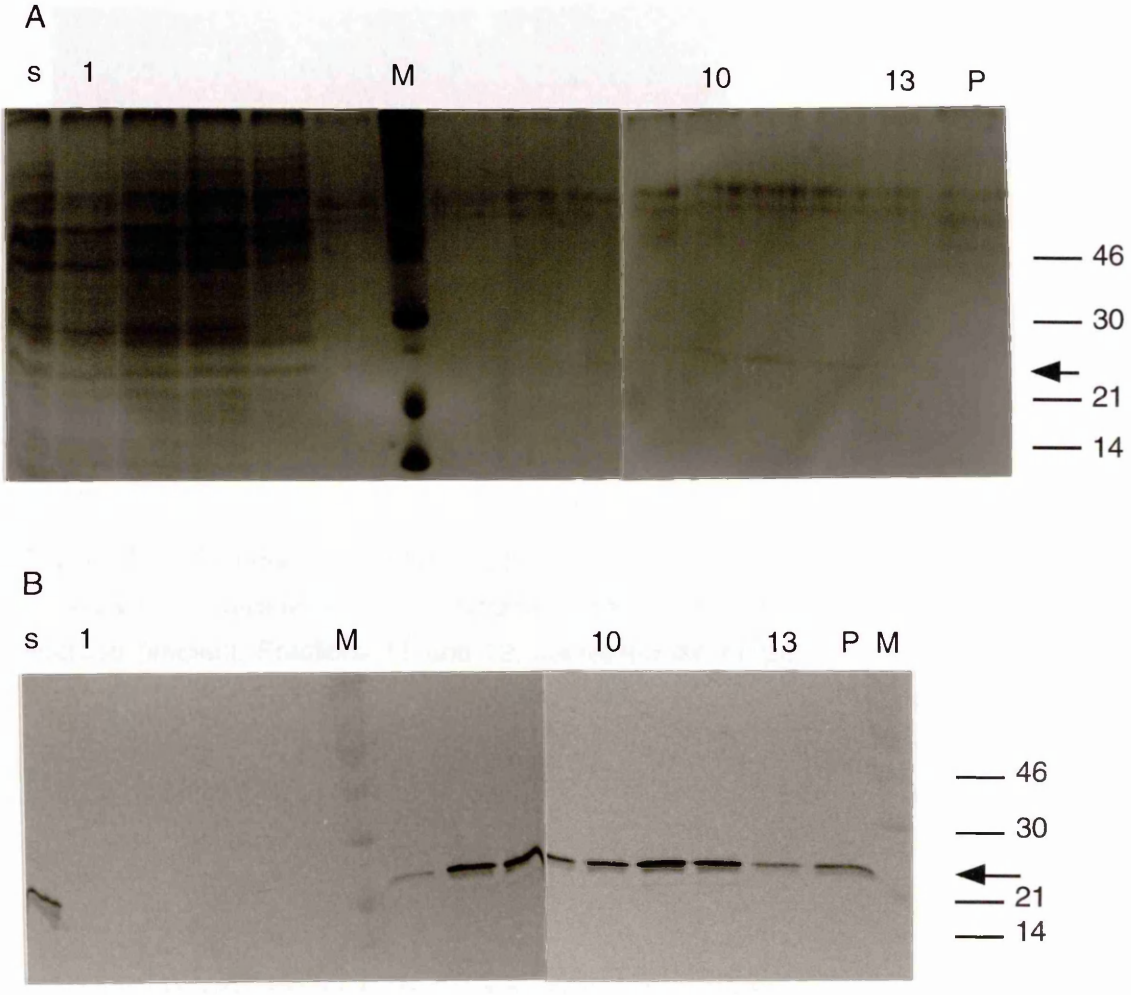


Figure 3B2c: Purification of wild-type core particles

Full-length core protein was expressed using plasmid pR1-11 and, after an ammonium sulphate precipitation step, was purified using a 10-50% sucrose gradient. Fractions were collected from the top (1) to the bottom (13) of the gradient. Any pelleted material was resuspended and collected (P) and an aliquot of soluble cellular extract (s) was also analysed. Panel A shows a Coomassie stained 12% polyacrylamide gel of the fractions and panel B shows specific detection of core protein-containing fractions on Western blots, using mAb 42B12. Molecular weights of the markers (M) used are shown and the position of the core protein is indicated by an arrow (←).

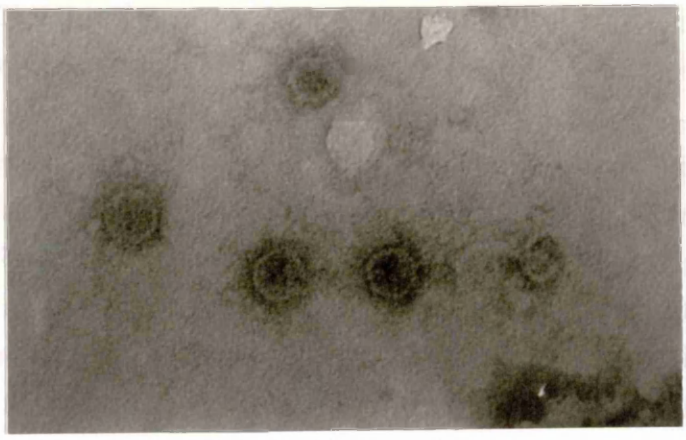


Figure 3B2d: Electron micrograph of pR1-11 core particles
Core particles were purified from extracts of pR1-11-transformed *E. coli* using a 10-50% sucrose gradient. Fractions 11 and 12, containing the highest levels of core protein as identified by Western blot analysis, were pooled and concentrated approximately five-fold using an Amicon 50 microconcentrator. An aliquot of the concentrated sample was negatively stained and examined by electron microscopy. Scale bar represents 50nm.

antibodies (Figure 3B2b, Panels B and C, respectively) also detected this protein in the extract from cells expressing full-length core protein. This protein showed greater reactivity in the extract from cells expressing core deletion protein Δ60-117 when the Dako polyclonal antibody was used (Panel A, Figure 2B2b). This protein may represent the core deletion protein, but as the concentration of acrylamide used was too low, complete resolution of these proteins may not have occurred. If these proteins are due to non-specific reactivity with the polyclonal antibodies used they could conceal detection of the core deletion proteins. Increasing the time taken for blocking the membrane gave no reduction in the number of proteins detected by these antibodies. It may have been that the detection method used on the blots was not sufficiently sensitive for detection of the core deletion proteins and a more sensitive method would have given better results.

Despite these difficulties in detecting the core deletion proteins, it was decided that for the pR1-11 plasmid containing the PCR amplified core gene, instead of the larger original HBV DNA fragment, validation of its ability to produce wild-type core particles would be worthwhile. Protein was expressed as before and clarified cellular extract was made up to 30% ammonium sulphate. After a centrifugation step, the precipitate was resuspended in TN buffer and centrifuged through a 10-50% sucrose gradient. The fractions removed from this gradient were analysed on Coomassie stained gels and by Western blot, using anti-core mAb 42B12 (see Figure 3B2c). A protein of approximately

23kDa was very faintly detectable on the Coomassie stained gel, in the lower region of the gradient (fractions 10-12). From the Western blot of the same fractions this band was conclusively identified as core protein and was found to be spread over a larger region of the gradient than was evident from the Coomassie stained gel. Core protein was present in fractions 6 to 13 and also in the pelleted material. The protein seemed to peak in fractions 11 and 12, corresponding to the core protein-containing fractions visible in the Coomassie stained gel. In order to determine whether the protein at this position in the gradient was in the form of core particles, these two peak fractions were pooled, dialysed and concentrated approximately five-fold before examination by negative staining electron microscopy (see Figure 3B2d). The particles shown here were approximately 30nm diameter, which was in agreement with the observed 28nm diameter of core particles and all the particles observed appeared to be uniform in size.

3B3 Discussion

After the problems encountered with purification of core particles using recombinant baculoviruses, expression of core protein and recovery of core particles using plasmid pR1-11 in a bacterial expression system has been successful. The purification protocol, comprising ammonium sulphate precipitation of bacterial extracts, followed by sucrose gradient centrifugation of the resuspended pellet, yielded particles of a similar diameter and morphology to core particles of HBV. The nature of the bands identified by Western blot above the region of the sucrose gradient where core particles were observed was only investigated at a later stage (see Section 3D2). Core protein was also observed in an equivalent region when particles were purified from *Xenopus* lysates (Zhou and Standring, 1991; Seifer *et al.*, 1993), but no comment was made on the possible nature of these proteins. However, a direct comparison between the results presented here and those of Zhou and Standring (1991) and Seifer *et al.* (1993) may not be relevant, as the overall gradient profiles for core protein differed from those presented here. Zhou and Standring showed a population of free core protein in the uppermost four fractions of the gradient, distinct from the core particle-containing fractions (fractions 7 to 11). The same gradient profile was shown by Seifer *et al.* (1993). Core protein was not detected in the uppermost fractions of the gradient in the results shown in this section, but was only present from fraction 6 down to the bottom of the gradient.

The problem of detection of core deletion proteins remained when they were expressed in this bacterial system. As full-length core protein was not evident as a major novel band on a Coomassie stained gel showing bacterial extract (see Figure 3B2c, panel A), it was necessary to use Western blotting to detect proteins expressed from pR1-11. Polyclonal antibodies were used in order to increase the probability of detection for these proteins on Western blots of cellular extracts, but this did not seem to be successful as numerous proteins were detected by all three antibodies. Proteins larger in size than the

predicted molecular weights of the core deletion proteins were probably detected due to a non-specific reactivity with the polyclonal antibodies used.

A more fundamental problem with the methods used in this section, observed by Dr. H. Marsden (Institute of Virology), was that the type of polyacrylamide gel used to analyse these proteins should have been altered. The core deletion proteins had predicted molecular weights in the range of 16-21kDa (see Table 2B1.11c). On 12% polyacrylamide gels, this meant that these proteins migrated almost at the dye front and were not resolved to any great extent, which made their analysis difficult. It was suggested that an increase in the acrylamide concentration would resolve this problem, therefore in the following sections all gels used were 17% polyacrylamide gels. Due to the number of additional proteins detected by all three polyclonal antibodies on Western blots, it was also decided to change to a different means of protein detection. Other methods of detection, such as immunoprecipitation of the proteins from cellular extracts, would probably suffer from the same problems as encountered previously: lack of binding by antibodies due to the size of the deletions or consequent changes in epitope conformations or non-specific binding by polyclonal antibodies to cellular proteins. Although a number of different monoclonal antibodies were available for testing which had detected the Δ84-109 core deletion protein (Section 3A4), these were not guaranteed to recognise the additional deletion proteins and, even if this was the case, any alterations we had planned to make to these proteins may have impaired this recognition. Preliminary results obtained by epitope-tagging core proteins (Section 3C) had shown that this might be a more useful means of detection for a wide range of core deletion proteins. It was also decided to use the more sensitive Amersham ECL system for detection of proteins on Western blots, to aid in detecting any proteins expressed at low levels.

3C Expression and purification of epitope-tagged core protein and core deletion proteins

3C1 Introduction

The use of small, discrete peptide sequences that are known to be epitopes recognised by specific monoclonal antibodies, has been invaluable in the analysis of a variety of biological processes. These epitope tags have been used to determine subcellular locations for proteins, such as the human ubiquitin-activating enzyme E1 (Handley-Gearhart *et al.*, 1994) and in the purification of biologically active proteins by immunoaffinity chromatography, as was achieved for human interleukin-5 receptor alpha chain (Brown *et al.*, 1995). It was decided that the use of an epitope tag attached to the core deletion proteins would greatly assist with the detection of these proteins and any modified proteins derived from them. The pp65 epitope tag was chosen as this sequence was short enough to be incorporated into an oligonucleotide primer for use in PCR amplification of these proteins and could be detected by a commercially available monoclonal antibody, mAb 9220. The ten amino acid pp65 epitope tag was derived from the HCMV UL83 gene product, the late nuclear protein pp65 (McLauchlan *et al.*, 1994). It was decided to fuse the pp65 epitope tag to the C-terminus of the core proteins by amplification of the genes using primer pair cd1/Bamtag. The resulting C-terminus of the proteins is shown in Figure 3C1. Addition of foreign sequences at this position has been previously shown to have no effect on core particle assembly (see Section 1E5). The core genes with C-terminal pp65 epitope tags were inserted into pR1-11 for subsequent expression studies.

Core protein sequence:

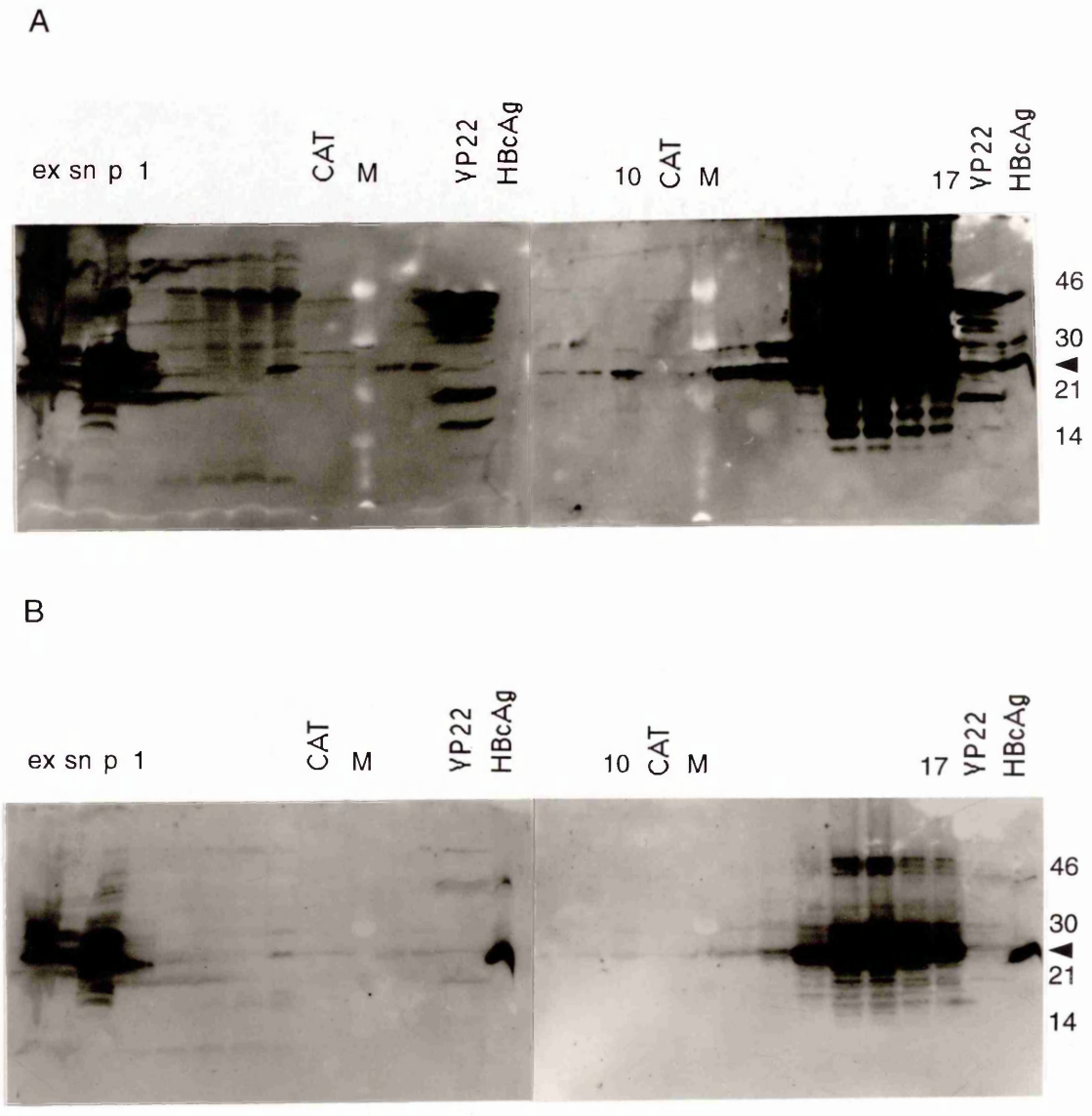
```
151  RRGRSPRRRT PSPRRRRSQS PRRRRSQSRE SQC*

151  RRGRSPRRRT PSPRRRRSQS PRRRRSQSRE SQCERKTPRV TGG*
```

Figure 3C1: The C-terminal end of core protein showing pp65 epitope tag sequence
The upper sequence shows the original core protein C-terminus and the lower sequence shows the addition of the ten amino acid pp65 tag sequence (underlined).

3C2 Expression and purification of pp65 epitope-tagged full-length core protein

As stated previously, addition of foreign sequences to the C-terminus of full-length core protein has no effect on the production of core particles. However, it was decided to purify the particles produced by the tagged full-length core protein to verify this and, more importantly, to ensure that detection was specific to the epitope tag sequences, with minimal non-specific binding. Full-length tagged core protein was expressed from pR1-11 and the resuspended ammonium sulphate pellet was centrifuged through a 10-50% sucrose gradient. Gradient fractions were analysed on Coomassie stained gels and Western blots, using both anti-core mAb 42B12 and anti-pp65 mAb 9220 (see Figure 3C2a). Controls for detection of the epitope tag were kindly supplied by Dr. J. McLauchlan: epitope-tagged CAT protein (20kDa) and tagged VP22 (40kDa), a tegument protein of HSV.

The results from the Coomassie gels of these fractions (Figure 3C2a, Panel C) showed that large amounts of protein were detected in the lower region of the gradient (fractions 14-17) and in the pelleted material. This protein corresponded well in size to that of the core protein expressed from pR1-11 with the additional tag sequences (predicted molecular weight 25kDa). The expression levels for this protein were greater than those of the similar protein expressed and purified in Section 3B2. This may be due, in some way, to the presence of the epitope tag, or may simply be due to a more efficient extraction procedure for this particular experiment. The protein-containing fractions in the lower region of the gradient were also relatively free of other contaminating proteins, with only two additional major staining bands of approximately 28 and 29kDa molecular weight present. Western blots of these fractions confirmed that the pp65 epitope tag was specifically detected by mAb 9220 (Figure 3C2a, Panel A). This detection was not due to cross-reaction with any core protein epitopes, as no reactivity was observed for the lanes containing an aliquot of untagged core particles, which were included as an additional molecular weight marker (the band observed in the final lane was thought to be due to overspill of sample from the adjacent lane after size comparison with the band present on the Panel B blot). The VP22tag control was also detected effectively by the 9220 mAb, but this was not the case for the CATtag control, the reasons for which were unclear. A major reactive tagged core protein band was detectable, with a mass of approximately 25kDa, which corresponded well with the band from the Coomassie gels. The peak fractions for this protein were 13-17, which also concurred with the Coomassie gels. As was observed for purification of untagged core particles (Section 3B2), this protein was also detectable towards the top of the gradient, from fraction 5. The nature of this protein was only investigated at a later stage (see Section 3D2). Several other higher molecular weight proteins were also detected by mAb 9220, including proteins of 28 and 29kDa. These particular proteins may be identical to those observed on the Coomassie stained



Controls:

CAT = epitope-tagged chloramphenicol acetyltransferase

VP22 = epitope-tagged HSV-1 tegument protein

HBcAg = wild-type purified core particles

Figure 3C2a: Purification of pp65 epitope-tagged core particles (continued with figure legend on following page)

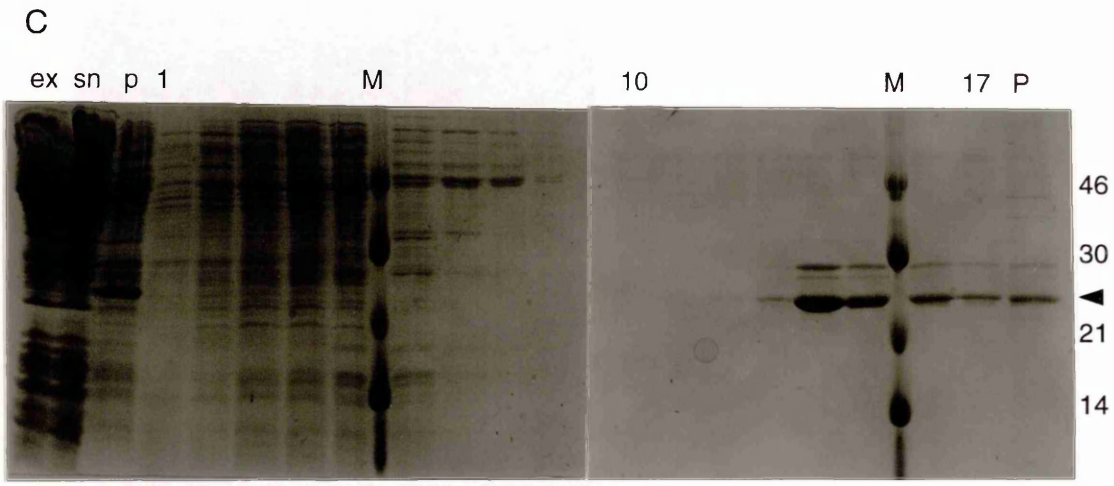


Figure 3C2a: Purification of pp65 epitope-tagged core particles

Full-length core protein with a C-terminal pp65 epitope tag was expressed using plasmid pR1-11 and purified on a 10-50% sucrose gradient, after an ammonium sulphate precipitation step. Fractions were collected from the top (1) to the bottom (17) of the gradient. Any pelleted material was resuspended (P) and analysed, along with aliquots of soluble cellular extract (ex), supernatant (sn) and pellet (p) from the ammonium sulphate precipitation step. Fractions were electrophoresed through 17.5% polyacrylamide gels which were Coomassie stained (Panel C) or Western blotted and incubated with anti-pp65 mAb 9220 (Panel A). This blot was stripped and reincubated with anti-HBc mAb 42B12 (Panel B). The position of the epitope-tagged core protein is indicated by an arrow (◄). Molecular weights of the markers used (M) are shown.

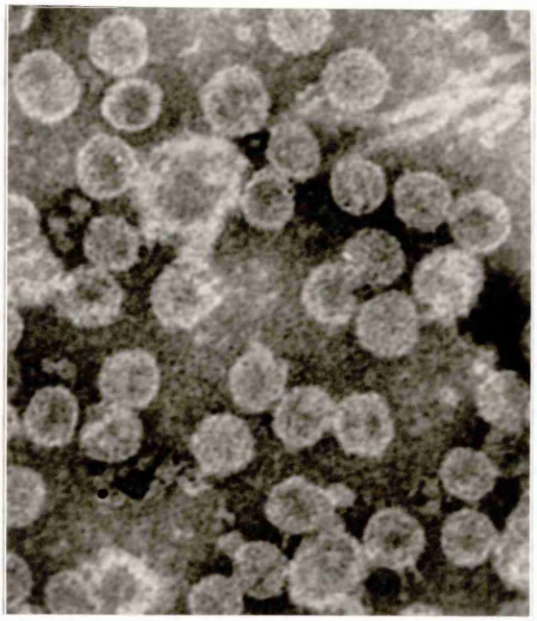


Figure 3C2b: Electron micrograph of pp65 epitope-tagged core particles
Core particles were purified from extracts of bacteria transformed with plasmid pR1-11 containing full-length core gene with a C-terminal pp65 epitope tag, using a 10-50% sucrose gradient. Fractions 14-16, containing the highest levels of core protein as identified by Western blot analysis, were pooled and concentrated using an Amicon 50 microconcentrator. An aliquot of the concentrated sample was negatively stained and examined by electron microscopy. Scale bar represents 50nm.

gels. Their reactivity with mAb 9220 suggests that, despite their larger size, they are related to the full-length core protein. These bands may, therefore, represent proteins produced by read-through of the stop codon encoded by the core gene. Proteins of approximately 44-50kDa were also detected and these may represent dimers of the core protein and other proteins with a slightly higher molecular weight, also detected by mAb 9220 (see Section 3D3). The blot incubated with anti-core mAb 42B12 (Figure 3C2a, Panel B) again showed that there was no antibody cross-reactivity with epitope tag sequences. This antibody was not as sensitive, at the dilution used, as the 9220 mAb, as protein towards the top of the gradient was less obvious. Fractions 14-16 were pooled, dialysed and concentrated as before for examination by negative staining electron microscopy (see Figure 3C2b). Large numbers of apparently normal particles were observed, confirming that the epitope tag made no obvious difference to particle assembly.

3C3 Expression and purification of pp65 epitope-tagged core deletion proteins

All four core deletion genes had been amplified by PCR to include the C-terminal epitope tag and inserted into pR1-11. Proteins were expressed as normal and samples of the ammonium sulphate precipitate were analysed by Western blot with anti-pp65 mAb 9220 (Figure 3C3a), where all four proteins were shown to be expressed and detected. Although all proteins were expressed for the same time and using equal culture volumes, differences in expression levels, as judged by the intensity of the bands on the blot, were observed. The number of other bands present on the blot was also variable for each protein. Bands of a lower molecular weight, observed for full-length core protein and core deletion Δ81-121, may have been due to degradation of the expressed protein. As was observed in a later section (see Section 3D3) a band of approximately double the size of the expressed protein represented protein dimers. however other bands were present of both smaller and larger molecular weights. Although the sample containing the core deletion Δ60-117 did not run well, subsequent gels for the purification of this protein showed that its size was the predicted 16kDa.

Purification of these expressed proteins on 10-50% sucrose gradients was undertaken, using the standard protocol (see Figure 3C3b). These results show that the four proteins differed in their movement through the gradients. Core deletion protein Δ84-109 (Figure 3C3b, Panel i) was only very faintly detectable in fractions 6-8 of the gradient, with none of this protein observed in fractions toward the bottom of the gradient in the region where core particles were detected. The lower level of expression of this protein, compared to the full-length protein and dilution of the protein in the larger volume of the gradient fractions probably both contributed to the poor levels of detection. When these fractions were pooled, concentrated approximately ten-fold and examined by electron microscopy, no particle-like structures were observed. Core deletion protein Δ81-121 (Figure 3C3b, Panel ii) was present in fractions 3-12, and also faintly in the pelleted material. It was likely that the protein present in the pellet represented high molecular weight aggregates formed during the extraction process. Again this protein is present in fractions nearer the top of the gradient than was observed for core particles. The intensity of the bands decreased in fractions 8 and 9 and then increased again in fraction 10. Fractions were pooled into two sets of fractions 5-7 and 10-12, in case the second peak of protein represented a distinct population, and concentrated for electron microscopy. For both sets of fractions, no evidence of particle-like structures was found. Core deletion protein Δ79-125 (Figure 3C3b, Panel iii) was only faintly detectable in fractions 6-9, which was, again, above the region where particles were detected. The lower levels of expression of this protein, when compared to the full-length protein, would have contributed to the difficulty in its detection. Electron microscopic examination of these concentrated fractions showed no particle-like structures. Core protein with the largest

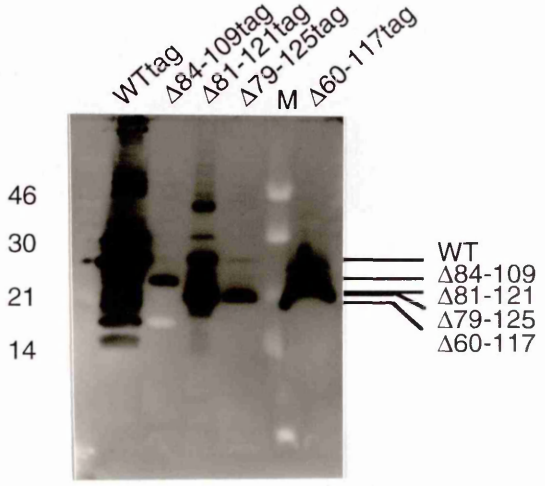
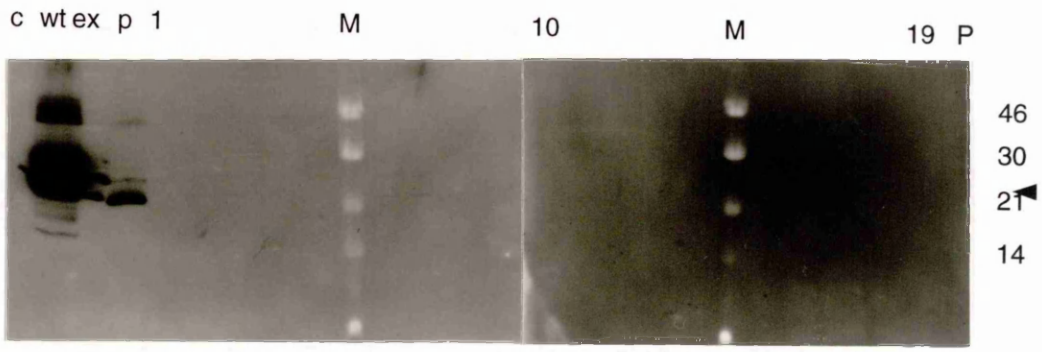


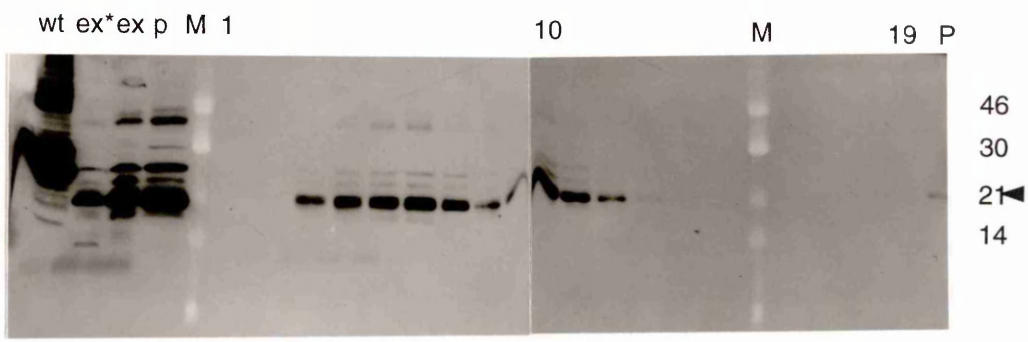
Figure 3C3a: Expression of pp65 epitope-tagged core deletion proteins

Core protein deletions with a C-terminal pp65 epitope tag were expressed using plasmid pR1-11 and precipitated with ammonium sulphate. An aliquot of the resuspended pellet was electrophoresed through a 17.5% polyacrylamide gel along with full-length tagged core protein (WTtag) and Western blotted. The blot was incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown.

i Δ84-109tag



ii Δ81-121tag



ex* - cellular extract of Δ79-125tag

iii Δ79-125tag

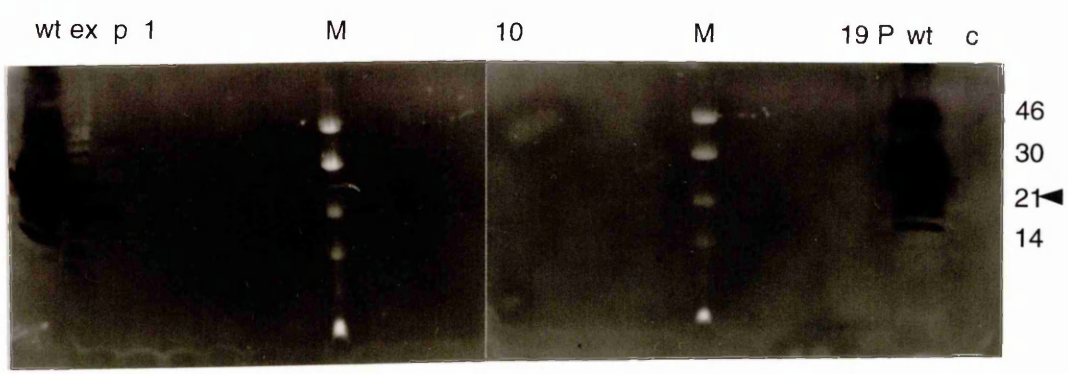


Figure 3C3b: Purification of pp65 epitope-tagged core deletion proteins (continued with figure legend on the following page)

iv Δ60-117tag

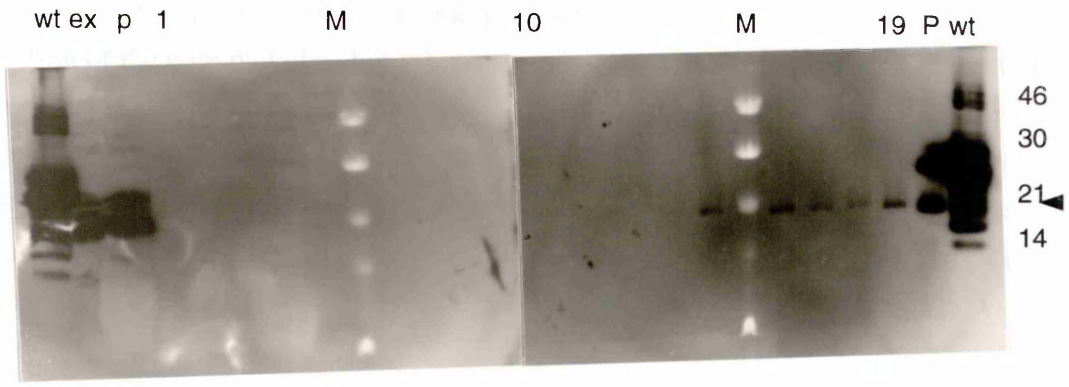


Figure 3C3b: Purification of pp65 epitope-tagged core deletion proteins (continued)
Core deletion proteins with C-terminal pp65 epitope tags were expressed using plasmid pR1-11 and purified on 10-50% sucrose gradients, after an ammonium sulphate precipitation step. Fractions were collected as before and analysed on Western blots using anti-pp65 mAb 9220. Lanes labelled 'c' in panels i and iii show extracts from untransformed bacteria. Molecular weights of the markers used are shown and the position of the protein in each blot is indicated by an arrow (◄).

deletion, Δ60-117 (Figure 3C3b, Panel iv), was present in the lower region of the gradient, in fractions 15-19 and in the pelleted material. This corresponded well to the position of the core particles, but when pooled, concentrated fractions were examined by electron microscopy, despite thorough efforts, no particle-like structures were observed.

These results have shown that all four core deletion proteins were incapable of assembling into a core particle-like structure that was visible by electron microscopic examination. It may have been that any particulate structures formed were unstable, explaining their absence in samples for examination. However, precautionary measures were taken to prevent this: samples were kept on ice at all times and concentrated and examined with minimum intervening periods and samples were not frozen at any stage after purification. The rationale for this last precaution came from studies with C-terminally truncated core protein, lacking the final 32 residues (Seifer and Standring, 1993). Mutant particles formed from this protein were originally thought to display 'particulate' HBe epitopes, but the results from this paper showed that this was probably due to the instability of these particles. These results also showed that one or two freeze/thaw cycles increased the level of HBe antigenicity, again by disrupting the particles. Sonication of bacterial cell suspensions could also have disrupted any core particle-like structures that had assembled. The possible effects of these procedures were analysed in a later section (see Section 3G).

3C4 Expression and purification of pp65 epitope-tagged Δcore-surface hybrid protein

From the results in the previous section it was concluded that core proteins with central deletions were unable to assemble into core particle-like structures. This raised the question of whether this was due to deletion of specific amino acid sequences required to direct folding of the protein into the correct conformation, and thus correct assembly of the particles, or purely due to the size of the deletion being too large, with the actual sequence deleted being irrelevant. In order to address this question it was decided to produce a modified core deletion protein, where the deleted region had been replaced by foreign sequence of exactly the same size. Core deletion Δ81-121 was chosen and the deleted region was replaced by a 123bp fragment of the HBV surface gene, using a PCR-based method (see Figure 2B1.11a), with the product being a Δcore-surface hybrid gene. The amino acid sequence of this protein is shown in Figure 3C4a. However, when this was expressed and purified on a sucrose gradient, it became apparent that there were problems with this experiment (Figure 3C4b, Panel A). Although protein of the correct size was produced it was expressed at very low levels and was only detectable after concentration of proteins by ammonium sulphate precipitation. Upon Western blot analysis of fractions from the sucrose gradient purification of this hybrid protein, no

```
  1  MDIDPYKEFG ATVELLSFLP SDFFPSVRDL LDTAAALYRD ALESPEHCSP
                                       81
                                        SRELVVSYVN VNMGLKIRQL
 51  HHTALRQAIL CWGDLMTLAT WVGDNLQDPA GMLPVCPLIP GSSTTSTGPC
                           121
     LWFHISCLTF GRETVLEYLV S
101  RTCTTPAQGI SMYPSCCCTK PFGVWIRTPP AYRPPNAPIL STLPETTVVR

151  RRGRSPRRRT PSPRRRRSQS PRRRRSQSRE SQC*
```

Figure 3C4a: Amino acid sequence of the Δcore-surface hybrid protein
The sequence of the region of surface protein inserted into the Δ81-121 deletion is underlined. Above it is shown the sequence present in full-length wild-type core protein.

protein could be detected at any position on the gradient. The failure to detect protein in the gradient fractions may have been due to the low level of protein present, which may have been below the sensitivity of the monoclonal antibody used. Another possibility, which seemed more likely, was that during this particular purification procedure the protein had become insoluble. Examination of protein which had precipitated during the overnight dialysis step showed this to be the case, as the Δcore-surface hybrid protein was present in this material. In order to attempt to overcome this problem, the purification procedure was modified in two ways. The first modification was to attempt to remove nucleic acids from the bacterial extract so they would not interfere with subsequent steps. This was achieved by the use of PEI, a positively charged polymer. PEI precipitation of nucleic acids is commonly used as the first step of a protein purification protocol. The PEI interacts with nucleic acids to form an insoluble complex which can then be removed by centrifugation. The bacterial extract was first made up to 0.5M NaCl before precipitation of nucleic acid by the addition of PEI. After removal of the PEI:nucleic acid complex, ammonium sulphate precipitation of proteins in the supernatant was carried out as before. The second modification was to load the resuspended pellet from the ammonium sulphate precipitation step onto the sucrose gradient, without the intervening dialysis step. When extracts from all stages of this protocol and fractions from the sucrose gradient were analysed, the protein was only detected in the crude bacterial extract (the apparent band in lane 10 was caused by forceps, used to handle membranes). The protein appeared to have been removed from solution during the nucleic acid precipitation step, possibly due to the increased salt concentration required (0.5M, compared to 0.1M in TN buffer).

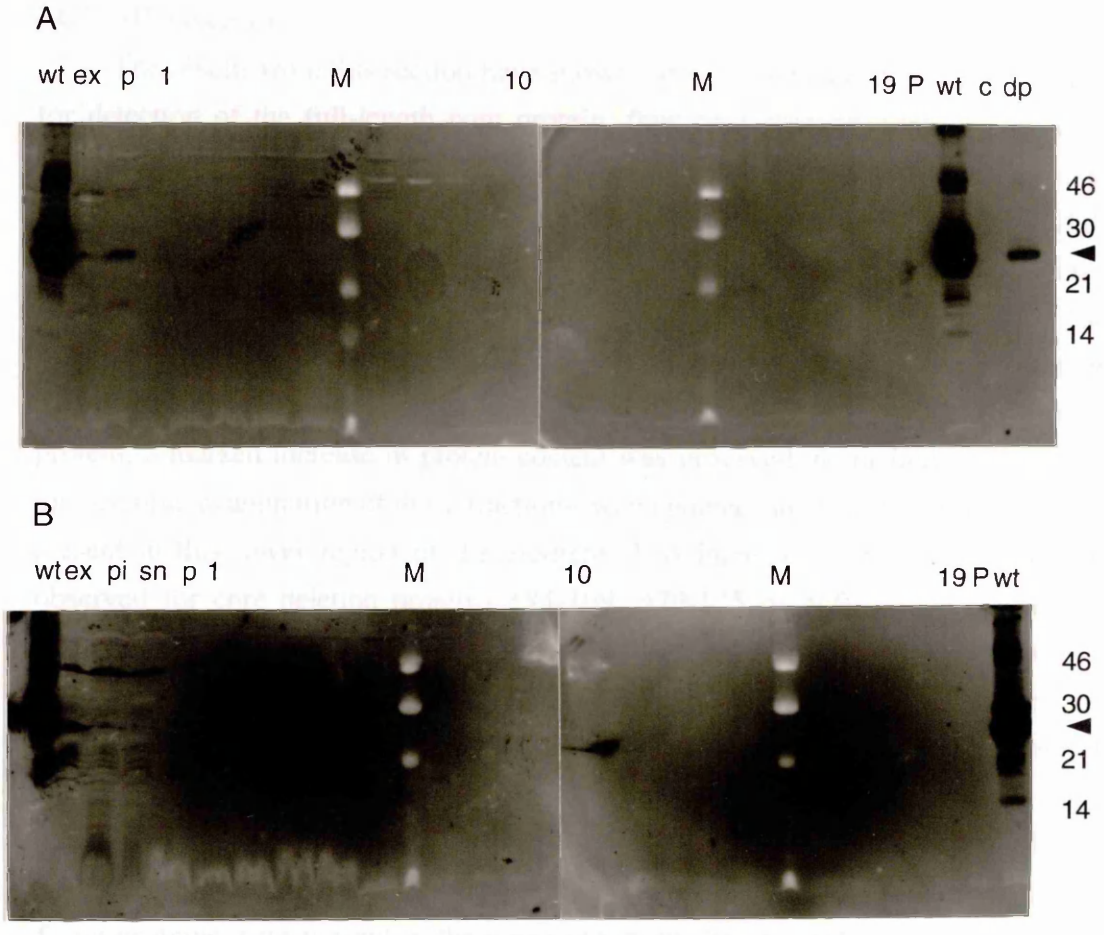


Figure 3C4b: Purification of epitope-tagged Δcore-surface hybrid protein

The Δcore-surface hybrid protein with a C-terminal pp65 epitope tag was expressed using plasmid pR1-11 and purified on a 10-50% sucrose gradient, after an ammonium sulphate precipitation step (Panel A) or with an initial additional precipitation of nucleic acid using PEI (Panel B). Fractions were collected from the top (1) to the bottom (19) of the gradient and analysed by Western blotting using anti-pp65 mAb 9220. Any pelleted material was resuspended (P) and analysed, along with aliquots of soluble cellular extract (ex) and pellet (p) from the ammonium sulphate precipitation step. Panel A shows the original purification method, where the lane labelled 'dp' contains material which precipitated during dialysis of the ammonium sulphate pellet and the lane labelled 'c' contains cellular extract from untransformed *E.coli*. Panel B shows a modification of the original method, as described above, where the lane labelled 'pi' contains an aliquot of the supernatant after the PEI precipitation and 'sn' contains an aliquot of the supernatant after the ammonium sulphate precipitation step. The position of the epitope-tagged protein is indicated by an arrow (◄). Molecular weights of the markers used (M) are shown. An aliquot of epitope-tagged full-length core protein (wt) was included as a control.

3C5 Discussion

The results from this section have shown the successful use of the pp65 epitope tag for detection of the full-length core protein, four core deletion proteins and a Δcore-surface hybrid protein. Full-length core protein formed large numbers of core particles when expressed showing that the presence of the epitope tag did not alter its ability to assemble into morphologically normal core particles. The core deletion proteins were expressed and purified, however none of them were able to assemble into any form of core particle-like structure. This was determined by a number of criteria, the simplest of which was the gradient profiles of these proteins. From the results with the full-length protein, a marked increase in protein content was observed in fractions 13-17. Electron microscopic examination of these fractions, when pooled, showed that core particles were present in this lower region of the gradient. This increase in protein content was not observed for core deletion proteins Δ84-109, Δ79-125 or Δ60-117, where all positive fractions on Western blots had approximately the same, albeit very faint, intensity. Core deletion protein Δ81-121 did show two distinct peaks of protein-containing fractions (3-9 and 10-12). However, no particles were detected in either region when they were examine separately by electron microscopy.

Examination of the relative positions of the core deletion proteins Δ84-109 and Δ79-125 in the gradient also led to the conclusion that they were unable to form particles. These proteins were present in the upper region of the gradient (fractions 6-8 and 6-9, respectively), above the fractions where core particles were detected. However, Western blots of fractions from the purification of full-length core protein (Figure 3C2a) showed that epitope-tagged core protein could also be detected in this upper region of the gradient, from fraction 5 down. Although this raised the possibility that protein in this region was in the form of particles, a more thorough investigation of core protein-containing fractions in this region by electron microscopy (see Section 3D2) failed to show evidence for this. In addition, purification of a C-terminally truncated core protein which was known to be unable to assemble into core particles showed that this protein was present in this upper gradient region (see Figure 3D2a). In contrast to the Δ84-109 and Δ79-125 proteins, core deletion proteins Δ81-121 and Δ60-117 were present in the lower region of the gradients (fractions 10-12 and 15-19, respectively). However, electron microscopic examination could not detect any particles. It may have been that any particulate structures formed were unstable, despite the precautions taken, and were unable to be detected. Aggregation of the particles could also have taken place, either during the purification or concentration stages, so hampering their detection by electron microscopy, despite thorough efforts. Despite these possibilities, it appeared that all four core proteins with central deletions were unable to assemble into core particles.

The construction of the Δcore-surface hybrid protein was designed with a view to determining the importance of the deleted core protein sequence in directing particle

assembly. However successful expression and purification of this protein was not achieved. The HBV surface protein region used to construct this protein may have affected its stability. Other modifications to the purification protocol for this protein could have been attempted, such as using DNase and RNase to remove nucleic acids instead of PEI or the use of larger volume bacterial cultures to provide increased amounts of protein. However, upon advice from other group leaders, it was decided to once more alter the expression system used. It was thought that achieving higher levels of protein expression would allow the ammonium sulphate precipitation step to be omitted, so removing its associated insolubility problems.

3D Expression and purification of epitope-tagged core proteins in the pET system

3D1 Introduction

The pET vectors were initially constructed by Studier and colleagues as pBR322 derivatives containing a bacteriophage T7 promoter (Studier *et al.*, 1990). Genes are cloned under the control of the strong T7 transcription and translation signals before the recombinant plasmid is transferred to the *E.coli* strain BL21(DE3)pLysS. Upon addition of IPTG to this strain, T7 RNA polymerase is induced and transcribes the cloned gene in the recombinant pET vector. These host cells also contain a plasmid encoding T7 lysozyme, which cleaves a specific bond in the peptidoglycan layer of the bacterial cell wall, and allows rapid and efficient lysis of cells by either freeze/thaw or mild detergent treatment. It was thought that the higher levels of expression recorded using this system would eliminate the requirement for an ammonium sulphate precipitation step, and the cell lysis method would allow higher levels of protein recovery. All core gene derivatives, along with full-length core gene, were amplified by PCR using primer pair cp3/Bamtag (see Section 2A3), which include *Nde* I and *Bam* HI sites respectively. In contrast to plasmid pR1-11, the 5' terminus of the products of these reactions contained the authentic core protein start codon, with the 3' terminus including the pp65 epitope tag.

3D2 pET expression and purification of epitope-tagged core deletion proteins

Initial expression experiments showed that the IPTG concentration used to induce expression had very little effect on the levels of soluble protein recovered, therefore cultures were induced with a final concentration of 300μM. Production of core deletion proteins Δ81-121, Δ79-125 and Δ60-117 was found to be more efficient when induced expression was carried out at 26°C, instead of 30°C or 37°C. Timecourse experiments showed that expression periods of 3hr for Δ81-121 and Δ79-125 or 2.5hr for Δ60-117 gave greater protein recovery. All other proteins were also induced at 26°C, after the temperature of induction was shown to have no effect on protein production, and expressed for a 4 hour period. The four tagged core deletion proteins, along with full-length core protein were shown to be expressed efficiently and detectable in Western blots of clarified cellular extracts (Figure 3D2a), although Δ79-125 consistently gave lower levels of expression. Additional bands were detected on these blots, the major one of which was approximately double the predicted size of the expressed protein. This was further investigated at a later stage (see Section 3D3).

After expression and overnight storage at -20°C, the cell pellet was resuspended in TNE II, to which was added DNase and RNase to remove the nucleic acids present. The

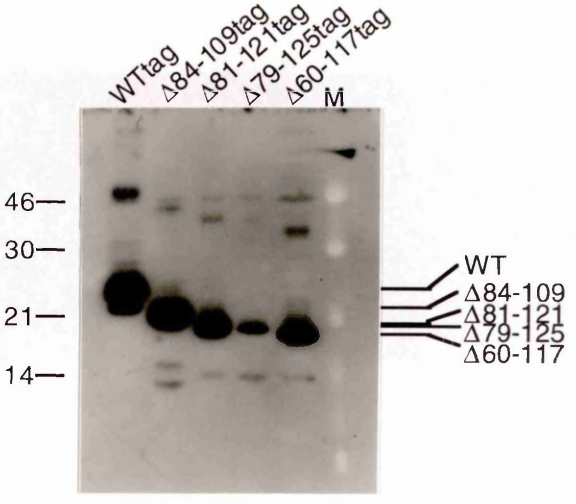


Figure 3D2a: Expression of pp65 epitope-tagged core deletions using the pET system
Core deletions with a C-terminal pp65 epitope tag were expressed in *E.coli* BL21(DE3)pLysS for 4 hours at 26°C, after induction with 300μM IPTG. Cellular extracts were prepared and aliquots electrophoresed through a 17.5% polyacrylamide gel alongside full-length tagged core protein (WTtag). The gel was Western blotted and incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown.

cellular extract was then clarified by centrifugation and soluble protein centrifuged through 10-50% sucrose gradients, after which, fractions were analysed by Western blot (see Figures 3D2b and 3D2d).

The full-length tagged protein was detectable from fraction 7 down to the bottom of the gradient, again with a marked increase in protein content occurring in fractions in the lower region of the gradient (fractions 14-19). When the peak fractions (14-17) were pooled, concentrated and examined by electron microscopy, large numbers of morphologically normal, 28nm diameter core particles were observed (Figure 3D2c). Core protein had been detected in gradient fractions above this region using both the pR1-11 vector (see Figures 3B2c and 3C2a) and in this section, with the pET3a vector. It was therefore decided to examine the fractions in this region to determine the state of this population of core protein. Fractions 9-12 were pooled, concentrated and examined by electron microscopy. However, no core particles could be detected.

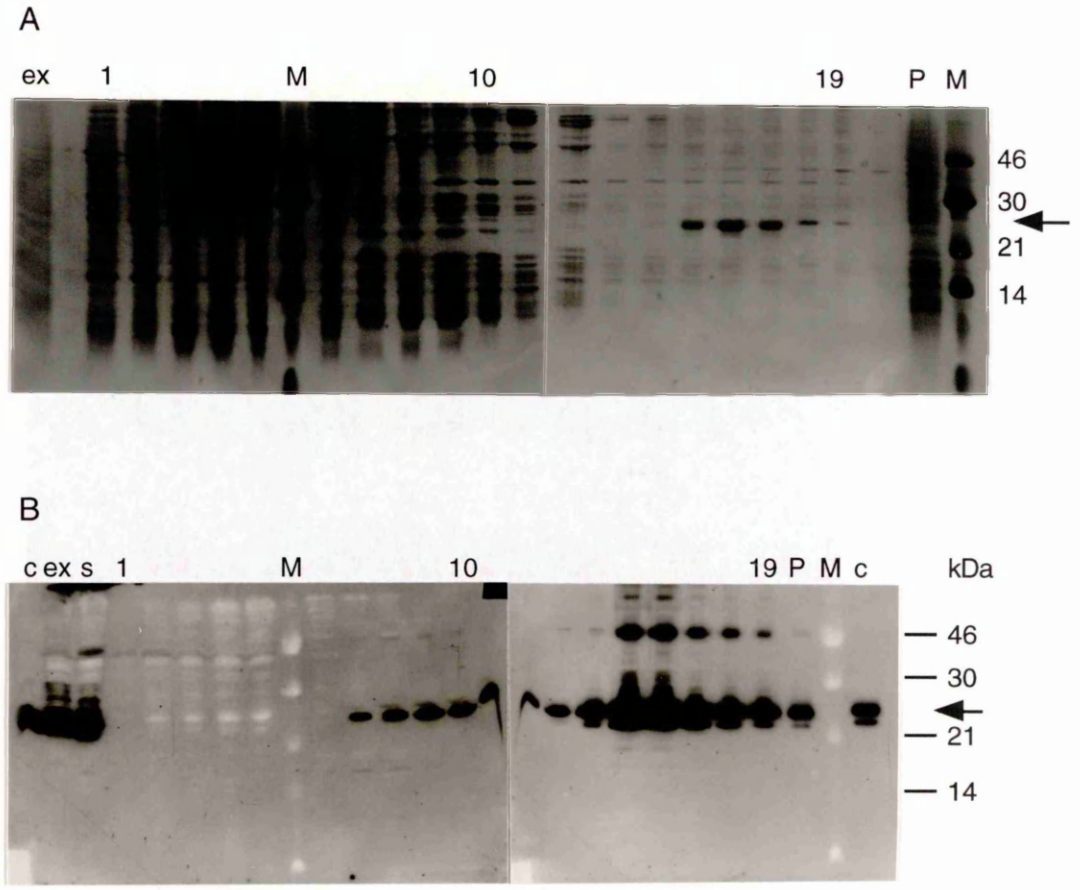


Figure 3D2b: Purification of pp65 epitope-tagged core particles expressed in the pET system

Full length core gene with a C-terminal pp65 epitope tag was expressed in *E.coli* BL21(DE3)pLysS for 4 hours at 26°C, after induction with 300µM IPTG. Total cellular extract (ex) was prepared and after centrifugation at 35k rpm for 20 min, soluble protein (s) was centrifuged through a 10-50% sucrose gradient. Fractions were removed from the top (1) to the bottom (19) of the gradient and any pelleted material was resuspended (P). Fractions were electrophoresed through 17.5% polyacrylamide gels along with an aliquot of tagged full-length core protein (c) expressed using plasmid pR1-11 (Section 3C2). These gels were then Coomassie stained (Panel A) or Western blotted with anti-pp65 mAb 9220 (Panel B). Molecular weights of the markers used (M) are shown and the position of the tagged core protein is indicated by an arrow (←).

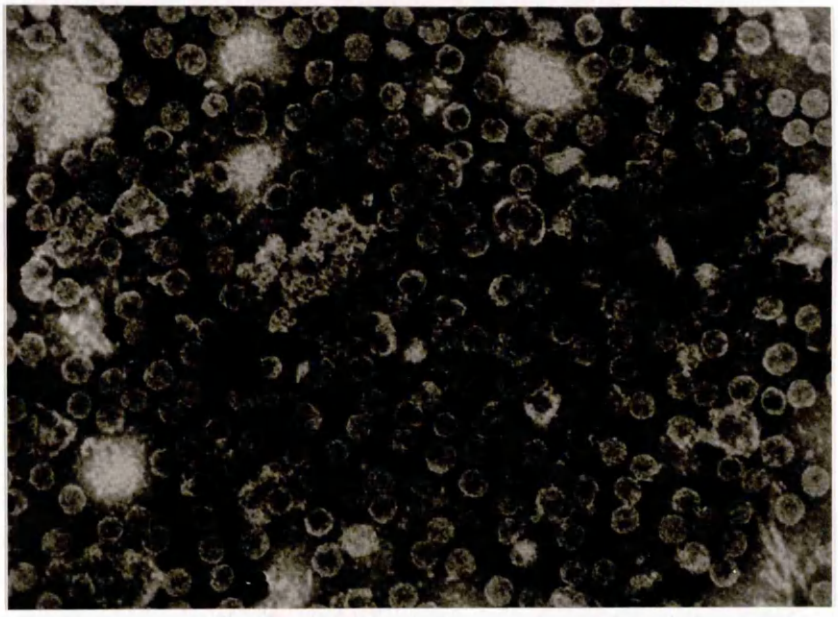
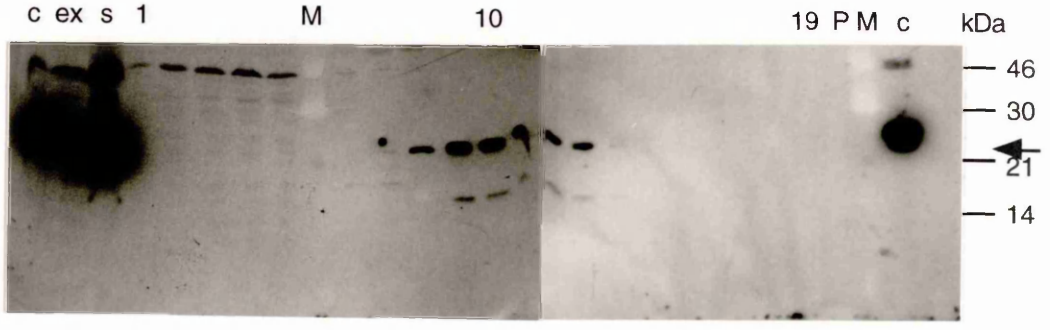


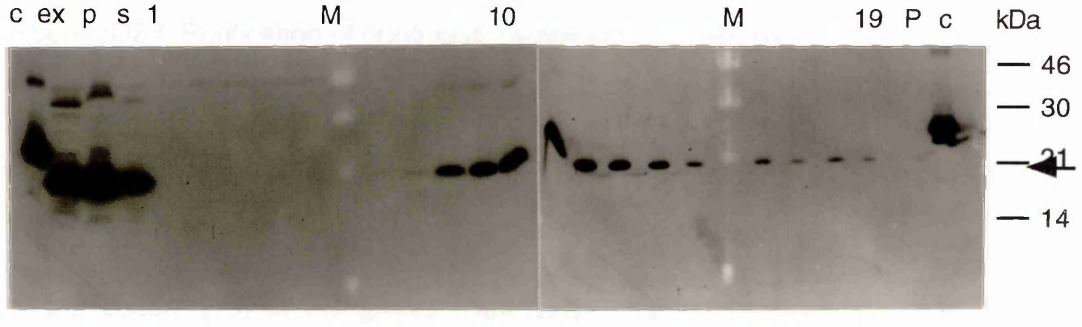
Figure 3D2c: Electron micrograph of pp65 epitope-tagged core particles expressed using the pET system

Core particles were purified from extracts of bacteria transformed with a pET3a plasmid containing full-length core gene with a C-terminal pp65 epitope tag, using a 10-50% sucrose gradient. Fractions 14-17, containing the highest levels of core protein as identified by Western blot analysis, were pooled and concentrated using an Amicon 50 microconcentrator. An aliquot of the concentrated sample was negatively stained and examined by electron microscopy. Scale bar represents 50nm.

i Δ84-109



ii Δ81-121



p = aliquot of resuspended pellet produced by 35k clarification of cellular extract

iii Δ79-125

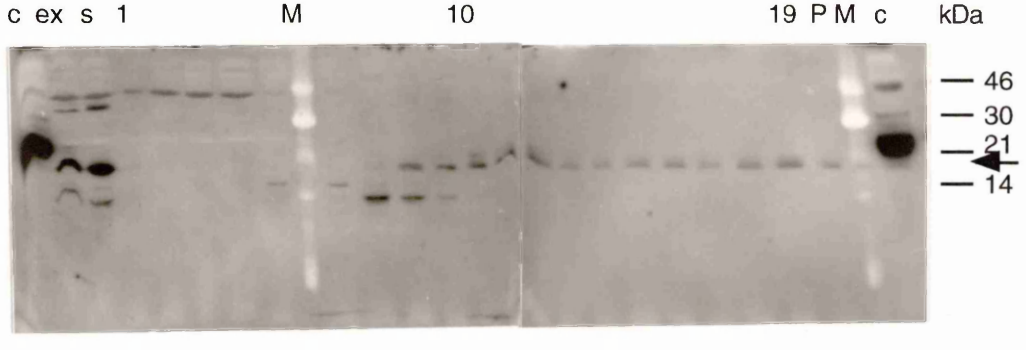


Figure 3D2d: Purification of pp65 epitope-tagged core deletion proteins expressed in the pET system (continued with figure legend on the following page)

iv Δ60-117

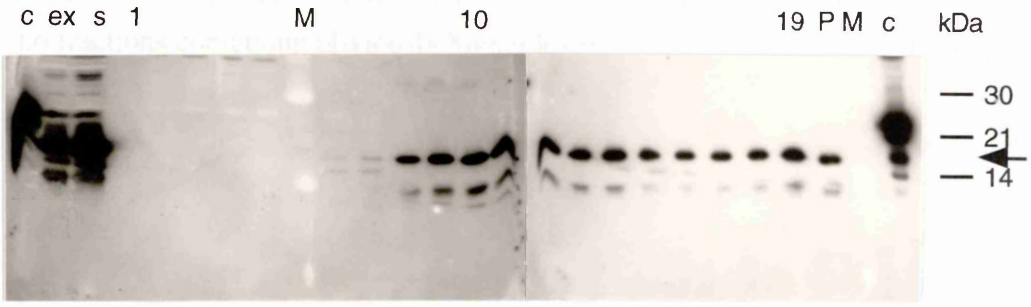


Figure 3D2d: Purification of pp65 epitope-tagged core deletion proteins expressed in the pET system

Core deletion proteins with a C-terminal epitope tag were expressed in *E.coli* BL21(DE3)pLysS at 26°C after induction with 300μM IPTG. Cellular extracts were prepared (ex) and after centrifugation at 35k rpm for 20 min, soluble protein (s) was centrifuged through a 10-50% sucrose gradient. Fractions were removed from the top (1) to the bottom (19) of the gradient and any pelleted material was resuspended (P). Fractions were electrophoresed through 17.5% polyacrylamide gels along with an aliquot of tagged full-length core protein (c). After Western blotting of the gels, membranes were incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown and the position of each core deletion protein is indicated by an arrow (←).

Core deletion protein Δ84-109 was detected only in the central region of the gradient (fractions 8-13), with all fractions having approximately equal intensities (see Figure 3D2d, Panel i). Concentration of these pooled fractions, approximately 20-fold, and examination by electron microscopy failed to detect core particle-like structures. A faintly reacting lower molecular weight protein (approximately 17kDa) was present in fractions 9-13. This may represent a protease degradation product of the core deletion protein, as it was present in the same fractions. An N-terminal region of the protein must have been removed in order to still allow detection of the C-terminal epitope tag by mAb 9220. Purification of Δ81-121 showed that protein was again present in fractions 8-13, in the central region of the gradient (see Figure 3D2d, Panel ii). However, in addition this protein was also faintly detectable in decreasing intensity towards the bottom of the gradient (fractions 14-19). When the peak fractions (fractions 8-13) were examined by electron microscopy after 20-fold concentration, no core particle-like structures were

observed. Core deletion protein Δ79-125 was detected to the same low intensity from fraction 8 down to the bottom of the gradient (see Figure 3D2d, Panel iii). As there were no fractions containing obviously higher levels of protein, fractions 9-11 and 12-14 were chosen for separate further examination as core particle-like structures formed from this protein would be expected to present in this region of the gradient, due to their smaller molecular mass. Upon examination, no such structures were observed. Smaller proteins of approximately 14kDa and 17kDa were also detectable in fractions 4-5 and 6-8, respectively. As no Δ79-125 deletion protein was present in fractions 4-6, the smaller proteins are unlikely to be core protein-related. Core deletion Δ60-117 was detected, again, in fraction 8 down to the bottom of the gradient, but at higher levels than Δ79-125 (see Figure 3D2d, Panel iv). A smaller protein of approximately 13kDa was also present in these fractions. This may have been due to proteolytic degradation of the N-terminus of the core deletion protein. The same fractions were chosen for electron microscopic examination as for the previous protein, with the same results.

The conclusion drawn from these expression studies was that all four core deletion proteins have again been shown to lack the ability to assemble spontaneously into any core particle-like structures. Although the Δ81-121, Δ79-125 and Δ60-117 deletion proteins were detected in the lower regions of the sucrose gradients (corresponding to fractions where core particles were detected), a marked increase in protein content was not observed in any of these fractions, contrasting with the result for the full-length protein.

Due to the continuing difficulties in determining the state of the protein present in gradient fractions above those where core particles were present and the state of the non-particulate protein from the core deletion purifications, it was decided that an additional control was necessary. It was shown by Birnbaum and Nassal (1990) that if core protein was C-terminally truncated beyond the leucine residue at position 140, this protein lost the ability to self-assemble into core particles. It was decided to reproduce this truncated core protein and purify it on the sucrose gradient. This result should define fractions where core particles would definitely not be present, thus excluding them from further analyses. It was decided to create the C-terminally truncated core protein by a single round of PCR using primer pair cp3/Thrtag. When expressed this would produce core protein truncated at Thr-128, with the additional ten amino acid epitope tag sequence. The extra truncation was intended to compensate for the ten amino acid epitope tag sequence, as it was not known whether this could function as an equivalent to missing core protein sequence and so, in fact, allow core particle assembly. The results for the purification of this core T128 protein are shown in Figure 3D2e. Protein of approximately the predicted size (16kDa) was produced and, as expected, was detected only in the upper region of the gradient, mainly in fractions 1-4. However, detectable amounts of protein were also present, down to fraction 10. Protein in these fractions must also have been in a non-particulate form.

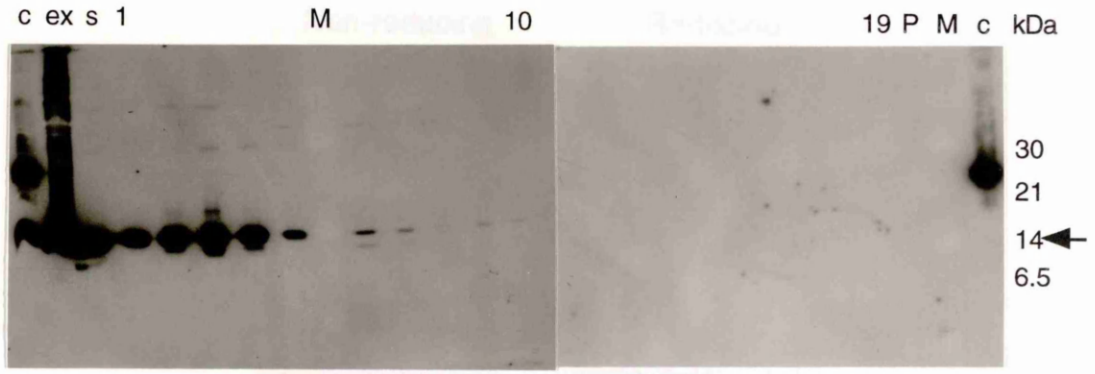


Figure 3D2e: Purification of C-terminally truncated core protein, T128
CoreT128 protein with a C-terminal pp65 epitope tag was expressed in *E.coli* BL21(DE3)pLysS as before. A cellular extract was prepared (ex) and after centrifugation at 35k rpm for 20 min, soluble protein (s) was centrifuged through a 10-50% sucrose gradient. Fractions were removed from the top (1) to the bottom (19) of the gradient and any pelleted material was resuspended (P). Fractions were electrophoresed through a 17.5% polyacrylamide gel along with an aliquot of tagged full-length core protein (c). After Western blotting of the gel, the blot was incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown and the position of the tagged coreT128 protein is indicated by an arrow (←).

3D3 Dimerisation of core deletion proteins

Studies of core particle assembly in *Xenopus* oocytes showed that dimerisation of core protein monomers, via disulphide bonds, was the initial step in the core particle assembly pathway (Zhou and Standring, 1992). As the core deletion proteins in this study did not appear to assemble into core particle-like structures, it was decided to determine at which stage the assembly pathway arrested: were the core proteins able to dimerise, but these dimers unable to assemble into particles, or did dimerisation not occur? In order to address this question, core deletion proteins were analysed on Western blots under non-reducing conditions for evidence of bands of approximately double the size of the monomeric protein (see Figure 3D3).

The results showed that, under non-reducing conditions, proteins of approximately the correct size for dimers of all four core deletion proteins were specifically detected by mAb 9220. Full-length core protein was included as a control for this experiment and under non-reducing conditions a major-staining band of approximately 48kDa was

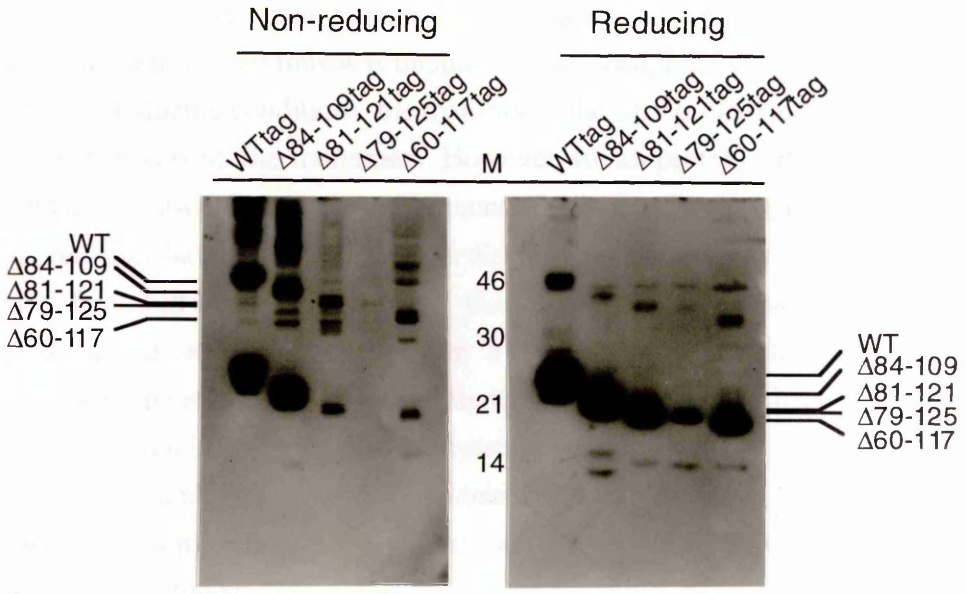


Figure 3D3: Dimerisation of core deletion proteins

Core deletion proteins with a C-terminal pp65 epitope tag were expressed using *E.coli* BL21(DE3)pLysS as before. Cellular extracts were prepared and mixed with non-reducing or reducing SDS PAGE sample buffer. Samples were heat-denatured at 100°C for 5 min, before being electrophoresed through a 17.5% polyacrylamide gel. After Western blotting of the gel, the blot was incubated with anti-pp65 mAb 9220. Tagged full-length core protein was also run (WTtag) as a control and the molecular weights of the markers used (M) are shown.

visible. This corresponded well to the predicted size of core protein dimer. This band decreased in intensity upon reduction, with a concomitant increase in the amounts of core protein monomer detected. In addition to the protein dimer, prominent lower molecular weight bands were observed under non-reducing conditions in all bacterial extracts expressing the core deletion proteins, with the exception of Δ79-125. These lower molecular weight proteins corresponded to the sizes of the respective protein monomers. It was not known if all the disulphide-linked dimers had remained intact, thus explaining the presence of the protein monomers. This may have been affected by the method of sample preparation, prior to electrophoresis: Gallina *et al.* (1989) heated samples for non-reducing electrophoresis at 60°C for 30min, whereas Zhou and Standring (1992) used 60°C for 5min, followed by 100°C for 5min. However, no reasons were given for these modifications and it seems unlikely that these differences in temperature could influence the stability of covalent bonds. The monomeric proteins may represent core proteins or core deletion proteins which had been denatured during the extraction process. This

denatured protein may have been present in a conformation where all four cysteine residues were buried and thus was unable to form disulphide bonds.

Under reducing conditions, the majority of the core protein and core deletion protein dimers were resolved into monomers. However, for all proteins, dimeric forms were still detectable. The use of an increased amount of β-ME in the sample buffer used for SDS PAGE gels may have eliminated protein dimers completely. Kann and Gerlich (1994) also detected faint full-length core protein dimers on Coomassie-stained gels when core particles treated with 5% β-ME were examined. Also visible under non-reducing conditions, in the results presented in this section, were bands of higher molecular weight than the core protein or core deletion protein dimers. These were particularly evident for the full-length and Δ84-109 core proteins and it may have been that these bands corresponded to higher order oligomers of core protein, formed by disulphide bonds between protein dimers.

3D4 Determination of the sensitivity of anti-pp65 mAb

Studies of the formation of core particles in the *Xenopus* system have shown that this process is highly co-operative and also dependent upon the concentration of accumulated core protein dimers (Seifer *et al.*, 1993). Core protein was required to accumulate to an estimated concentration of 0.7-0.8μM (14-16μg/ml) before particle assembly would initiate. As results from the purification of the core deletion Δ79-125 showed only small amounts of protein detectable in the gradient fractions, it was decided to quantify the detection sensitivity of the anti-pp65 mAb 9220, to ensure that the smaller quantity produced for this protein was not a limiting factor in its assembly into particles.
The concentration of protein in a sample of purified epitope-tagged wild-type core particles was determined by Bradford assay to be 1μg/μl. This sample was serially diluted two-fold and aliquots separated by SDS PAGE and Western blotted, using the standard 1:1000 dilution of mAb 9220 (see Figure 3D4). Under these conditions the antibody was capable of detecting a minimum concentration of 3.9μg/ml of tagged protein. Although this was lower than the minimal value required for particle assembly, the protein from the purification of Δ79-125 was distributed over twelve 0.65ml fractions. Therefore the original concentration of protein in the bacterial extract was greater than that required.

3D5 Discussion

The results from this section have shown that all four core deletion proteins studied were unable to self-assemble into stable core particle-like structures. The same probabilities existed for these results as those discussed in Section 3C5.

Core deletion protein Δ84-109 was detected only in the central region of the gradient, above the fractions where core particles were observed when full-length protein was purified. However, it was possible that this deletion protein was capable of

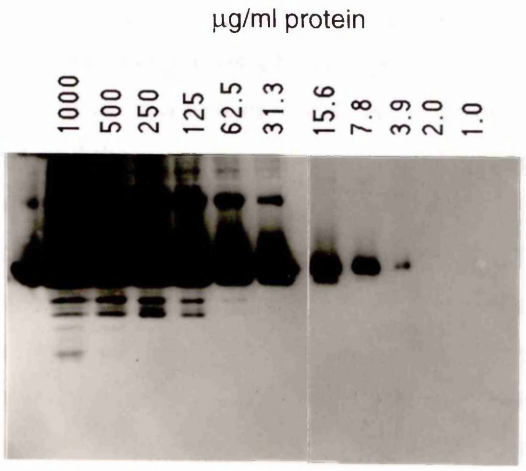


Figure 3D4: Detection sensitivity of mAb 9220 on a Western blot
Decreasing concentrations of pp65 epitope tagged core particles were electrophoresed on a 17.5% polyacrylamide gel, which was then Western blotted. After blocking, the blot was incubated with a 1:1000 dilution of anti-pp65 mAb 9220.

assembling into particles but that the particles were extremely unstable and rapidly dissociated into dimers, before the purification step. A similar possibility existed for core deletion proteins Δ81-121, Δ79-125 and Δ60-117, which were detected in the lower region of the gradient, where wild-type core particles were observed. Core particle-like structures could have been formed from these proteins, but were unstable and therefore undetectable during electron microscopic examination. This possibility was catered for to some extent by rapid and careful examination of samples after purification, with no freezing steps in the intervening periods, which may have decreased the stability of particle-like structures. However, the gradient profiles for these deletion proteins also differed from that of the full-length protein. For all four deletion proteins, none of the fractions in the lower region showed as substantial an increase in protein content as was observed for the full-length protein. This showed that 'banding' of particles at a specific sucrose density was not occurring and that no distinct particulate population of protein was present. Instead, only a protein population with densities which increased in a gradual manner was observed.

A broad gradient profile, similar to that obtained from purification of the core deletion proteins Δ81-121, Δ79-125 and Δ60-117, was also observed by another group who were investigating the properties of core protein insertion mutants (Beames and Lanford, 1995). Insertions of two to four amino acids were made throughout the first 140 residues of the protein and these proteins were expressed in insect cells using recombinant baculoviruses. Expressed proteins were then purified on 5ml 10-50% sucrose gradients to assay for core particle formation. Under the conditions used, full-length core protein (in the form of core particles) was detected in two 500μl fractions in the centre of the gradient. Full-length proteins with two or three amino acid insertions after Ala-36 or Ser-44, respectively, were detected in fractions from the top of the gradient down to the central fractions. Insertion of three, two or four amino acids after residues Thr-67, Ala-69 and Val-89, respectively, produced proteins which were detected in fractions below those to which full-length core protein migrated. The proteins with insertions after Thr-67 and Ala-69 were unable to assemble into core particles. These results were verified by expression of these two core protein insertion mutants in Huh7 cells transfected with core-defective HBV genomes. This experiment assayed for *trans*-complementation of HBV replication and, again, showed that no core particles were produced by these two proteins. This group concluded that the broader distributions of core proteins observed in the gradients suggested the formation of large aggregates of core protein insertion mutants or the formation of unstable core particles by these insertion mutants.

Two separate investigations were carried out in order to attempt to further interpret the results of the purification of the core deletion proteins. Electron microscopic examination of fractions 9-12 from the purification of the full-length core protein was performed in order to determine whether particles were present in this region of the gradient, in addition to the lower region (fractions 14-17). The results showed that core protein in these fractions was in a non-particulate state. The second investigation involved purification of the C-terminally truncated core protein, T128. After purification, this protein was present in fractions at the top of the gradient, but also in fractions as low as fraction 10. These results implied that for the core deletion proteins, protein which was observed in fractions from the top of the gradient down to fraction 12 were unlikely to represent core particle structures. Zhou and Standring (1992b) demonstrated that free core protein (fractions 1-4 of the fourteen fractions from a 10-60% sucrose gradient) existed as dimers. Therefore, the equivalent fractions on the 10-50% sucrose gradients used in the work presented in this thesis (fractions 1-7) are also likely to contain protein dimers. Unfortunately the information derived from all these results does not conclusively clarify the nature of proteins found to be present below fraction 12. Therefore, the conclusion of Beames and Lanford (1995) who stated that protein present in this region represented large protein aggregates, appears to be the only alternative explanation. In order to answer this question more satisfactorily, it would be useful to analyse the core deletion proteins

on 3-25% sucrose sizing gradients, similar to those used by Zhou and Standring (1992a). Comparison with the migration of proteins of known molecular weight, such as carbonic anhydrase (29kDa) and BSA (66kDa), on these gradients would have allowed determination of the molecular weights of the core deletion protein species. This could then be used to determine whether these proteins were present as dimers or higher order structures.

The correlation between the results obtained from expression of core deletion proteins using plasmid pR1-11 (Section 3C3) and the results presented in this section is variable. Core deletion protein Δ84-109 was present in fractions 6-8 for pR1-11 and 8-13 for pET purifications. Deletion protein Δ81-121 did show some degree of overlap with the two results (fractions 3-12 for pR1-11 and 8-19 for pET). Δ79-125 was present in fractions 6-9 for pR1-11 and 8-19 for pET purifications. Although core deletion Δ60-117 was detected in fractions 15-19 for both purifications, this protein was also present in fractions 8-14 when expressed using the pET system and purified. These differences between results from the two systems are difficult to explain. The promoters used to express protein from each plasmid were different: pR1-11 utilises the lacUV5 promoter, whereas the stronger T7 polymerase promoter is used in pET vectors. Expression from plasmid pR1-11 resulted in the production of a core:β-galactosidase protein, with ten N-terminal β-galactosidase residues replacing the two authentic N-terminal core protein residues. In contrast, the core protein produced from the pET3a plasmid initiates with its authentic methionine residue. Another factor which differed between the two sets of results was the strain of *E.coli* used for protein expression: BL21(DE3)pLysS was used for the pET3a plasmid compared to DH5α for pR1-11. The protein purification protocol for pR1-11 also included 'variable' steps, such as the efficiency of sonication and ammonium sulphate precipitation whereas results from the pET system were consistently reproducible. All these factors could possibly have contributed to the differences in results observed with the two plasmids used.

In an investigation of protein dimerisation, full-length core protein and all four core deletion proteins were shown to be able to form disulphide-linked dimers, under non-reducing conditions. The position of the deletions with respect to the four cysteine residues present in core protein also provides confirmatory information on the importance of individual residues for disulphide bond formation. As all four deletions removed Cys-107, this residue cannot be the only residue to participate in the formation of intermolecular disulphide bonds (see Section 1E2). The removal of Cys-61 by deletion Δ60-117 also shows that one of the two remaining eligible cysteine residues must be involved. At this stage the question as to which type of disulphide bond the Cys-183 residue in full-length core protein participates in becomes pertinent: is it involved in a disulphide bond within a core protein dimer or a bond between dimers? If the latter situation is true, why should dimers be the predominant species detected under non-

reducing conditions? It would seem more likely that the existence of both intra-dimer (involving Cys-61 and possibly Cys-48) and inter-dimer (involving Cys-183) disulphide bonds would enable the core particle to retain its conformation under non-reducing conditions. This would make it unlikely to enter a resolving polyacrylamide gel, as was observed by Gallina *et al.* (1989) and Kann and Gerlich (1994). However, the model resulting from the analysis of core protein expressed in *Xenopus* oocytes shows that Cys-183 participates in a disulphide bond within the dimer (Seifer and Standring, 1994). This allowed core particles to be broken down into dimer subunits under non-reducing conditions. Other investigators have observed a partial effect with bacterially-expressed core protein: dimers can enter resolving polyacrylamide gels under non-reducing conditions, but intact particles are still observed at the interface of the stacking gel and the resolving gel (Nassal *et al.*, 1992; Zheng *et al.*, 1992). In these cases, the dimeric protein which entered the resolving gel may simply represent the population of free core protein dimers present within the cells. In view of these possibilities, core deletion protein Δ60-117 could be held together by intra-dimer disulphide bonds involving Cys-48 and/or Cys-183. However, examination of this protein under non-denaturing conditions does not show convincing evidence for the existence of a higher molecular weight species that would result from the presence of Cys-183 inter-dimer disulphide bonds.

The fact that all four core deletion proteins were capable of forming dimers indicates that their inability to assemble particles may be due to the failure of subsequent stages in the particle assembly process to occur normally. Although dimers did form, this may require only the region of the polypeptide chain surrounding amino acids 48-61 to be surface exposed. The remainder of the protein may have been in a conformation that would not allow the further interactions required for particle assembly.

The inability of the core deletion proteins to assemble into particles was not thought to be due to the concentration of protein being below that estimated to be necessary for core particle assembly (Seifer *et al.*, 1993). The anti-pp65 mAb, under standard conditions of use, was able to detect protein at concentrations below this minimum.

3E Modifications of the core deletion protein Δ81-121

3E1 Introduction

From the results in the previous section, it was concluded that core deletion proteins were unable to assemble into any type of core particle-like structure. It was decided to investigate further what were the determining factors in this inability to form core particles. As was discussed in Section 3C4, the initial question was which core protein sequences are essential for directing particle assembly, either through correctly folding the protein, or through participating in protein:protein interactions, and whether these are missing in core deletions. This question was to be investigated in two ways: firstly, by replacing the deleted core protein sequence Δ81-121 with foreign sequence of exactly the same length (as described in Section 3C4) and secondly, by sequentially replacing core sequence, that had been deleted in the Δ81-121 protein, in blocks of ten amino acids from either the 5' or 3' termini of the deletion. The latter approach would hopefully allow determination of the minimum deletion, in this central region of the protein, that would still allow particle assembly and would also identify core protein sequences that played an important role in directing particle assembly. The first experimental approach is described in Section 3E2 and the second approach in Section 3E3.

3E2 pET expression and purification of pp65 epitope-tagged Δcore-surface hybrid protein

The Δcore-surface hybrid gene constructed in Section 3C4 was amplified by a single round of PCR using primer pair cp3/Bamtag and cloned into the pET3a vector. In contrast to the pR1-11 purification, this protein was efficiently expressed from the pET vector and was more stable under the conditions used for purification. The results of the sucrose gradient centrifugation are shown in Figure 3E2. The protein expressed was equal in size to full-length core protein and was detected from fraction 7 down to the bottom of the gradient. Although fractions 14-17 did seem to have an increased amount of protein present, this increase was far less than was observed for the purification of wild-type core particles. In addition, upon repetition of this experiment, this observation was not consistent: although the Δcore-surface hybrid protein was present in the same fractions, the same apparent increase in protein content in fractions 14-17 was not evident. Examination of these pooled fractions, again, gave no evidence for the assembly of particles from this protein. Although on the blots, some bands of lower molecular weight were observed in the lanes containing cellular extract and soluble protein, these represented a minority proportion of the protein detected, and were not thought to contribute to the lack of particle detection, for example, due to protein degradation. The conclusion drawn from this experiment was that this Δcore-surface hybrid protein was

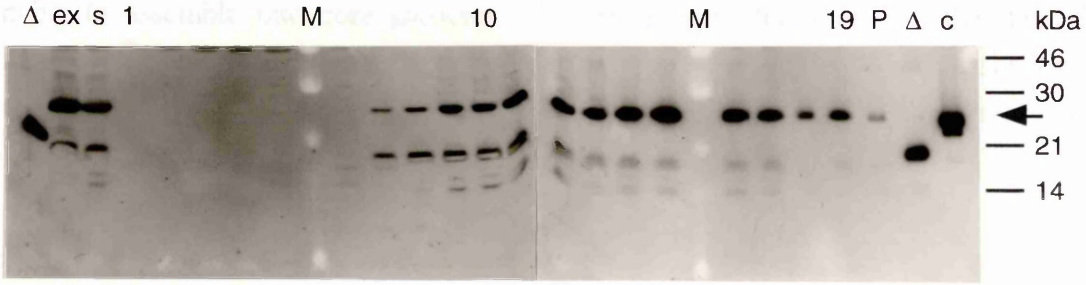


Figure 3E2: Purification of Δcore-surface hybrid protein expressed in the pET system
Δcore-surface hybrid protein with a C-terminal epitope tag was expressed in *E.coli* BL21(DE3)pLysS for 4 hr at 26°C after induction with 300μM IPTG. A cellular extract was prepared (ex) and after centrifugation at 35k rpm for 20 min, soluble protein (s) was centrifuged through a 10-50% sucrose gradient. Fractions were removed from the top (1) to the bottom (19) of the gradient and any pelleted material resuspended (P). Fractions were electrophoresed through a 17.5% polyacrylamide gel along with an aliquot of tagged full-length core protein (c) and tagged core deletion protein Δ81-121 (Δ). After Western blotting of the gels, membranes were incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown and the position of the tagged Δcore-surface hybrid protein is indicated by an arrow (←).

unable to assemble into core particles. This result indicated that, for the naturally-occurring core deletions, the physical size of the deletion was not the sole reason for the inability to form particles. Rather, specific core sequences, as yet undetermined, must play an important role in the particle assembly pathway.

3E3 Expression and purification of Δ81-121 fill-in proteins

The results from previous sections had shown that deletions in the central region of the core protein disabled the ability of the protein to correctly form core particles, and that this was due to the removal of specific protein sequence essential to this process. It was decided to attempt to determine the location and the identity of these sequences, by gradually replacing increasing amounts of original sequence in the Δ81-121 core deletion, from either end of the deleted region. This was achieved using a PCR-based method to construct these 'fill-in' proteins and is detailed in Figure 2B1.11b. This method generated three core deletion genes, 5'F10, 5'F20 and 5'F30, with 31, 21 and 11 amino acids, N-terminal to amino acid 121, deleted respectively and with the pp65 epitope tag sequence at the C-terminus (see Table 2B1.11c). An analogous method was used to construct the 3' fill-in proteins.

After expression, proteins were detected by anti-pp65 mAb 9220 (Figure 3E3a). All proteins expressed corresponded well to their predicted sizes. Proteins were then expressed and purified on 10-50% sucrose gradients. These results are shown in Figure 3E3b.

The results from purifications of both the 5'F10 and 5'F20 proteins showed gradient profiles for these proteins that were similar to that observed for the full-length core protein (see Figure 3D2b). Both showed protein predominantly in the lower region of the gradient: 5'F10 protein was detectable from fraction 8 with increased levels in fractions 14-19 and 5'F20 was also present from fraction 8, but peaked in fractions 15-16. However, the next protein in this series, 5'F30, did not conform to this profile, as protein was predominantly detectable in the central region of the gradient, in fractions 6-9. The peak fractions for all three proteins were pooled and concentrated before examination by electron microscopy. However, despite thorough efforts, no particles were detected for any of these samples.

The 3' fill-in proteins were also constructed and proteins of the correct size were produced and detected by mAb 9220 (Figure 3E3c). All three proteins were expressed and purified on 10-50% sucrose gradients, as before, and the results shown in Figure 3E3d.

Comparison of the results from this 3' fill-in series was treated with caution, as the 3'F20 protein was found to be expressed poorly initially and, after further timecourse experiments this was found to be due to poor stability. Consequently, expression took

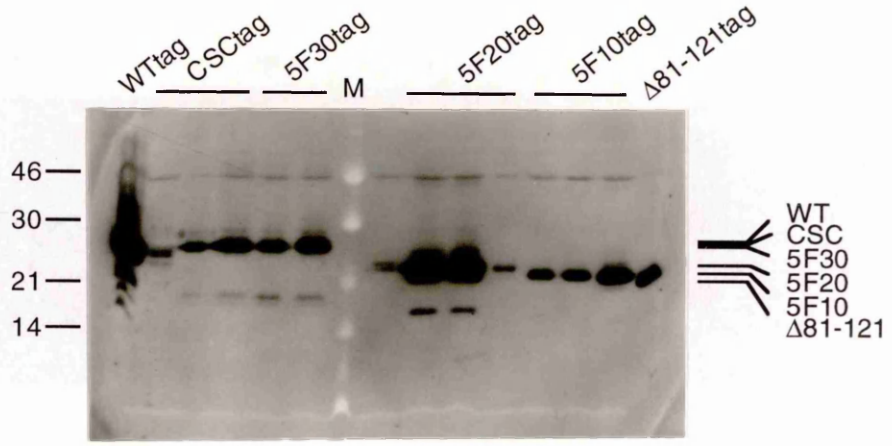


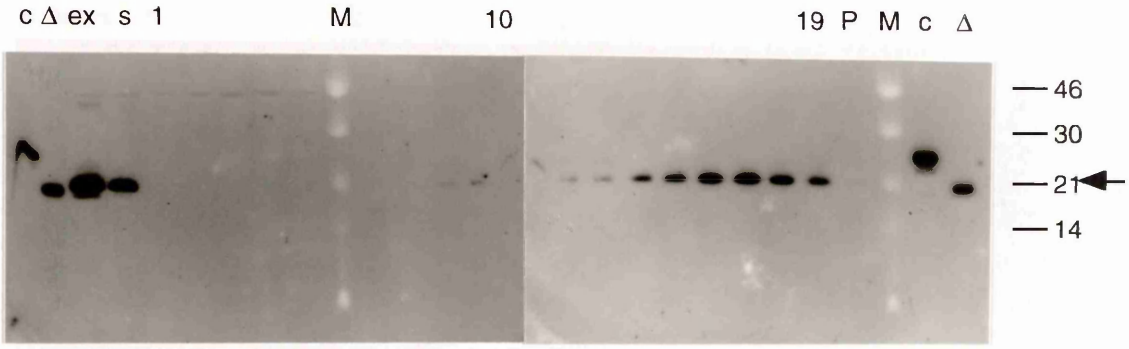
Figure 3E3a: Expression of Δ81-121 5' fill-in proteins

Δ81-121 5' fill-in proteins with a C-terminal pp65 epitope tag were expressed in *E.*coli BL21(DE3)pLysS for 4 hours at 26°C after induction with 300μM IPTG. Cellular extracts were prepared and aliquots electrophoresed through a 17.5% polyacrylamide gel alongside tagged full-length core protein (WTtag), tagged Δ81-121 core deletion protein (Δ81-121tag) and tagged Δcore-surface hybrid protein (CSCtag). After Western blotting of the gel, the blot was incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown.

place for only 2hr, which may also have accounted for the lower levels of protein produced. 3'F10 protein was detected from fraction 4, but was present predominantly in fractions 9-14. As it was probable that any core particle-like structure formed from these deletion proteins would have a smaller size in comparison to wild-type core particles, it was decided to examine these pooled fractions by electron microscopy. However, again despite thorough efforts, it was not possible to detect any particulate structures. The gradient profile for the 3'F30 protein, although protein was present in the lower region of the gradient (fractions 8 on), showed none of the 'banding' effect in any region, as was observed for the wild-type particle purification. Rather, the protein content of each fraction was approximately equal. Analysis of two sets of pooled, concentrated fractions (10-13 and 14-17) by electron microscopy showed that no particles were present.

The conclusion from the results in this section was that core proteins with deletions of only 11 amino acids at two separate positions in this central region, are still incapable of assembling into stable core particles.

i 5'F10tag



ii 5'F20tag

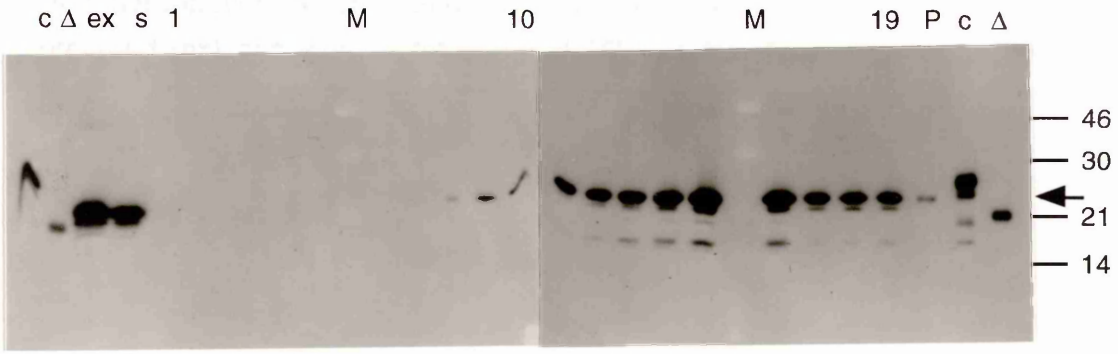


Figure 3E3b: Purification of Δ81-121 5' fill-in proteins (continued with figure legend on the following page)

iii 5'F30tag

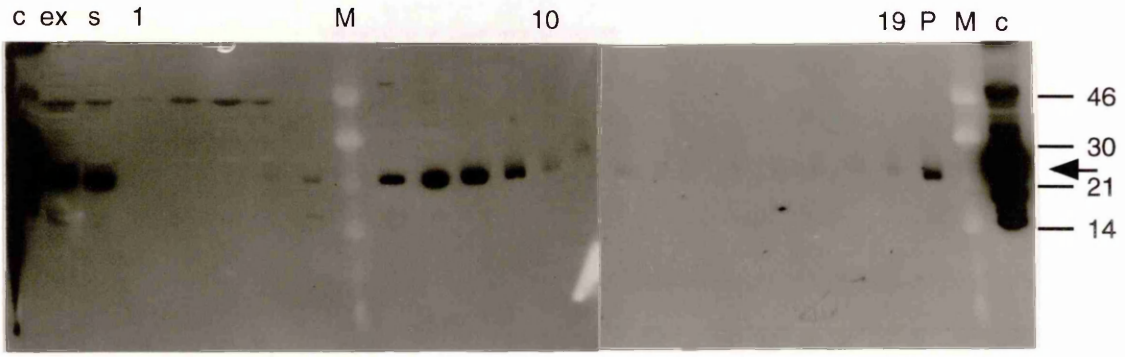


Figure 3E3b: Purification of Δ81-121 5' fill-in proteins (continued)
Δ81-121 5' fill-in proteins with C-terminal pp65 epitope tags were expressed in *E.coli* BL21(DE3)pLysS at 26°C after induction with 300μM IPTG. Cellular extracts were prepared (ex) and after centrifugation at 35k rpm for 20min, soluble protein (s) was centrifuged through a 10-50% sucrose gradient. Fractions were collected from the top (1) to the bottom (13) of the gradient and any pelleted material was resuspended (P). Fractions were electrophoresed through 17.5% polyacrylamide gels along with aliquots of tagged full-length core protein (c) and tagged core deletion protein Δ81-121 (Δ). After Western blotting of the gels, membranes were incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown and the position of the tagged proteins is indicated by an arrow (←).

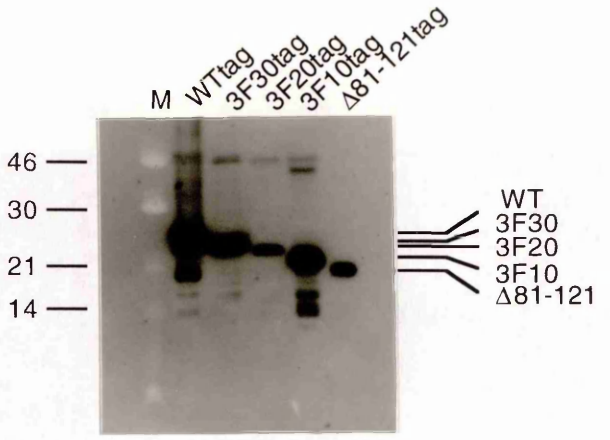
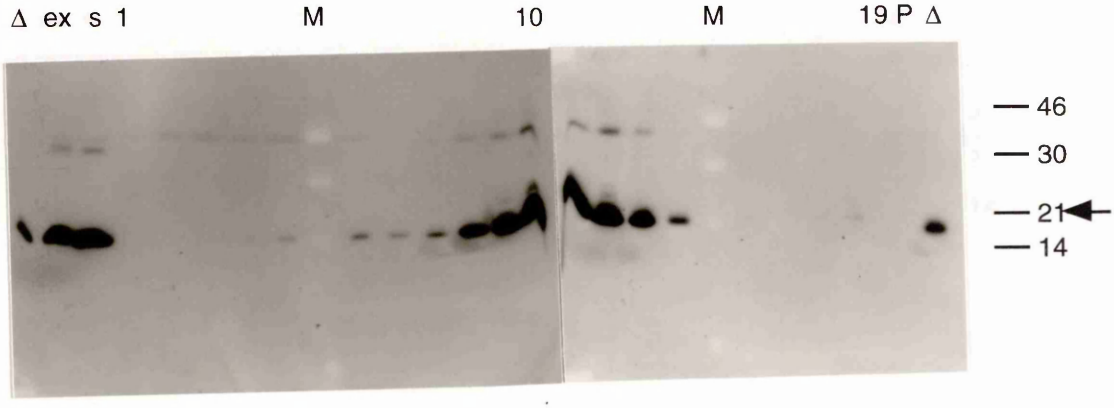


Figure 3E3c: Expression of Δ81-121 3' fill-in proteins

Δ81-121 3' fill-in proteins with a C-terminal pp65 epitope tag were expressed in *E.*coli BL21(DE3)pLysS for 4 hours at 26°C after induction with 300μM IPTG. Cellular extracts were prepared and aliquots electrophoresed through a 17.5% polyacrylamide gel alongside tagged full-length core protein (WTtag) and tagged Δ81-121 core deletion protein (Δ81-121tag). After Western blotting of the gel, the blot was incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown.

i 3'F10tag



ii 3'F20tag

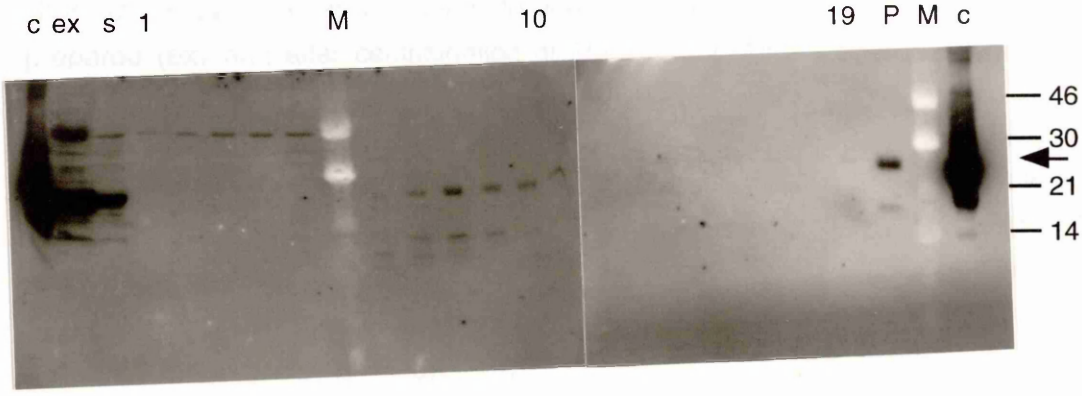


Figure 3E3d: Purification of Δ81-121 3' fill-in proteins (continued with figure legend on the following page)

iii 3'F30tag

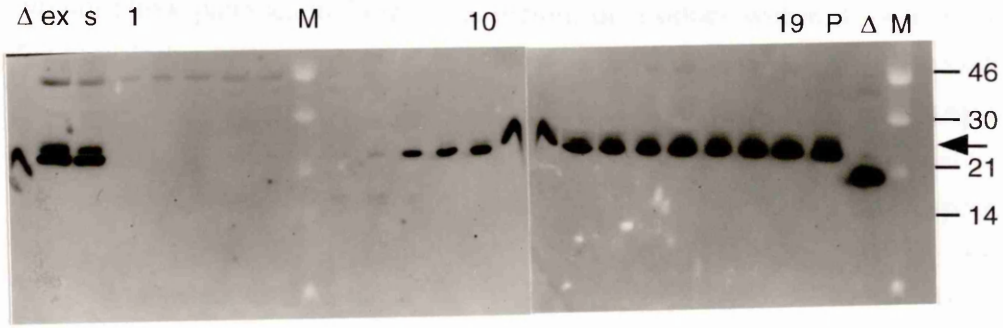


Figure 3E3d: Purification of Δ81-121 3' fill-in proteins (continued)
Δ81-121 3' fill-in proteins with C-terminal pp65 epitope tags were expressed in *E.coli* BL21(DE3)pLysS at 26°C after induction with 300μM IPTG. Cellular extracts were prepared (ex) and after centrifugation at 35k rpm for 20min, soluble protein (s) was centrifuged through a 10-50% sucrose gradient. Fractions were collected from the top (1) to the bottom (19) of the gradient and any pelleted material was resuspended (P). Fractions were electrophoresed through 17.5% polyacrylamide gels along with aliquots of tagged full-length core protein (c) and tagged core deletion protein Δ81-121 (Δ). After Western blotting of the gels, membranes were incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown and the position of the tagged proteins is indicated by an arrow (←).

3E4 Discussion

The results from this section have shown that the central region of the core protein, amino acids 81-121, is very important for correct assembly of core particles. Neither alteration of the amino acid sequence in this region of the protein nor deletion of smaller sections of this region allowed the resulting proteins to assemble into particles, as was determined by sucrose gradient analysis.

The rationale behind these experiments came from the observations of the 'structural flexibility' of the core protein (see Section 1E5). Certain fusion protein derivatives of core protein described in Section 1E5 (Schödel *et al.*, 1992) involve the removal of central core protein residues (amino acids 76-82) and replacement of these regions with foreign amino acid sequences. These core protein derivatives were able to assemble into core particles

and suggest that not all core protein residues are required for particle assembly. However, from the results presented here, replacement of the larger region from amino acid 81-121 did not allow particles to form. This region, or residues within it, is therefore important for particle formation. It was possible, however, that the region of the surface protein, used to replace the deleted core protein residues, contained amino acid sequences which would have inhibited core particle assembly. At the time of construction of the Δcore-surface gene, the only detailed structural information available on core protein was the structure predicted by Argos and Fuller (1988). This predicted that the region between amino acids 81-121 included two β-sheet structures (amino acids 89-95 and 99-115). However, the more recent biophysical analysis and high resolution structural data obtained for the core protein by electron cryomicroscopy (see Section 1E3) have shown that the previous predicted structure of core protein was incorrect. Instead, the region between amino acids 81-121 includes sections from two α-helices (helix 3: residues 82-110 and helix 4: residues 112-128). As the region from the surface protein inserted into the core protein deletion included six proline residues (see Figure 3C4a), these would have disrupted the α-helical nature of this region. Therefore, it is possible that the presence of these residues was responsible for the inability of this protein to assemble into core particles. In order to overcome this problem, another Δcore hybrid protein could be constructed. This could either be the Δcore-surface hybrid protein used in this section, but where the proline residues had been mutated to alternative residues, or a Δcore hybrid protein where the region chosen to replace amino acids 81-121 possessed a similar hydropathy profile to amino acids 81-121 of core protein.

The results from the purifications of the 5' and 3' fill-in proteins, which contained deletions of decreasing size, showed that none of the proteins with a deletion in the central region were competent for particle assembly. The smallest deletions, in the 5'F30 and 3'F30 proteins, were still sufficiently large to prevent particle assembly. The fact that both of these small deletions, which removed core protein regions with dissimilar amino acid sequences, were capable of preventing core particle assembly suggests that the size of the deletion was the critical factor. Therefore, further reducing the sizes of these deletions may have allowed particles to form.

Although the results from this section have identified the core protein region between amino acids 81-121 as important for core particle assembly, the N- and C-terminal limits of this functionally important region have not been precisely delineated. It may have been that small deletions, equal in size to the 5' and 3'30 deletions, made on either side of residues 81-121 would have actually allowed particle assembly to occur. One possible way to further define this region would be to construct a range of core protein deletion mutants, with overlapping deletions of a fixed size covering the entire core gene. Subsequent assaying of these proteins for particle assembly would allow identification of regions of the protein which were non-essential for this process. This

type of experiment has been performed for the larger core protein from DHBV (Yang *et al.*, 1994). In view of the results presented in this thesis, it was of interest that, although deletions were tolerated in the C-terminal region of the DHBV core protein, protein with N-terminal or central deletions did not assemble of core particles when expressed in *E.coli*..

3F Studies of co-assembly between full-length and core deletion proteins into core particles

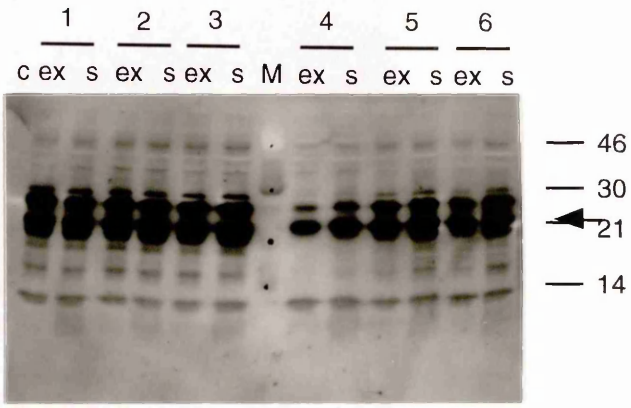
3F1 Introduction

As discussed in Section 1E6, in all but two reported cases of detection of core gene deletions from patients with chronic hepatitis investigators have, in addition, always been able to detect full-length core gene in these samples. However, because these results were all derived from the study of serum samples, the relationship between these two populations in the cellular context is still unclear. For example, it is not known whether both full-length core gene and core deletions can arise and co-exist in the same infected cell, or whether core deletions can represent the sole population in an infected cell. From the results in previous sections, where such core deletions have been shown to be incapable of forming stable core particles, it is unlikely that they would be able to sustain the viral infection themselves. Lacking the ability to produce functional core particles, which play important roles in the viral lifecycle, the virus could spread no further than its present host cell. However, if the alternative scenario in fact exists, where both virus with full-length core gene and that with core gene deletions occur in the same cell, then the possibility exists of novel core particles being assembled, containing both types of protein.

3F2 Co-expression of full-length and core deletion proteins

In order to investigate the possibility of co-assembly, it was decided to attempt to produce both proteins in the same bacterial cell and analyse the particles produced, to determine whether they contained both types of core protein. This necessitated modification of the existing bacterial expression system to introduce means of ensuring that cells were actually expressing both types of protein and means of specifically detecting each protein expressed. The full-length core gene was amplified by PCR using primer pair cp3/cd2 to give full-length core protein lacking the C-terminal pp65 epitope tag sequences. This gene was then cloned into the pET9a vector, which is identical to pET3a but with a kanamycin resistance gene instead of one conferring ampicillin resistance. pET3a plasmids containing the epitope tagged core deletion gene were as used previously. For the initial experiment the pET9a and pET3a (containing the gene encoding the Δ84-109 core deletion protein) recombinant plasmids were co-transformed into *E.coli* BL21(DE3)pLysS and co-transformants selected for in the presence of ampicillin and kanamycin. After growth of single colonies and induction of protein expression, total cellular extracts and soluble protein extracts were made and analysed by Western blot. Anti-core mAb 42B12 was used to detect core protein lacking the pp65 epitope tag and anti-pp65 mAb 9220 was used for the detection of the epitope-tagged core deletion proteins. The results of the co-expression of full-length core protein and the Δ84-109 core

A Anti-core mAb 42B12



B Anti-pp65 mAb 9220

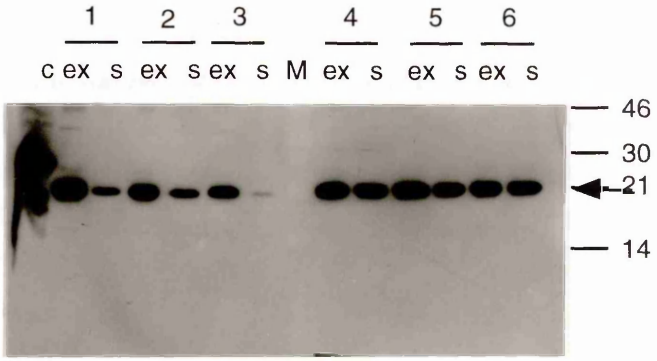


Figure 3F2a: Co-expression of full-length core protein and pp65 epitope-tagged core deletion protein Δ84-109

Full-length core protein and core protein deletion Δ84-109 were co-expressed in *E.coli* BL21(DE3)pLysS at 26°C for 4 hr after induction with 300μM IPTG. Cellular extracts (ex) and soluble protein samples (s) were prepared from six clones and electrophoresed through 17.5% polyacrylamide gels. After Western blotting, the blots were incubated with anti-core mAb 42B12 (Panel A) or anti-pp65 mAb 9220 (Panel B). An aliquot of tagged Δcore-surface hybrid protein (c) was included as a control. Molecular weights of the markers used (M) are shown and the position of each protein is indicated by an arrow (←).

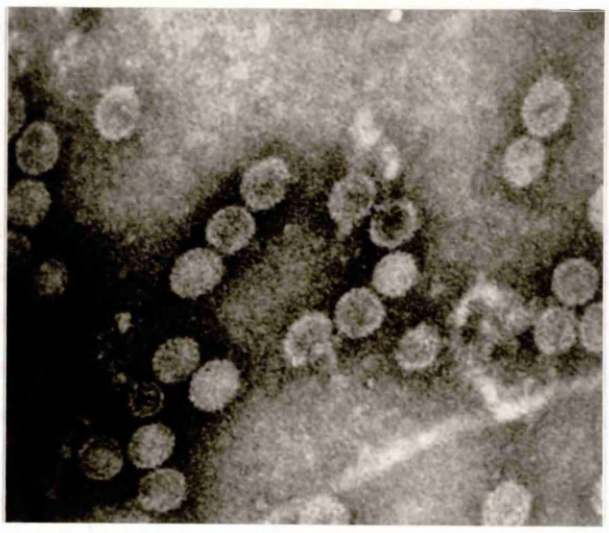


Figure 3F2b: Electron micrograph of core particles expressed using the pET system
Core particles were purified from extracts of bacteria transformed with a pET9a plasmid containing full-length core gene, using a 10-50% sucrose gradient. Fractions containing the highest levels of core protein as identified by Western blot analysis, were pooled and concentrated using an Amicon 50 microconcentrator. An aliquot of the concentrated sample was negatively stained and examined by electron microscopy. Scale bar represents 50nm.

deletion protein are shown in Figure 3F2a. A protein of the correct predicted size for the core deletion protein Δ84-109 was specifically detected on the blot using mAb 9220. However the blot on which mAb 42B12 was used showed two major bands for all samples of approximately 23kDa and 27kDa. The smaller of these bands was the correct size for the full-length core protein, but the identity of the larger band was undetermined. This protein may have been translated due to read-through of the core gene stop codon. Cellular extracts from clones 1, 2, 3 and 5 also contained a protein of approximately 29kDa, which reacted with the anti-core mAb 42B12. In order to determine whether the presence of these extra proteins would affect core particle assembly, *E.coli* BL21(DE3)pLysS cells were transformed with the pET9a plasmid containing the untagged full-length core gene. Following sucrose gradient centrifugation of induced cellular extracts and Western blot analysis, fractions with the greatest amounts of core protein were pooled, concentrated and examined by electron microscopy. The results showed that the full-length core protein which had been expressed was able to assemble into morphologically normal core particles (see Figure 3F2b). The results from the

experiments in this section showed that expression of both proteins in the same cell was possible.

3F3 Purification of co-expressed full-length and Δ84-109 core proteins

Proteins were expressed from bacteria which had been co-transformed with plasmids containing the full-length core and tagged core deletion protein Δ84-109, as described in the previous section. Cellular extracts were prepared and, after clarification, soluble protein was centrifuged through a 10-50% sucrose gradient. The results are shown in Figure 3F3. Full-length protein expressed from the pET9a plasmid was detected strongly in fractions 13-17, which corresponded well to the position of core particles from previous experiments (see Figure 3D2b). The tagged Δ84-109 protein was present in two peaks in the gradient: one in fractions 7-11 and the other in fractions 13-17. The co-existence of the Δ84-109 protein in fractions containing core particles suggested that both the full-length core protein and the core deletion protein were present within the same particles. Further evidence for this came from a comparison with the results obtained from the purification of the Δ84-109 protein alone (Figure 3D2d). These results showed that the Δ84-109 protein was detected in fractions 8-13. However, in the results shown here, the Δ84-109 protein was also present in the lower region of the gradient, in fractions 13-17. This position corresponded very well to that of the core particles produced in the same cells. In addition, the fact that the Δ84-109 protein was not distributed equally throughout fractions 7-17, but gave a distinct peak in fraction 13-17 suggested that it was associated with the core particles in these fractions.

3F4 Purification of co-expressed full-length and Δ81-121 core proteins

It was decided to perform the same type of co-expression experiment for the core deletion protein Δ81-121 and full-length core protein, to determine if a similar result would be obtained. This protein was co-expressed with the full-length protein and the particles produced were purified (see Figure 3F4). The full-length protein was present in fractions 13-17, as before, and the Δ81-121 deletion protein was detected in two peaks: fractions 9-11 and fractions 14-17. Again the location of the Δ81-121 protein differed from the results obtained when this protein was expressed by itself and purified (Figure 3D2d). These results showed that the Δ81-121 deletion protein migrated to fractions 8-19, with more protein present in fractions 8-14. These results differ markedly to the results presented in this section, where the Δ81-121 protein was detected in two distinct peaks. The position of the second peak corresponded well with the position of the core particles produced, which again led to the possibility that these two proteins had associated and formed novel 'mixed' core particles.

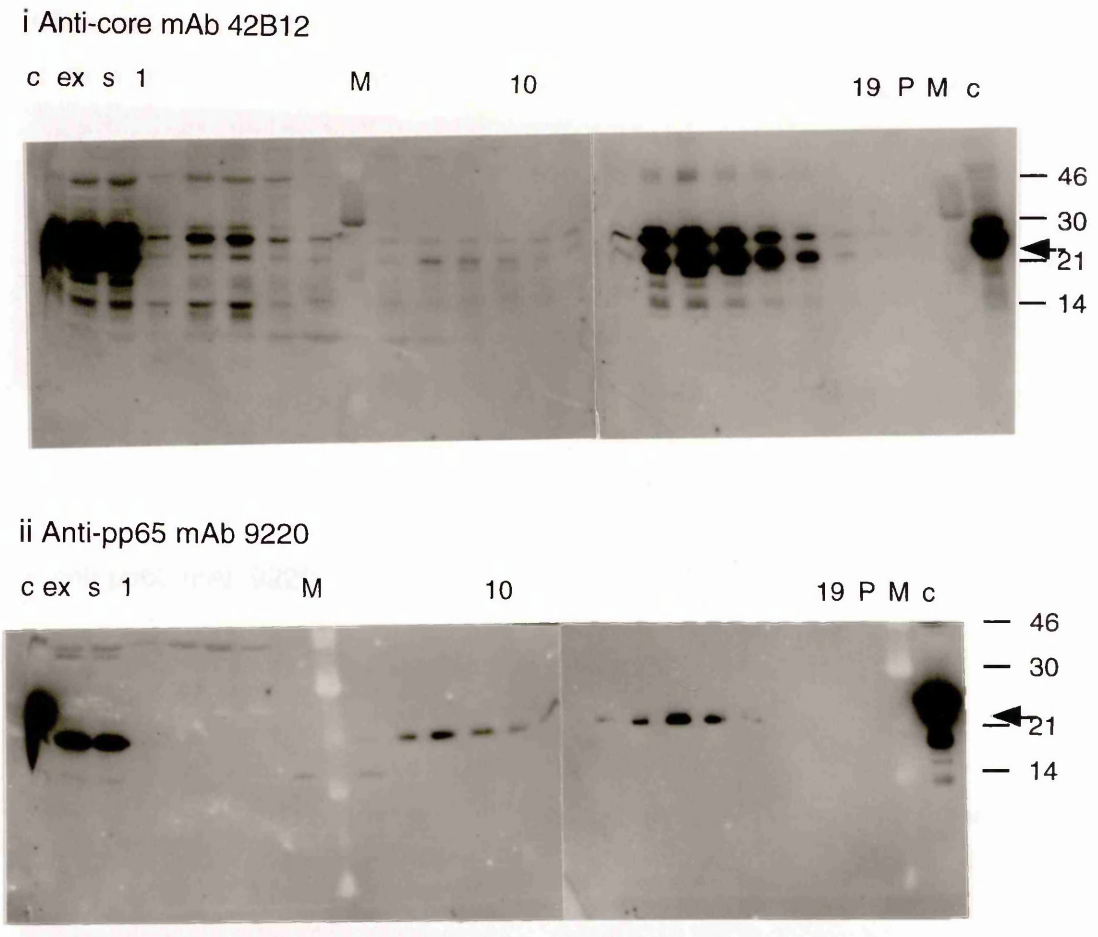


Figure 3F3: Purification of co-expressed coreWT and Δ84-109tag

Full-length core protein and pp65 epitope-tagged core deletion protein Δ84-109 were co-expressed in *E.coli* BL21(DE3)pLysS at 26°C for 4 hours after induction with 300μM IPTG. Cellular extracts were prepared (ex) and after centrifugation at 35k rpm for 20 min, soluble protein (s) was centrifuged through a 10-50% sucrose gradient. Fractions were collected from the top (1) to the bottom (19) of the gradient and any pelleted material was resuspended (P). Fractions were electrophoresed through 17.5% polyacrylamide gels along with tagged full-length core protein (c). After Western blotting of the gels, membranes were incubated with either anti-core mAb 42B12 (panel i) or anti-pp65 mAb 9220 (panel ii). Molecular weights of the markers used (M) are shown and the position of the relevant protein is indicated by an arrow (←).

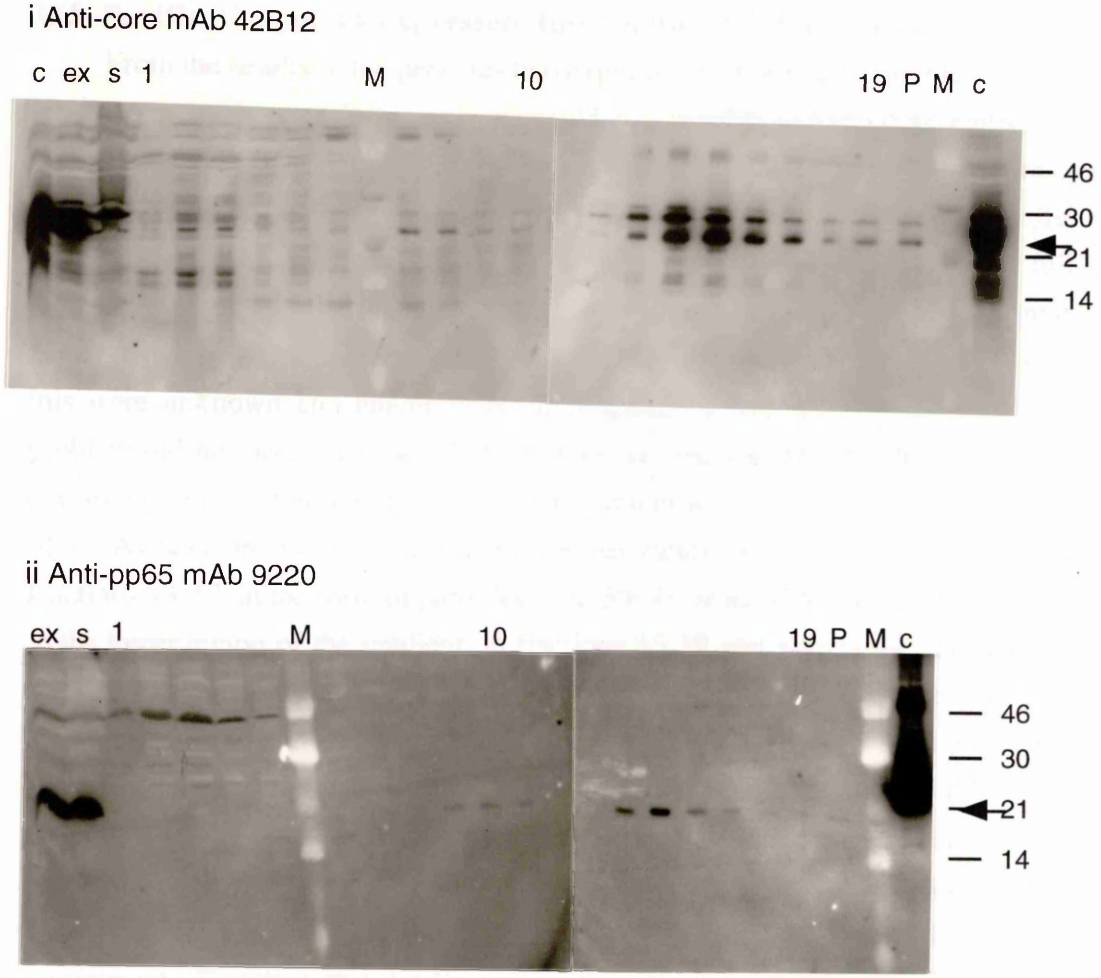


Figure 3F4: Purification of co-expressed coreWT and Δ81-121tag

Full-length core protein and pp65 epitope-tagged core deletion protein Δ81-121 were co-expressed in *E.coli* BL21(DE3)pLysS at 26°C for 3 hours after induction with 300μM IPTG. Cellular extracts were prepared (ex) and after centrifugation at 35k rpm for 20 min, soluble protein (s) was centrifuged through a 10-50% sucrose gradient. Fractions were collected from the top (1) to the bottom (19) of the gradient and any pelleted material was resuspended (P). Fractions were electrophoresed through 17.5% polyacrylamide gels along with tagged full-length core protein (c). After Western blotting of the gels, membranes were incubated with either anti-core mAb 42B12 (panel A) or anti-pp65 mAb 9220 (panel B). Molecular weights of the markers used (M) are shown and the position of the relevant protein is indicated by an arrow (←).

3F5 Purification of co-expressed full-length and 5'F30 core proteins

From the results of the previous two experiments it seemed possible that full-length core protein and core deletion proteins could co-assemble to form core particles. It was decided to see whether a more conclusive result could be obtained by using proteins with smaller deletions. The 5'F30 and 3'F30 fill-in proteins from Section 3E3 were consequently used in these co-expression experiments. When co-expression of full-length and 3'F30 proteins was attempted it was found that, although full-length protein was produced effectively, detection of the 3'F30 protein was not possible. The reasons for this were unknown and unable to be investigated further, due to time restraints. This problem did not occur for the 5'F30 protein and the particles produced from cells co-expressing this protein and full-length core protein were able to be purified (see Figure 3F5). As was observed in the previous experiments, the full-length protein banded in fractions 13-17, in the form of particles. The 5'F30 protein, however, was detected only in the lower region of the gradient, in fractions 15-19 and also in the pelleted material at the bottom of the centrifuge tube. Another lower molecular weight protein of approximately 17kDa, was detected in fractions 6-8. The location of the 5'F30 protein in this co-expression experiment differed from the results obtained when this protein was expressed by itself (Figure 3E3b). When expressed in isolation and purified on a sucrose gradient, the 5'F30 protein was present in fractions 6-9. These fractions correspond only with the fractions in which the 17kDa protein was detected in the co-expression experiment. Therefore this 17kDa protein may represent a degradation product of the 5'F30 protein. The shift in location that was observed for the 5'F30 protein, upon co-expression with the full-length protein, to a region further down the gradient appeared to suggest that these two proteins had associated to form particles. However, upon closer examination of the results it was concluded that this was not the case. The fractions containing full-length core protein (fractions 13-17) did not correlate exactly with those containing the 5'F30 protein (fractions 15-19). This is not the expected result if the two proteins are covalently linked to form a particulate structure.

3F6 Discussion

The results in this section have shown the possibility that core protein and certain core deletion proteins could co-assemble to form novel core particles. Upon co-expression with full-length protein, core deletion proteins Δ84-109 and Δ81-121 migrated to different regions of the sucrose gradient, compared to when either deletion protein was expressed alone and purified. These new regions corresponded well to the fractions where purified core particles were observed. In this lower gradient region these proteins were also observed to be present in relatively few fractions, in a manner analogous to purified core particles. However, the results from the purification of co-expressed full-length core and 5'F30 proteins did not follow this same pattern. It was concluded that

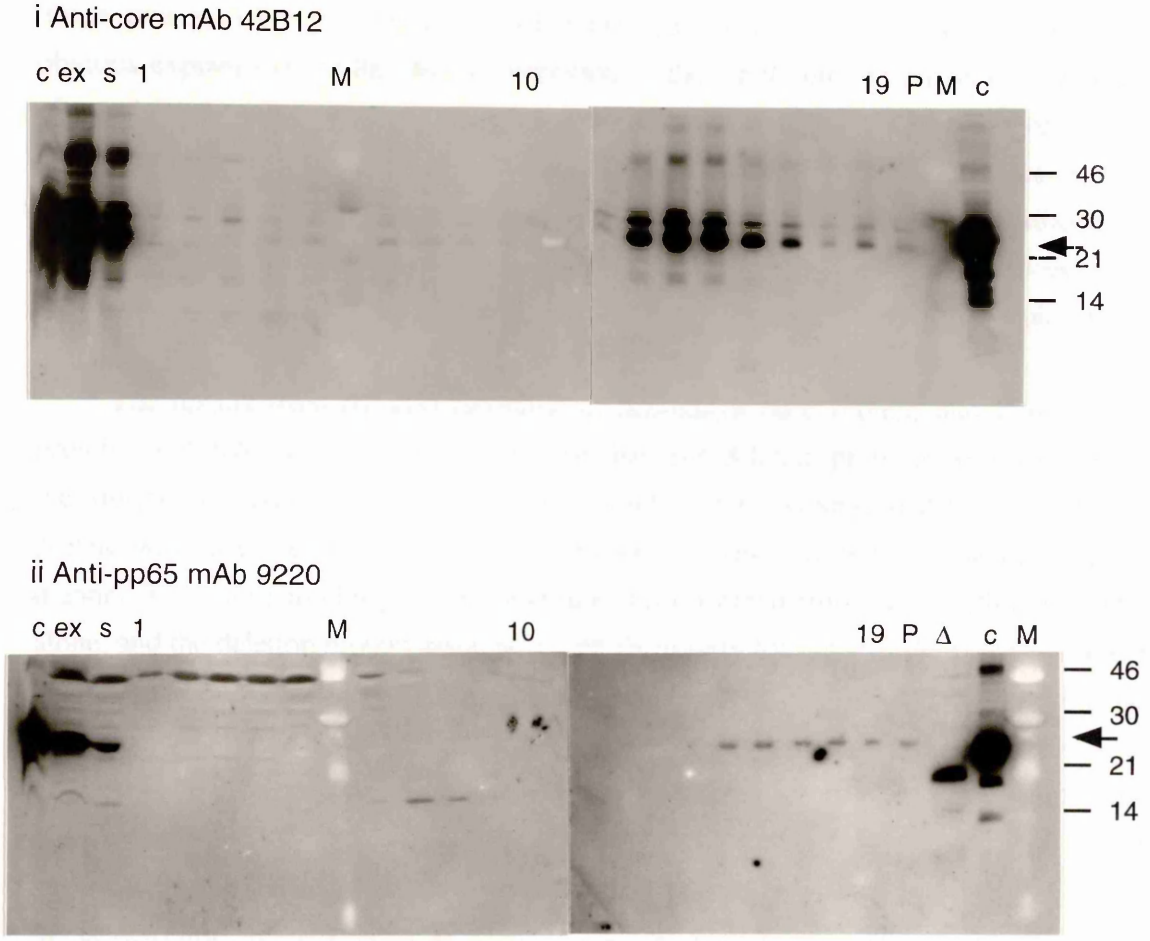


Figure 3F5: Purification of co-expressed coreWT and 5'F30tag

Full-length core protein and pp65 epitope-tagged Δ81-121 5'F30 were co-expressed in *E.coli* BL21(DE3)pLysS at 26°C for 4 hours after induction with 300μM IPTG. Cellular extracts were prepared (ex) and after centrifugation at 35k rpm for 20 min, soluble protein (s) was centrifuged through a 10-50% sucrose gradient. Fractions were collected from the top (1) to the bottom (19) of the gradient and any pelleted material was resuspended (P). Fractions were electrophoresed through 17.5% polyacrylamide gels along with tagged full-length core protein (c). After Western blotting of the gels, membranes were incubated with either anti-core mAb 42B12 (panel i) or anti-pp65 mAb 9220 (panel ii). Molecular weights of the markers used (M) are shown and the position of the relevant protein is indicated by an arrow (←).

these two proteins did not co-assemble into particles. This conclusion provided no obvious explanation for the shift in migration of the 5'F30 protein on sucrose gradients after being co-expressed with the full-length protein. It may have been that the presence of the two proteins in the same cell could have contributed to an increased aggregation of the 5'F30 protein, with these larger aggregates migrating further down the gradient. The assembly of core particles was not interfered with presumably because this process is co-operative, so leaving less time for other proteins to interfere with the rapid particle assembly.

The results from the co-expression of full-length core protein and core deletion proteins Δ84-109 and Δ81-121 only suggest that core deletion proteins co-assemble with the full-length protein to form novel core particles. It is possible that the core deletion protein was only associated with wild-type core particles in a non-integrated, peripheral manner. The core particles present could have been formed from full-length core protein alone, and the deletion protein associated with them only by non-covalent interactions. In order to provide more evidence for the occurrence of co-assembly, it would be necessary to demonstrate that both proteins were still present in samples of 'mixed' particles which had been previously incubated in buffers of increasing salt concentrations and then repurified on sucrose gradients. These increasingly stringent conditions should disrupt any non-covalent interactions, while leaving the disulphide-linked particle intact. A more direct detection method, such as immune-electron microscopy, would also have been useful. With this method, anti-core mAb 42B12 could be conjugated to a large diameter gold particle and anti-pp65 mAb 9220 conjugated with a gold particle of smaller diameter. Preparations of particles re-purified after being exposed to the above conditions would be incubated with both antibody-conjugates and examined by electron microscopy. The direct visualisation of both conjugated antibodies binding to the same particles would prove the presence of both proteins in the particle.

The formation of hybrid or mixed core particles has been demonstrated recently by Scaglioni *et al.* (1997). Mammalian cells were transiently co-transfected with a plasmid expressing-length core protein and one expressing an epitope-tagged p22 pre-core protein. When immunoprecipitation experiments were carried out on these transfected cells, both proteins were found to co-immunoprecipitate, with the conclusion that hybrid core particles had formed. In addition, purification of core particles from these transfected cells on sucrose gradients showed that both proteins were present in the same fractions, as detected by Western blot analysis. In similar immunoprecipitation experiments, an 18kDa core protein containing ten C-terminal pre-core residues and truncated at Pro-144 was also shown to be able to co-assemble with full-length core protein into mixed core particles.

3G Modifications of the core particle purification protocol

3G1 Introduction

Throughout the course of these studies, the purification protocol for the core deletion proteins had been altered a number of times to take various factors into account. When these results were compared it was discovered that the method of preparation of cellular extracts and soluble proteins, prior to sucrose gradient centrifugation, could significantly affect the conclusion drawn from the experiment. In order to illustrate this the results obtained for each modification of the purification of core deletion protein Δ84-109 are shown in Figure 3G. For all these modifications the results obtained for purification of the full-length protein were identical to those presented in Section 3D.

3G2 Modification of the purification protocol for core deletion protein Δ84-109

The initial extract preparation method consisted of overnight freezing of the induced *E.coli* BL21(DE3)pLysS cell pellet. After thawing of the pellet on ice, addition of TNE II buffer and DNase/RNase treatment, the extract was sonicated to disrupt any remaining cellular material. The extract was then clarified of cellular debris by a 4°C centrifugation step at 13k rpm for 20 minutes, prior to sucrose gradient purification. The results are shown in Figure 3G, Panel A.

The first modification to this protocol was to increase the centrifugation speed during the clarification step to 35k rpm. This was intended to ensure that only protein which was truly soluble was loaded onto the sucrose gradients, which would hopefully eliminate some of the bands detected in the lower region of the gradient. From the results which had accumulated at this stage in the study, it was already suspected that these bands represented aggregates of expressed protein, rather than particles. The results are shown in Figure 3G, Panel B.

Due to the extremely efficient lysis of the *E.coli* BL21(DE3)pLysS cells after the freezing stage, it was thought that the sonication step was no longer required and may, in fact, have caused damage to any particle-like structures which had formed. This was of obvious concern, especially if these particles were less stable than the wild-type particles. These results are shown in Figure 3G, Panel C.

The final modification was designed to cater for the possibility that, despite sucrose gradient fractions not being frozen at any stage, the freezing step required for lysis of the BL21(DE3)pLysS cells was responsible for disrupting any particles that had actually assembled inside these cells. This modification involved overnight storage of the induced cell pellet at 4°C, rather than -20°C, followed by lysis of the cells using lysozyme extraction buffer (see Figure 3G, Panel D). Although protein was detected in the gradient factions, this extraction protocol was not as efficient as the previous method. It was

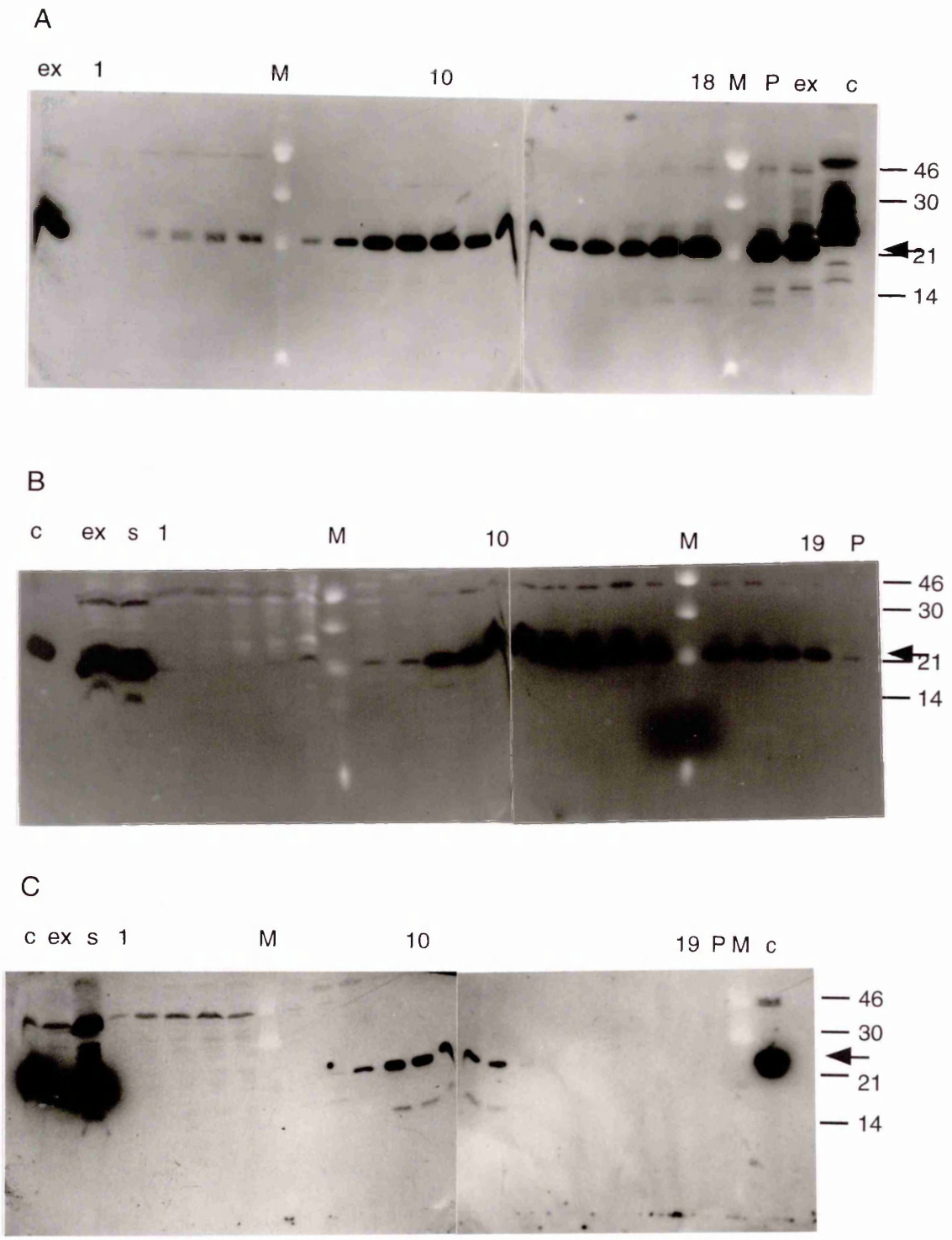


Figure 3G: Modifications to the core particle purification method (continued with figure legend on the following page)

D

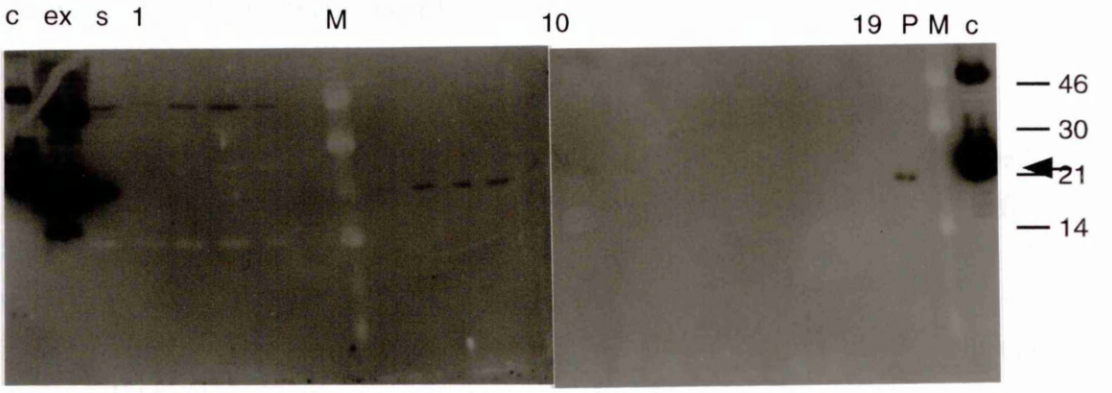


Figure 3G: Modifications to the core particle purification method (continued)
Core deletion protein Δ84-109 with a C-terminal pp65 epitope tag was expressed in *E.coli* BL21(DE3)pLysS at 26°C for 4 hr after induction with 300μM IPTG. Cellular extracts (ex) and soluble protein samples (s) were prepared as described in the text and according to the summary table below. Clarified extracts were centrifuged through 10-50% sucrose gradients. Fractions were collected as normal and analysed by Western blotting using anti-pp65 mAb 9220. Lanes marked 'c' contain tagged full-length core protein. Molecular weights of the markers used (M) are shown and the position of the tagged protein is indicated by an arrow (←).

		Extract preparation				Extract clarification (rpm)	
		F/T	4°C	S	L	13k	35k
Panel	A	X		X		X	
	B	X		X			X
	C	X					X
	D		X		X		X

F/T - overnight freezing at -20°C, followed by thawing on ice
4°C - overnight storage at 4°C
S - disruption by sonication
L - extraction with lysozyme

thought that this may have caused problems in the detection of proteins that had been expressed at comparatively lower levels.

3G3 Discussion

The results from this section have shown that the method employed for the preparation of cellular extracts and soluble protein extracts could greatly influence the results obtained from the subsequent sucrose gradient purification step. The final extraction method, using lysozyme extraction buffer (Panel D), was thought to be least likely to disrupt any potentially unstable particle-like structures formed. However, the decreased amounts of protein extracted using this protocol reduced the sensitivity of the protein detection. For this reason, the modification involving freeze/thaw of the bacterial cell pellet, followed by clarification at 35k rpm (Panel C), was chosen for routine use and yielded all the results shown in previous sections. Despite the freeze/thaw step involved in this protocol, which may have disrupted unstable particles, protein purified using this modification was detected in exactly the same fractions of the gradient as protein purified following lysozyme extraction. In addition, increased amounts of protein could be extracted from bacterial cells using this protocol, when compared with using the lysozyme extraction method. This aided detection of the proteins in the subsequent purification steps.

3H Agarose gel assay for core particle formation

3H1 Introduction

During the last stage of the experimental work presented in this thesis we were made aware of an additional method for the detection of core particles. This was described by Birnbaum and Nassal (1990) and involved electrophoresis of core particle samples through a 1% agarose gel. The particles were able to enter the gel, which was then first stained with ethidium bromide, to detect the RNA known to be incorporated into core particles. After destaining, the gel was stained with Coomassie blue to detect the position of the protein. It was decided to use this assay on the samples for which concentrated material remained after electron microscopic examination.

3H2 Results

The results for the selection of samples assayed are shown in Figure 3H. These results showed that wild-type core particles could be detected by this method. The position of the band after ethidium bromide staining of the gel (Panel A, lanes 1-4) corresponded exactly with the position of the band after Coomassie staining(Panel B, lanes 1-4). This indicated that core particles had formed and were capable of encapsidating nucleic acid. Investigation of the samples of concentrated fractions resulting from the purification of the core deletion proteins (Sections 3D and 3E), however, did not give equivalent results. Both nucleic acid and protein bands were observed at the same positions after ethidium bromide and Coomassie staining for core deletion protein Δ84-109 fractions 8-10 (lane 5) and also to a lesser extent for core deletion proteins Δ79-125 fractions 9-12 (lane 9) and Δ60-117 fractions 9-12 (lane 11). However the earlier results for these samples, from examination by electron microscopy, showed that no particulate structures were detected. Furthermore, these assay results were not consistent when equivalent samples from duplicate purification experiments were tested, as can be seen from a comparison of the samples in lanes 6-8 with those in lanes 21-24.

The concentrated fractions from the purification of the 5' fill-in proteins did give results similar to those observed for the full-length protein, with bands visible at the same position after both staining procedures (lanes 13, 15 and 17). This suggested that particulate structures had formed and were able to encapsidate nucleic acid. However, again, these results did not concur with those obtained from the electron microscopic examination of these samples, where no particles were detected (Section 3E).

3H3 Discussion

The results obtained from this type of assay gave further evidence to the conclusion that the core deletion proteins examined in this study were unable to assemble into stable core particle-like structures. No bands were visible at the same positions when the gel

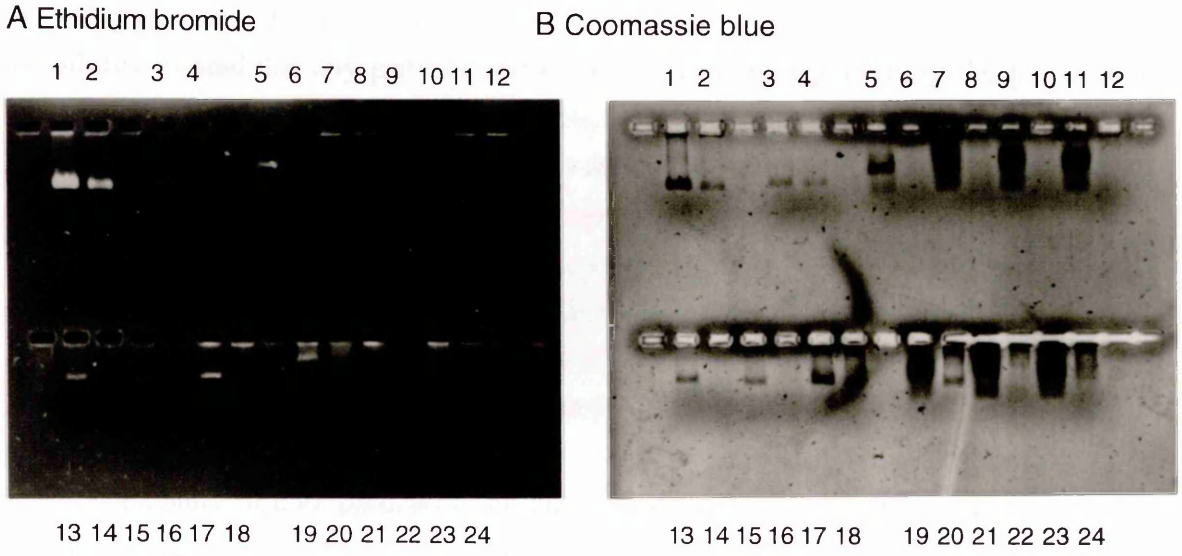


Figure 3H: Agarose gel assay for core particle detection

Peak fractions from a number of sucrose gradient purifications were pooled and concentrated using a Centricon-50 microconcentrator. Samples were loaded onto a 1% TAE agarose gel, according to the legend shown below, and electrophoresed. Panel A shows the gel after staining with ethidium bromide and panel B shows the same gel after destaining and subsequent staining with Coomassie blue.

Lane	Sample	Lane	Sample
1	Wild-type core - 10μg	13	5'F10 fractions 12-14
2	Wild-type core - 3μg	14	5'F10 fractions 15-17
3	Wild-type core - 1μg	15	5'F20 fractions 12-14
4	Wild-type core (pET9a) - 1μg	16	5'F20 fractions 15-17
5	Δ84-109 fractions 8-10	17	5'F30 fractions 12-14
6	Δ84-109 fractions 15-17	18	5'F30 fractions 15-17
7	Δ81-121 fractions 8-11	19	Wild-type core fractions 9-12
8	Δ81-121 fractions 16-19	20	Wild-type core fractions 14-17
9	Δ79-125 fractions 9-12	21	Δ84-109 fractions 9-12
10	Δ79-125 fractions 15-18	22	Δ84-109 fractions 13-16
11	Δ60-117 fractions 9-12	23	Δ81-121 fractions 9-12
12	Δ60-117 fractions 16-19	24	Δ81-121 fractions 13-16

was stained with ethidium bromide and subsequently with Coomassie blue. Again the possibility existed that any particles formed were either damaged during the purification process, or were so unstable as to be undetectable using the methods described in this thesis. The results for the 5' fill-in proteins did appear to show convincing evidence for the assembly of particles that possessed the ability to encapsidate nucleic acid. However, these putative particles were not observed by electron microscopy. For all non-particulate samples, large amounts of protein were detectable after Coomassie staining as a smear. This was of a different size to the band of core particles and agreed with the previous conclusion that large aggregates, rather than particles, were formed when these deletion proteins were expressed.

The ability of core particles to encapsidate RNA could have been exploited further as an assay for particle assembly by the core deletion proteins. Bacterial cells could have been grown in the presence of α^{32}P-UTP and protein expression induced. Following the preparation of cellular extracts, unprotected nucleic acids would be digested with DNase and RNase treatment of the extract. The clarified extract would be centrifuged through a sucrose gradient and fractions analysed by Western blot. Encapsidated RNA could be detected by exposure of a dried SDS PAGE gel of the gradient fractions to X-ray film or by scintillation counting of the gradient fractions. If RNA had been encapsidated within particulate structures formed from core deletion proteins, the peak gradient fractions containing core deletion protein should correspond with the peak α^{32}P-UTP-RNA-containing fractions.

It may be possible that RNA could be present within an aggregate of proteins due to interactions with the basic C-terminal region of the proteins. This may result in the RNA being protected from the nuclease treatment of cellular extracts. This would explain the correlation in the position of the bands for the 5'F10, 5'F20 and 5'F30 core deletion proteins after both staining procedures. However, if this was the case, similar results would be expected for other deletion proteins.

3J Investigation of the cellular localisation of core deletion proteins

3J1 Introduction

Core protein has been shown by a number of groups to localise to the nucleus, cytoplasm or both, in cells either infected with HBV or in cells transfected with the core gene alone (see Section 1E4). The signal directing nuclear localisation of core protein was contained in the C-terminal arginine-rich region of the protein and may be comprised of two distinct regions. This localisation has also been shown to be cell cycle stage-dependent (Yeh *et al.*, 1993). Core protein localised to the nucleus in the G_0/G_1 phase of the cell cycle, with the amount localising increasing during G_1 phase. When cells were in S phase, however, the protein was undetectable in the nucleus and instead accumulated in the cytoplasm.

It was decided to examine the subcellular localisation of the core deletion proteins using immunofluorescence assays. The location of the protein can be determined at a specific stage in the cell cycle by the use of aphidicolin, which arrests cells at the G_1/S boundary (Pedrali-Noy *et al.*, 1980). The epitope-tagged core gene deletions, along with a full-length tagged core gene, were amplified by PCR using primer pair cpk/Hintag. These primers included *Eco* RI and *Hind* III sites respectively, and primer Hintag also included the sequences for the pp65 epitope tag. The PCR products were digested and cloned into the vector pRK5 (see Section 2A1) such that the core genes were under the control of the CMV IE promoter. After transfection of aphidicolin-treated HepG2 cells with these plasmids, cells were fixed and labelled with either a Zymed anti-core polyclonal antibody or the anti-pp65 mAb 9220. Labelling was detected using FITC-conjugated second antibodies.

3J2 Expression of core deletion proteins in mammalian cells

An initial experiment was carried out to determine the correct expression of protein from this vector and detection by anti-pp65 mAb 9220. Plasmids were transfected into COS 7 cells and total cellular extracts prepared. These extracts were then analysed by Western blot using the anti-pp65 mAb 9220. The results are shown in Figure 3J2. No bands are visible in the lanes containing extracts from cells transfected with vector alone or untagged core protein (kindly provided by Ed Dornan - Institute of Virology). The core deletion proteins were all specifically detected, but were shown to be expressed at different levels, with the Δ79-125 protein expressed at the lowest level. An additional protein of approximately 44kDa was detected in the extract from cells transfected with the plasmid expressing epitope tagged full-length core protein. This may represent the dimeric form of this protein, as was examined in Section 3D3. Additional proteins of lower molecular weight were also detected in extracts from cells transfected with the plasmids

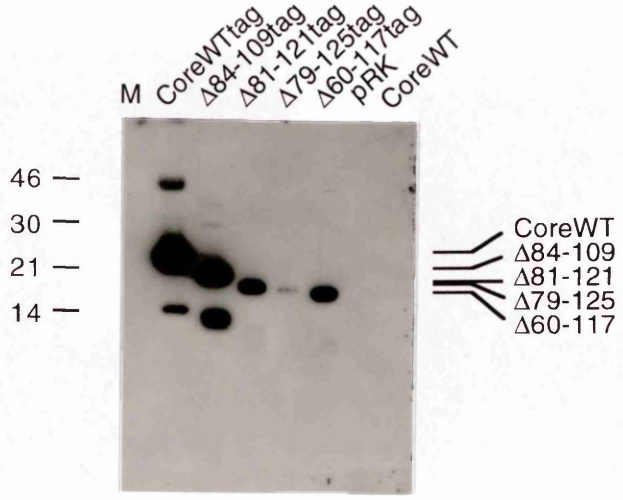


Figure 3J2: Expression of epitope-tagged core deletion proteins in transfected COS 7 cells

COS 7 cells in 60mm dishes were transfected with 5μg of pRK5 plasmid expressing tagged full-length core protein (CoreWTtag), tagged core deletion proteins (Δtag), untagged full-length core protein (CoreWT) or vector only (pRK), as a control. After a 24hr incubation following replacement of the transfection medium, cells were scraped and pelleted by centrifugation. The cell pellet was washed twice with PBS, finally resuspended in SDS PAGE sample buffer and heat-denatured at 100°C for 5 min. Samples were electrophoresed through a 17.5% polyacrylamide gel and Western blotted. The blot was incubated with anti-pp65 mAb 9220. Molecular weights of the markers used (M) are shown.

expressing epitope tagged full-length core protein and core deletion protein Δ84-109, but these may have only been due to protein degradation during preparation of the extracts.

3J3 Immunofluorescence assays

Although the core protein has been shown to localise to the nucleus under the conditions used here, the results obtained in this section were found to differ (see Figure 3J3). The full-length core protein did not localise exclusively to the nucleus, but was present in both nuclear and cytoplasmic regions of the cells, when examined with the anti-pp65 mAb 9220 (Panel A). The same localisation pattern was detected when the polyclonal antibody was used (Panel G). Upon closer examination of the amino acid sequence of this particular protein it was observed that a Gly>Cys mutation was present in the C-terminal region of the protein, at amino acid 153. As the nuclear localisation

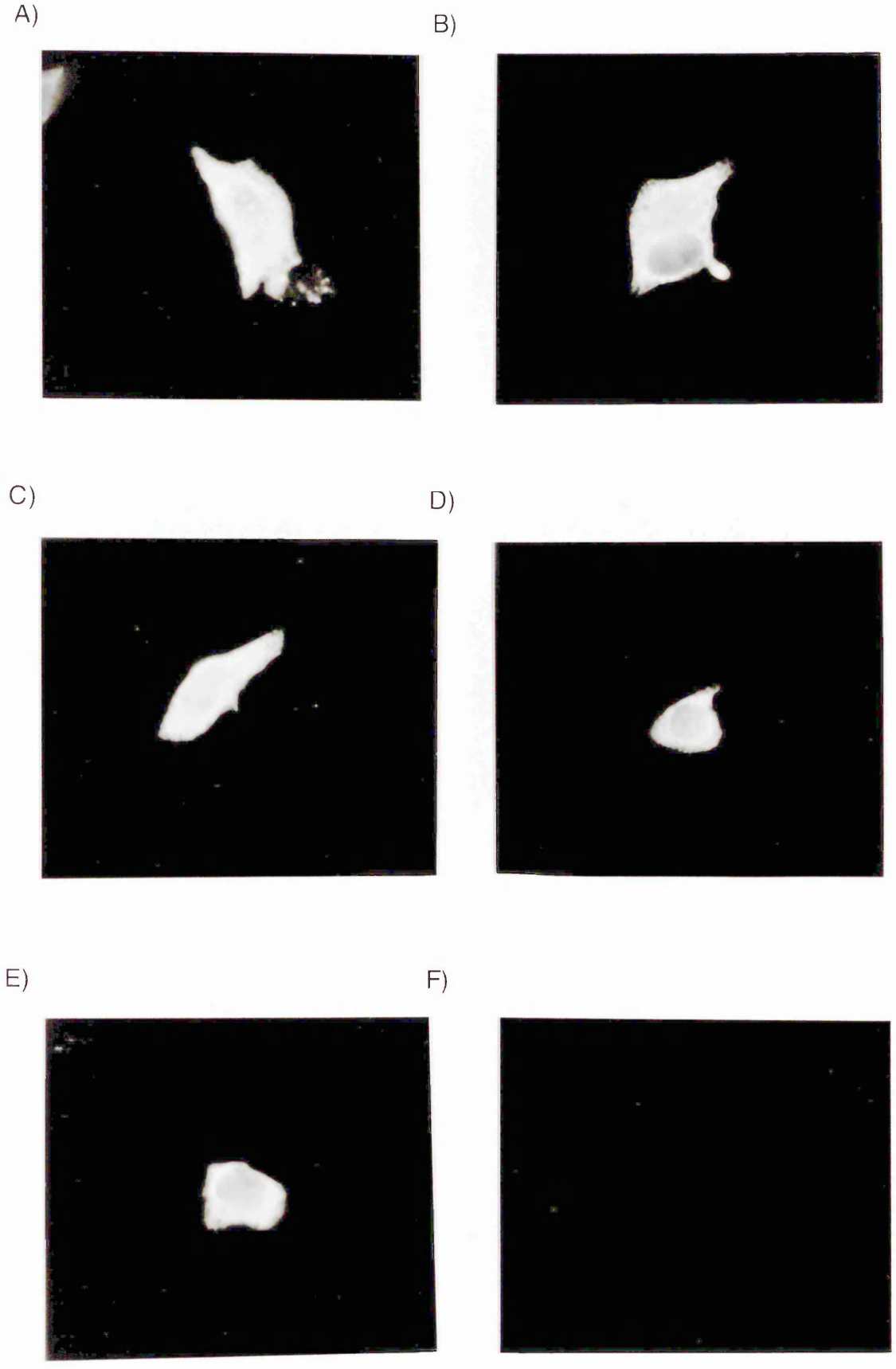


Figure 3J3: Immunofluorescent labelling of cell-cycle arrested HepG2 cells transfected with pp65 epitope-tagged core genes (continued on following pages)

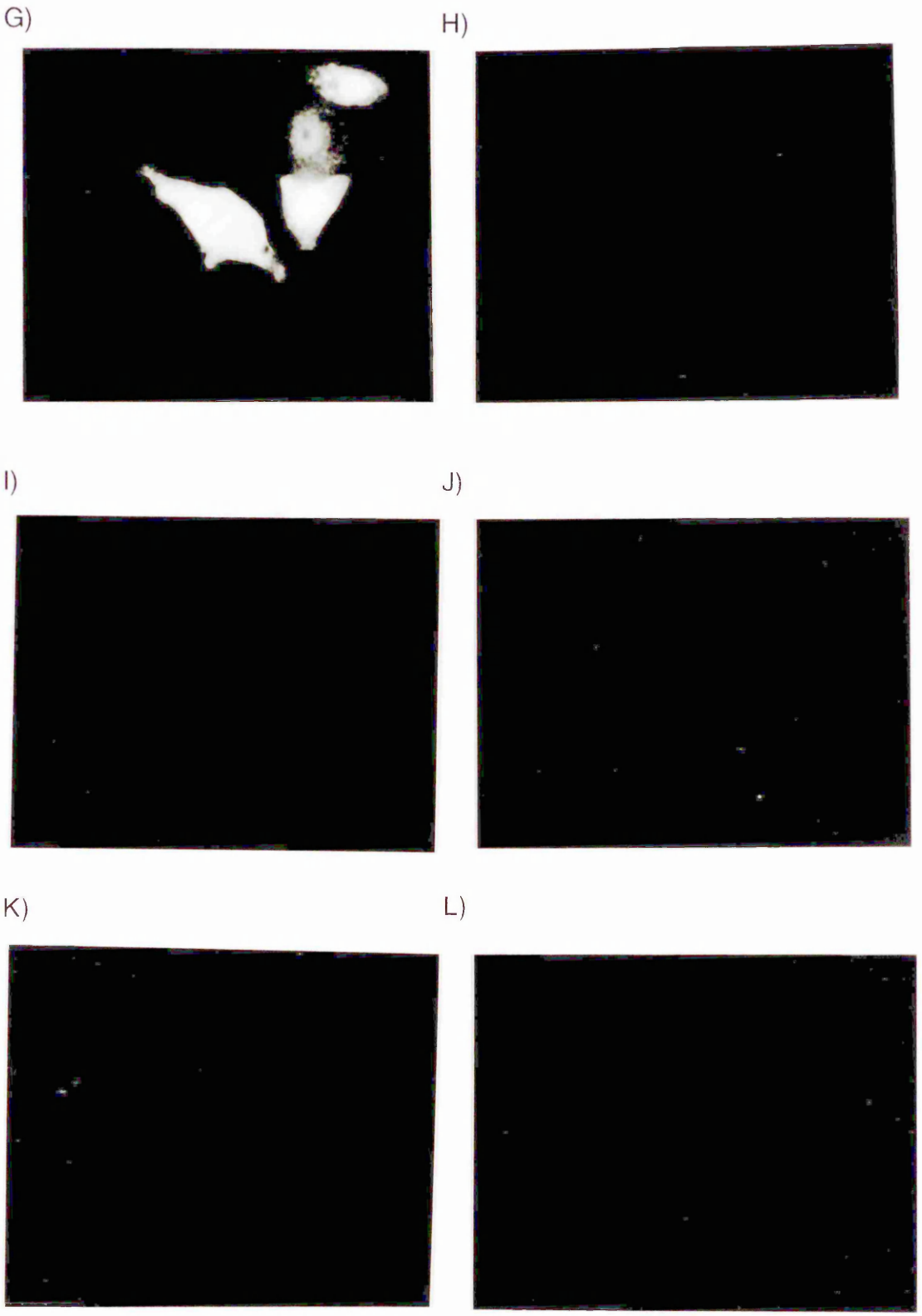


Figure 3J3: Immunofluorescent labelling of cell-cycle arrested HepG2 cells transfected with pp65 epitope-tagged core genes (continued with legend on the following page)

M)

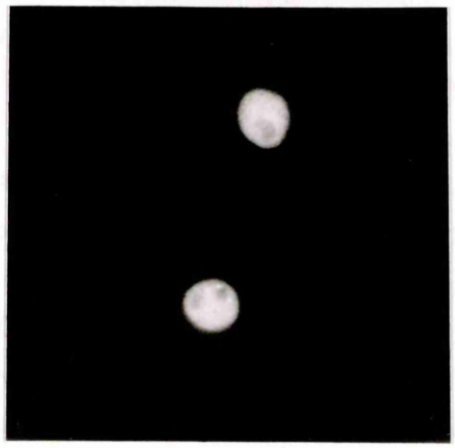


Figure 3J3: Immunofluorescent labelling of cell-cycle arrested HepG2 cells transfected with pp65 epitope-tagged core genes

HepG2 cells arrested at the G_1/S boundary in the presence of aphidicolin were transfected with plasmid pRK5 containing either full-length core gene with a C-terminal epitope tag, epitope-tagged core deletion genes or no insert. Following transfection, cells were incubated with either a 1:300 dilution of anti-pp65 mAb 9220 or undiluted Zymed polyclonal antibody. After washing, cells were incubated with either a 1:100 dilution of FITC-conjugated goat anti-mouse antibody, or a 1:60 dilution of goat anti-rabbit FITC-conjugated antibody to enable detection of the pp65 sequence and core protein, respectively.

A) Full-length core protein - mAb 9220
B) Δ84-109 - mAb 9220
C) Δ81-121 - mAb 9220
D) Δ79-125 - mAb 9220
E) Δ60-117 - mAb 9220
F) pRK5 - mAb 9220
G) Full-length core protein - Zymed pAb
H) Δ84-109 - Zymed pAb
I) Δ81-121 - Zymed pAb
J) Δ79-125 - Zymed pAb
K) Δ60-117 - Zymed pAb
L) pRK5 - Zymed pAb
M) Untagged full-length core protein - Zymed pAb

signal of the core protein is located within this C-terminal region of the protein, the Gly153Cys mutation may have affected the localisation of this protein.

The core deletion proteins were all detected effectively using mAb 9220 (Panels B-E), but none were detected with the anti-core polyclonal antibody (Panels H-K). The localisation of all four core deletion proteins was found to be similar, with protein present in both the cytoplasm and the nucleus of the cells, in a manner similar to that observed for the epitope-tagged full-length protein used in these experiments. The C-terminal amino acid sequences of these proteins were examined for mutations, but none were detected. It was concluded that the core deletion proteins did not display the same subcellular localisations as wild-type core protein. Rather, the proteins showed a distribution that was both nuclear and cytoplasmic, even when cells were arrested at the G_1/S boundary by incubation with aphidicolin. Cells transfected with plasmid alone showed minimal fluorescence with either antibody (Panels F and L). Panel M shows the nuclear fluorescence observed when untagged full-length core protein with no C-terminal amino acid mutations was expressed (this plasmid construct was kindly supplied by Ed Dornan - Institute of Virology). The core protein was detected using the Zymed polyclonal antibody.

3J4 Discussion

The results from this section have shown that the core deletion proteins showed a subcellular distribution that was altered from that of full-length wild-type core protein. At the G_1/S boundary of the cell cycle, all the deletion proteins were located in the cytoplasm in addition to the nucleus. This altered distribution may have been due to the presence of the C-terminal epitope tag sequences, which could interfere with the normal function of the core protein nuclear localisation signal. The more likely possibility was this signal was affected by the deletion of the upstream core protein residues, or by the new amino acid sequence formed as a result of the deletion. From the results of previous sections, where the core deletion proteins were shown to be incapable of forming as core particles, it may have been expected that they would not display other properties which characterise the full-length protein.

The epitope-tagged full-length core protein studied here, again in contrast to published reports, did not localise exclusively to the nucleus, but displayed a distribution throughout the nucleus and cytoplasm of the cell, similar to that observed for the core deletion proteins. This was thought to be due to a Gly153Cys mutation in the C-terminal region of the protein. A previous mutation (Arg151Thr) in this region was shown to inactivate the core protein nuclear localisation signal (Eckhardt *et al.*, 1991).

The implications of core protein C-terminal mutations and the related alterations in subcellular localisation of the proteins are unclear at present. Recent results from our laboratory have suggested a correlation between core protein localisation in the cytoplasm

and ongoing disease (Dornan *et al.*, 1996). Sequential serum samples from e antigen positive patients, who subsequently seroconverted to anti-HBe antibody, and patients who were continually anti-HBe positive were analysed. The core gene was amplified by PCR (using an antisense primer which did not include the nucleotide sequences encoding the pp65 epitope tag) and cloned into pRK5. Following transfection of HepG2 cells the location of the core protein was analysed in a manner identical to that used in this section, using the Zymed polyclonal antibody. Patients who seroconverted to anti-HBe antibody and went into clinical remission showed no change in the localisation of the core protein between sequential serum samples: all samples showed strong nuclear fluorescence for core protein. However, patients with ongoing disease after seroconversion to anti-HBe antibody showed increased cytoplasmic core protein in samples taken after seroconversion. Increased cytoplasmic core protein expression was also detected in later sequential samples from patients who were continually anti-HBe positive, but remained viremic and had episodic reactivations of hepatitis.

As the localisation of the core deletion proteins is both nuclear and cytoplasmic (for the particular samples tested), the disease observed in the patient from which these variants were isolated may be caused by a similar mechanism to that observed in the patients described above. Work is ongoing to attempt to further determine the possible details of such a mechanism.

3K Sequence analysis of co-existing full-length core genes

3K1 Introduction

For all but two cases where the existence of core gene deletions has been described, investigators have also been able to detect full-length core gene in the same samples. The relative proportions of the two populations has not been well-documented, but greater attention has been focused on the core deletion genes. The rationale behind this was that as the regions deleted in these genes corresponded approximately to antigenic regions of the protein, it was hypothesised that these deletion variants may have represented a form of immune escape from one or more arms of the immune response. By allowing the virus to persist in this way, these deletion variants may have contributed to the chronic nature of the disease which affected these patients. However this attention may have been wrongly focused, and it may have been that the co-existing virus was in fact responsible for the disease state, in a manner similar to chronic patients without core gene deletions.

With this hypothesis in view it was decided to isolate and sequence full-length core genes, including the pre-core region, from the serum of patients with core gene deletions. This was made possible by the very kind gift of an aliquot of serum from Dr. N. Naoumov, from the Kings College group in London. This serum came from an e antigen-positive Chinese patient with CAH, from whom the core gene deletions Δ81-121, Δ79-125 and Δ60-117 had been isolated. After extraction of DNA from the serum, the pre-core/core region was amplified by PCR using primer pair C1/C2. When the products of this reaction were analysed by agarose gel electrophoresis, a broad band at approximately 600bp could be detected, representing core gene deletions, with only a fainter smear visible at approximately 730bp, corresponding to full-length core gene. However, when both bands were excised, gene-cleaned and cloned into the PCR-cloning vector pT7-Blue, no full-length genes could be isolated from recombinant plasmids. It was then decided to gene-clean total PCR products and make them blunt-ended using Klenow polymerase. The products could then be cloned into *Sma* I-digested plasmid pTZ18R, full-length genes being identified by *Eco* RI/*Hin* dIII digests of recombinant plasmid mini-prep DNA which were subsequently sequenced.

3K2 Results

A total of 84 pre-core/core genes were analysed following cloning of the PCR products as described above. Restriction enzyme digest fragments that were smaller in size than approximately 800bp, as determined by agarose gel electrophoresis, were excluded from further study. The full-length genes of the remaining 33 clones were then completely sequenced. Upon closer sequence analysis, it was found that only 10 of the 33 'full-length' genes actually represented complete genes. The remaining genes all had small

		Core amino acid position															
Clone	Pre-core	5	26	31	45	46	52	57	60	67	85	130	147	155	157	158	182
4	-		S→A							N→T	V→I	P→T	T→A	S→P		R→G	Q→*
4a	-		S→A		P→R	E→G	H→R				V→I	P→T	T→A		R→G	R→G	
6	-		S→A						L→V								
26	-		S→A						L→V								
5	W28*	P→T	S→A						L→V		V→I						Q→*
8	W28*	P→T		L>I					L→V			P→T	T→A	S→P			Q→*
16	M1F W28*	P→T	S→A								V→I		T→A	S→P			Q→*
18	W28*								L→V								
20	W28*							Q→K	L→V								
43	W28*								L→V				T→A	S→P			Q→*

* = stop codon formation

Table 3K: Sequence analysis of full-length pre-core/core genes

A total population of pre-core/core genes was amplified from the serum of a patient with HBV core gene deletions. After blunt-ending and cloning of PCR products, full-length genes were identified by restriction enzyme digestion and agarose gel electrophoresis and sequenced. After translation, full-length sequences were compared to consensus subtype sequences (*adr*, *adw*, *ayr* and *ayw*) to identify the position and nature of any novel mutations.

deletions from 29-38bp between nucleotides 385 and 431 (inclusive) of the core gene. Among these clones, the predominant deletion was a 37bp deletion of nucleotides 387-423, inclusive. Mutations observed in the amino acid sequences derived from the full-length genes are shown in Table 3K.

From these results, the sequences can be separated into two groups: sequences with a pre-core stop mutation and those without. However there was no significant difference in the number of mutations between these two groups (χ^2 test; $P<0.1$). In addition several mutations were common to both groups: S26A, L60V, V85I, P130T, T147A, S155P and Q182Stop. Frequent mutations were observed which were specific to either group, such as R158G and P5T, in addition to sporadic mutations at other positions. The observed mutations do not appear to cluster in any of the regions suggested by other groups (see Section 1D2), with the exception perhaps of the region between amino acids 130 and 156 suggested by Ehata *et al.* (1992, 1993). Of the 46 mutations observed, 12 (26%) are located in this region, but as no functional or immunological significance it has been assigned to this portion of the protein the relevance of this is unclear. These results also contrast with those of Akarca and Lok (1995) who observed an absence of mutations in the C-terminus of the protein which overlapped with the polymerase reading frame.

It was of interest that no mutations occurred within the region between amino acids 86-129, as this contains the hydrophobic heptad repeat motif required for correct particle assembly. Cysteine residues were also unaltered, with the exception of Cys-183, which was missing in clones 4, 5, 8, 16 and 43 due to the creation of a stop codon at residue 182. However, as these residues are not essential for particle assembly their absence is not significant.

3K3 Discussion

The results presented in this section have attempted to analyse mutations occurring in full-length core genes from a patient with co-existing core gene deletions. However the low number of full-length clones isolated has made this type of analysis less effective than was possible; the remaining 23 clones all having small regions deleted. This does show that, at this time-point, viruses with core gene deletions are present at greater levels than those with full-length core genes (10 full-length:73 deletions). Further longitudinal analysis of this particular patient by the group at Kings College has shown that after seroconversion, viruses with core gene deletions were lost from the serum, in common with other patients having deletions (Patient WT; Marinos *et al.*, 1996).

As the significance of mutations in the core gene remains undetermined, it is difficult to comment on the absence of any mutational clustering in the sequences from this section. It may be that examination of a greater number of clones would shed light on this and maybe reveal significant regions.

In conclusion, the presence of full-length core genes containing no mutations which would be likely to disrupt the functional role played by the protein, provides the means for the propagation of viral genomes with core gene deletions. Further analysis of the relevance of core mutations, particularly in relation to the immune response, is required before their significance is determined.

CHAPTER 4 DISCUSSION

Hepatitis B viruses with deletions in the core gene can represent a large population of viral variants observed during chronic infections. As the significance of many HBV variants remains unclear, the work presented in this thesis has attempted to identify and characterise any possible functional roles for core gene deletions in the life cycle of the virus or in the pathology of the disease.

4A Do core deletion proteins form core particles?

An important initial question regarding the relevance of core gene deletions was whether or not the resulting proteins could assemble into core particle structures. That core protein can tolerate substantial deletions of various regions as well as insertions of large heterologous sequences and still form core particles shows its tremendous structural flexibility. It was therefore hypothesised that proteins with central deletions would also retain this ability, especially those with smaller deletions.

The results from Sections 3C, 3D and 3H have shown that, under the conditions used, no core particle-like structures could be detected when core proteins containing naturally-occurring deletions were expressed and purified. Despite the fact that Western blot analysis of sucrose gradient fractions showed the presence of these proteins in fractions where particles would have been expected, none were visible by electron microscopic examination. Two possibilities exist: (i) that no particles, only large aggregates of protein, were formed or (ii) that particulate structures did actually form but, despite all precautions taken, they were insufficiently stable to allow detection by electron microscopic examination. The latter possibility seems likely for the Δ81-121 5'fill-in proteins, where particles with encapsidated nucleic acid were also suggested by the agarose gel assay (see Section 3H), but not visible by electron microscopy.

Other experimental strategies could have been pursued to address this possibility had time allowed. A cross-linking agent such as glutaraldehyde, added to the protein solution immediately after the sucrose gradient fractionation stage, may have stabilised any particulate structures. Another possibility was to immunoprecipitate proteins from the bacterial extract, but problems may have arisen as major core protein epitopes were deleted (see Section 3B). It may have been possible to use the anti-pp65 mAb 9220 for these experiments, but unfortunately this antibody/epitope tag combination does not work well in such experiments (John McLauchlan, MRC Virology Unit - personal communication).

The most obvious reason for the failure of particle formation is that the missing regions completely disrupt the tertiary structure of the protein, making it impossible for the protein dimers to form particles. The observation that the core deletion proteins retain

the ability to dimerise (see Section 3D3) does not guarantee that these dimers are in the correct conformation for the subsequent core particle assembly: the formation of disulphide bonds may only require cysteine residues to be in close proximity and exposed at the surface of the folded protein.

It is also now obvious that the central region of core protein does not form a separate, external, globular domain, which could have been removed with minimal disturbance to the remainder of the polypeptide chain. Even the smallest regions, which were deleted in the Δ81-121 5'F30 and 3'F30 proteins, do not seem to form separate domains. These particular regions must therefore be located within the main domain of the protein. This has now been shown experimentally by the structural studies of core particles published during the production of this thesis (Conway *et al.*, 1997; Böttcher *et al.*, 1997; see Section 1E5). These studies showed that four α-helices were formed by the polypeptide backbone of the core protein. Comparison of the polypeptide chain conformation obtained from these results with the locations of the natural core protein deletions studied in this thesis showed that the deletions removed at least the entire third helix (Δ84-109) and, in the case of Δ60-117, large sections of helices 2 and 4 were removed as well. As helices 2 and 3 form part of the novel four-helix bundle found in the core particle, central deletions destroy this structurally important region of the protein.

These structural studies also provide a possible explanation for the inability of the Δcore-surface hybrid protein to form stable particles (see Section 3E2). Examination of the surface protein sequence used to replace core residues 81-121 showed that this region contained seven proline residues. It was previously thought that this region of the protein contained mainly β-sheets (Argos and Fuller, 1988) and so would not be disrupted by the presence of these residues. However, their presence would certainly disrupt the third and fourth α-helices now known to be present in this region of the protein. These two helices are also affected by the two smallest deletions created, Δ81-121 5'F30 and 3'F30, with the N-terminal third of helix 3 and the N-terminal half of helix 4 removed, respectively. Disruption of these regions of secondary structure may be the primary factor in prohibiting particle assembly. The deletion reported by Valliammai *et al.* (1995) removed amino acids 41-51 which contain only a region of unstructured polypeptide and the N-terminal residue of helix 2 (according to the structure from Böttcher *et al.* (1997)). As the hydrophobic heptad repeat (see below) is also undisturbed by this deletion, this particular deletion protein may well be assembly-competent.

It may be of interest to create deletions which reduce the length of helices 2 and 3 equally, by removing residues 75-85 for example and determine whether this would allow particle assembly. Deletion of this region should only reduce the length of the spike on the external surface of the particle and therefore may have an increased probability of assembling altered core particles. A logical mutagenesis approach such as this, based upon knowledge of the structure of the core protein, may have a greater chance of

defining the residues required for particle assembly, rather than the more random manner in which the deletions were created in this thesis.

The identification of the hydrophobic heptad repeat motif in core protein (Yu *et al.*, 1996), published toward the end of the work in this thesis, provided an additional explanation for the inability of core deletion proteins to form particles. This important motif for particle assembly was disrupted by all four natural deletions, with at least two motif residues missing or, in the case of Δ79-125, all four residues. When viewed in the light of the importance of this motif, such gross alterations as the observed deletions could only be expected to prohibit particle assembly. The two smallest created deletions, Δ81-121 5'F30 and 3'F30, removed one or none of the motif residues, respectively, and would therefore not have interfered significantly with the function of this motif. The more general disruption to the tertiary structure of the core protein described above would however, still explain the absence of core particle assembly when these proteins were expressed (see Section 3E3).

If further work was to be carried out to address the ability of core deletion proteins to form particles, a change of experimental strategy would probably allow more definitive results to be generated. Although core particles can be assembled by expression of the core gene in heterologous systems, comparisons of the efficiency of this process and the equivalent process occurring during the infectious cycle have not been carried out. It is possible that the latter system operates at a higher level of efficiency due to the presence of additional interactions. For example, interactions between the core protein and the full-length pgRNA may catalyse the assembly of the core particle during the encapsidation process and perhaps act to stabilise its final structure. Although Birnbaum and Nassal (1990) showed that core particles assembled in *E.coli* could encapsidate RNA, the major RNA species present within the particles was only approximately 0.9kb, in contrast to the 3.5kb pgRNA. The core protein may also interact with the polymerase protein, bound to the ε sequence. This interaction may initiate the core particle assembly process. As these potential interactions are not all necessarily present in the bacterial systems commonly used, opting for an experimental system which closely mimics events occurring during infection would help increase the authenticity of the results obtained. Amplification of viral DNA from serum using the PCR method developed by Günther *et al.* (1995) would allow the isolation of entire HBV genomes containing core gene deletions. Individual variant genomes could then be cloned and transfected into hepatocyte cell lines, in the absence of any other viral genomes, to allow viral DNA replication and virion production to occur. The advantage of this type of system is that all the cellular and viral components required in the lifecycle, subsequent to viral entry and uncoating, are present. If functionally normal core particles were produced from genomes containing in-frame core gene deletions, any resulting extracellular virions could be purified from the culture medium and the nucleic acid analysed to confirm the presence of the deletion. However,

if transfection of these genomes failed to produce virions, intracellular contents could be examined for the presence of core particles, thus identifying whether core particle assembly or core particle envelopment was the end-point in virion assembly for these variants.

4B Are core deletion proteins incorporated into mixed particles?

As viruses with core gene deletions co-exist with wild-type virus in the majority of reported cases, the core deletion proteins may co-assemble with the full-length core protein to form altered core particles. This assumes that both types of virus are present in the same hepatocyte, which has never formally been proven. However, the results discussed above coupled with the detection of these variants in the serum strongly suggest this is realistic.

As core particles can be composed of 90 or 120 dimer subunits, a certain degree of flexibility is probably permitted in the overall particle structure. For example, the presence of a small number of 'altered' subunits may be permitted, with the wild-type core protein dimers acting as a scaffold, maintaining the structure of the particle. As the N- and C-terminal regions of the protein form the shell of the particle, perhaps only these would be required in altered subunits, removing only the spike from the particle surface. Thus, it may be possible to include a significant proportion of altered subunits whilst still maintaining the structural integrity of the particle.

From the results of Section 3F, it appears that this scenario may exist for the two naturally-occurring core deletion proteins studied, Δ84-109 and Δ 81-121. However for a definitive conclusion to be reached, a more direct experimental approach should be employed, such as the immunoelectron microscopy discussed in the results section. Comparison with the results of the 5'F30 protein co-expression experiment also stresses the need for confirmation of the previous results, as this deletion is included in the Δ81-121 region yet does not appear to assemble into mixed particles. The extra polypeptide sequence in this protein compared to the Δ81-121 protein may alter the conformation so that it cannot fit into the subunit packing arrangement of the particle.

If the results for the naturally-occurring deletions are correct, the core particle would appear to be capable of tolerating a substantial proportion of altered subunits in its structure. In these cases, the original helix 2 may fold back on itself, forming a shorter spike but still supplying a sufficient length of polypeptide chain to form the domain interacting with the particle shell. In an analogous manner, it may again be of interest to determine whether or not a core deletion protein with only the upper region of the spike removed would form mixed particles. It is also unknown if mixed particles can encapsidate nucleic acid and be enveloped correctly.

If such mixed particles do assemble, encapsidate viral pgRNA and are enveloped during infection, what would their relevance be? As discussed previously, core gene

deletion variants have been proposed to represent a viral subpopulation capable of escaping detection and elimination by the immune system. However, no mechanism has been proposed in any great detail to explain this hypothesis. Had core deletion proteins been able to assemble into particulate structures by themselves, these would have escaped recognition, or at best been poorly recognised, by B-cells. This would have reduced or eliminated the anti-core humoral response. The efficient initiation of a Th cell response would also have been affected in this scenario, as Th epitopes were either removed or disrupted by the deletions. In addition, a consequent decrease in the production of anti-core protein and anti-surface protein antibodies would result. As the anti-core Th response is held to be the primary factor in determining the outcome of infection, the deletions could cause a higher incidence of chronic disease.

However, from the results mentioned previously, the only possibility appears to be the formation of mixed core particles. The relevance of these in the course of virus infection is more difficult to determine. In this scenario, full-length core protein remains present and accessible to all branches of the immune system, making 'immune escape' highly unlikely. This does not necessarily mean that the viral deletion variants are incapable of influencing the outcome of infection. It may be that the relative strength of the immune response is important, in contrast to the generation of an all-or-nothing response. In the former case, mixed particles would contain fewer epitopes for antibody and Th recognition, thus reducing these responses and their consequent effects.

4C Do core deletion proteins show altered subcellular localisations?

From the results of the immunofluorescence assays on transfected HepG2 cells in Section 3E4, it appears that core deletion proteins do not share the same subcellular distribution as the wild-type protein. Whereas full-length wild-type protein is reported to localise to the nucleus in the G_1 phase of the cell cycle, all four core deletion proteins were detected in both the cytoplasmic and nuclear compartments in this phase. This apparent lack of cell-cycle regulation of core deletion proteins was not caused by disruption of the NLS regions, as no mutations were present in the C-terminus of the core deletion proteins. However, other explanations exist: the presence of the epitope tag sequences or the absence of the deleted residues could both affect the NLS function.

Full-length epitope-tagged core protein, used as a control, also showed both nuclear and cytoplasmic distribution under cell-cycle arrest conditions. This suggested that the epitope tag sequences may have prevented cell cycle regulation of the location of the protein, possibly by occluding NLS regions or phosphorylation sites. However, the particular full-length protein used in these experiments contained a Gly>Cys mutation at amino acid 153. Similar C-terminal point mutations within the NLS have been shown to be responsible for significantly altering the localisation pattern of the core protein (Ed Dornan, Division of Virology, University of Glasgow - unpublished results): core protein

with a Gln→Lys mutation at amino acid 169 localised preferentially in the cytoplasm. After this single mutation was corrected by site-directed mutagenesis, the protein localised to the nucleus when expressed under identical conditions. Therefore, it seems probable that the unexpected core protein distribution observed in the results obtained for the full-length protein was due to the Gly153Cys point mutation. In support of this conclusion, other full-length core proteins have been examined under identical conditions and have all been shown to localise to the cell nucleus (Ed Dornan - personal communication). The failure of the core deletion proteins to display the expected subcellular distribution is therefore likely to be due to the presence of the deletions.

A similar alteration in localisation caused by the deletion of residues separate from the NLS has previously been shown for the murine type IV c-*abl* gene product (Van Etten *et al.*, 1989). Removal of 53 amino acids in the N-terminal region of the protein, upstream from the NLS sequence, changed the localisation of the protein. The nuclear protein now located to the cytoplasm, with an apparent increased concentration at the cell membrane. In a separate experiment, insertion of a four amino acid linker at approximately the same position also resulted in this altered localisation. Deletion of the core protein residues upstream to the NLS could alter the tertiary structure of the protein so as to prevent protein kinases from phosphorylating target residues or inhibiting the interaction of the NLS with its nuclear targeting protein. As the precise role of the nuclear localisation of core protein and its regulation by the cell cycle are unclear, it is difficult to comment on the meaning of these results. Since the deletion proteins do not form particles they can play little, if any, role in genome delivery to the nucleus during infection, except perhaps in the form of mixed core particles.

4D Are core gene deletion variants an important factor in chronic hepatitis?

Large amounts of information have been generated from studies of the core protein and core particles over recent years. Studies in the contexts of such fields as immunology, structural biology and molecular biology, have allowed a more informed assessment of the relative importance of HBV core gene deletion variants to be made.

1) Do core gene deletion variants represent immune escape?

A hypothesis exists which proposes that deletion of core gene regions which encode epitope sequences would prevent or decrease recognition of the protein by the immune system. The possibility that the chronic disease observed in patients is a consequence of the presence of these variants has been proposed by a number of groups (Wakita *et al.*, 1991; Takayanagi *et al.*, 1993; Ackrill *et al.*, 1993; Fiordalisi *et al.*, 1994; Akarca and Lok, 1995b; Valliammai *et al.*, 1995; Zoulim *et al.*, 1996). However, other more thorough studies have cast significant doubt on this hypothesis.

Günther *et al.* (1995) detected core gene deletions in viral genomes cloned from a group of renal transplant patients. However, all patients from which deletion variants were isolated had been kept under continuous immunosuppressive treatment. Results from this study and a similar subsequent study (Günther *et al.*, 1996b) suggested that core gene deletions may be characteristic of this subgroup of HBV-infected patients. However, the continuous immunosuppressive therapy received by these patients makes immune-mediated selection an unlikely explanation for the emergence of core gene deletion variants. This possibility cannot be ruled out entirely, as some immune functions may have been retained by these patients.

A longitudinal analysis of changes in core gene deletion variants was carried out by Marinos *et al.* (1996) to determine the effects of both an enhanced immune response, in patients undergoing IFNα treatment, and seroconversion to anti-HBe antibody. The study showed that viruses containing core gene deletions were preferentially eliminated after either IFNα-induced or spontaneous seroconversion to anti-HBe. This failure to persist under enhanced immune pressure contrasts with the results obtained for viruses with full-length core genes and makes it unlikely that these deletion variants remain undetected by the immune system.

Another possibility does exist where core gene deletion variants need not completely evade the immune response, but could instead inhibit core-specific T cell responses in a manner analogous to that described by Bertoletti *et al.* (1994a, b). Rather than involving a wild-type T cell epitope and a variant epitope containing point mutations, the novel peptides formed as a result of juxtaposition of sequences external to the deletion may also act as inhibitors of the T-cell responses evoked by the wild-type peptide. Alternatively, competition for HLA-binding could occur between novel and wild-type peptides, or the novel peptides could render T cells anergic. However, examination of the novel sequences formed by the deletions shows they display no homology to T cell epitopes identified in core protein. Nevertheless, until tested experimentally, using a system similar to that of Bertoletti *et al.*, this scenario cannot be discounted.

2) Are core gene deletion variants responsible for increased disease severity?

As discussed previously (Section 1E6) in the majority of studies where viruses containing core gene deletions have been detected, patients have suffered from CAH/severe liver damage as opposed to being diagnosed as CPH or ASC cases. This correlation does suggest a role, as yet undefined, for this type of variant in augmenting disease severity.

An alternative explanation for this link is that, rather than representing immune escape, core deletion proteins may be directly cytopathic to the hepatocytes in which they are expressed. Core protein has previously been associated with such an effect when

expressed in a HepG2 clone produced by transfection with cloned circular HBV DNA (Roingeard *et al.*, 1990). This clone did not produce extracellular virions, despite expressing high levels of surface proteins and e antigen. This clone also displayed a high level of cell death, in contrast to an identical HBV-producer clone. The reason for this cytopathic effect was found to be an accumulation of core protein within the cells.

A similar scenario may exist for core deletion proteins: accumulation of the protein may occur, producing a cytopathic effect. Alternatively, the core deletion proteins could interfere with particle assembly from full-length protein, leading to an accumulation of protein due to its inability to be enveloped and secreted from the cell. Günther *et al.* (1996b) showed that viruses with core gene deletions could persist and even increase with time, presumably leading to increased production and accumulation of core deletion proteins with the concomitant cytopathicity. However, such proteins cannot be highly cytotoxic as they are able to be transiently expressed in bacterial and mammalian cells with no visual evidence of cell death or growth inhibition. In addition, Günther *et al.* observed only a limited increase in serum markers indicative of hepatocyte destruction, again suggesting only a weak cytopathic effect.

From these observations the likelihood of a link between the presence of core deletion proteins and disease severity is possible. However, it is unlikely to fully explain the increased severity observed in these patients. In addition, other studies of disease state in patients with core gene deletions have found no correlation (Okamoto *et al.*, 1987b; Akarca and Lok, 1995b).

3) Do core gene deletion variants represent defective interfering particles?

Defective interfering particles (DIPs) are replication-defective viral deletion mutants and are associated with most RNA viruses and some DNA viruses (Holland, 1990). DIPs can interfere with the growth of homologous wild-type virus by a number of possible mechanisms and it has been suggested that they may play a role in several viral infections (Holland, 1990). The presence of DIPs may affect viral replication due to the modification of viral proteins in the DI genome (dePolo *et al.*, 1987); by alteration of the host immune response (Chattopadhyay *et al.*, 1989; Morgan and Dimmock, 1992); or by alteration of the expression of viral genes due to novel regulatory proteins being encoded by the DI genome (Saïb *et al.*, 1993). It is possible that hepadnaviral deletion variants can act as DIPs, affecting the wild-type virus with which they co-exist. Experiments involving co-transfection of full-length DHBV genomes and defective genomes into Huh7 cells have shown that the presence of the defective genome can suppress the production of wild-type virions (Horwich *et al.*, 1990). Defective genomes encoding either C-terminally truncated polymerase and surface proteins or C-terminally truncated polymerase, surface and core proteins were able to reduce the production of virions from transfected cells, after co-transfection. This reduction was approximately thirty-fold for

the latter defective genome and was thought to be due to disruption of the virion assembly process by the truncated proteins. It is important that similar experiments are performed for the naturally-occurring HBV deletion variants in order to assess the likelihood that they produce DIPs.

The possible effects that core deletion proteins have on the immune responses occurring during HBV infection also need to be examined. Core protein can inhibit the production of IFNβ (see Section 1C3) and has also recently been implicated in the decreased expression of the type I interferon-inducible MxA protein observed in Huh7 cells stably transfected with a defective 2.2kb viral genome (Rosmorduc *et al.*, 1996). Core deletion proteins may have increased inhibitory effects on the expression of these important immune system proteins or may reduce the responsiveness of infected cells to the host immune response. These alterations may be an important factor in determining the outcome of HBV infection.

4) Do core gene deletions have effects on other viral genes?

The possibility exists that the deletions originally attributed to the core gene actually primarily affect other viral genes. This seems a likely explanation, especially as the deletions cluster around the region encoding the core protein hydrophobic heptad repeat motif, which is essential to the function of the core protein. However, the central clustering of the deletions avoids the polymerase gene. Were the core gene deletions to extend much farther in the 3' direction, 5' regions of the polymerase gene would also be removed. Instead, studies of two different patient groups have proposed a different explanation for the observed clustering. This involves the J- and C2-AUG codons (nucleotides 2163-2165 and 2177-2179, respectively), located 5' to the authentic polymerase AUG codon as described by Fouillot *et al.* (1993).

Deletions clustering in the central region of the core gene remove sequences encoding both of these AUG codons, but do not affect the authentic polymerase start codon. Deletion analysis of this region showed that removing the J and C2 codons allowed increased polymerase translation from the downstream ORF. Therefore, in viruses with core gene deletions which remove this region, increased polymerase expression probably occurs. Due to the requirement for polymerase in the encapsidation process, its predicted increased expression in these viruses may explain the high viremia, the preferential encapsidation and accumulation of this type of viral variant in renal transplant patients (Günther *et al.*, 1995, 1996b). This enhanced expression also provides a possible explanation for a selection process for these variants that is independent of immune system involvement.

However, Marinos *et al.* (1996) found that patients with core gene deletions had only low levels of viremia, in contrast to the results from the renal transplant patients. This was despite the fact that all the core gene deletions had also removed the J- and C2-

AUG codons. In this case the core deletion proteins were proposed to interfere with core particle assembly by full-length core protein dimers. This may have occurred by the deletion proteins preventing dimerisation of the full-length protein or preventing full-length protein dimers from multimerising. Due to the *cis*-preferential nature of core protein dimerisation, full-length core protein:core deletion protein dimers are unlikely to form, and results from this thesis suggest that the latter possibility may also be unable to explain the prevention of particle assembly. Direct interactions between the polymerase and core deletion proteins, leading to the inhibition of polymerase function, may instead explain the reduced DNA levels. Such interactions may also interfere with the encapsidation process, causing further reductions.

Increased presentation of polymerase-specific T cell epitopes may be an additional consequence of higher polymerase expression caused by core gene deletions. As this protein contains several CTL epitopes, increased levels of CTL activity will occur with a resulting increase in hepatocyte lysis and liver damage. This secondary effect of the deletions may alone explain the more severe form of disease suffered by these patients, and also the preferential elimination of cells harbouring these variants.

5) Is the co-existing wild-type virus population responsible for the disease state?

Perhaps the most unsatisfactory answer to the question concerning the relevance of core gene deletion viruses is that they have no significant role to play in causing the disease state observed. This is so unsatisfactory because it consigns this class of variant to the same group as all other core gene variants: it is possible to isolate and describe them, but their significance remains unknown, as does any possible mechanism linking them to chronic hepatitis.

Sequencing of the limited number of full-length genomes revealed no promising clues to explain the disease state. It is clear that more detailed studies of this subpopulation are required. Such studies should include HLA-typing, in order to refine the analysis of mutations; studies of mutations in relevant epitopes for immunological differences; assessment of core particle assembly by individual variants; an assessment of the stability of particles formed; and quantification of the efficiency of pgRNA encapsidation and envelopment by these same variants.

4E Mechanism of generation of core gene deletions

Despite the many reported observations of deletions in the HBV core gene, no mechanism for their generation has been discovered. At least two mechanisms are possible: splicing of pgRNA and, by analogy to the generation of deletions in a retroviral vector-based system (Pathak and Temin, 1990; Pulsinelli and Temin, 1991), a mechanism involving errors in plus-strand viral DNA synthesis.

Splicing of the pgRNA has been discussed previously in Sections 1B3.1 and 1E6 (2). The intron boundaries contain the consensus 5' GT/C and 3' AG donor and acceptor sites. However, examination of the boundaries of the other observed core gene deletions shows that neither site is consistently present at these positions. Therefore, splicing can only explain the generation of HBV core gene deletion variants if non-consensus donor and acceptor sites can be recognised by the splicing machinery. As introns with highly conserved, but non-consensus sequences at both splice sites are present in genes from a variety of species (Tarn and Steitz, 1997), splicing of pgRNA may be the mechanism whereby deletions in the core gene are generated. However, if this is the case, deletions might be expected to occur throughout the genome, which has not been observed (Ackrill *et al.*, 1993).

The mechanism of HBV DNA replication may provide an alternative explanation for the generation of the core gene deletion variants observed. As reverse transcriptase is an error-prone enzyme, base-pair substitutions, insertions and deletions can all occur during a single round of replication (Roberts *et al.*, 1989; Pathak and Temin; 1990). Detailed characterisation of deletions and other mutations which occurred in a spleen necrosis virus-based vector allowed the elucidation of the mechanisms whereby these mutations were generated (Pulsinelli and Temin, 1991). Two classes of deletion mutant were generated by a mechanism which may be applicable to the generation of HBV core gene deletion variants. The first class of retroviral-based vector deletions arose due to a misalignment of the 3' terminus of the plus-strand strong stop DNA as it was being extended by the reverse transcriptase (see Figure 4A, steps 9 and 10). Sequences at the 3' terminus of this strand hybridised to sequences which were homologous to those being copied, but which were located further downstream in the minus-strand DNA. As extension of the plus-strand DNA continued at this new site, a stretch of intervening minus-strand nucleotide sequence was deleted. In some cases the homology between the sequence being copied and the misalignment site was only two nucleotides. The second class of retroviral vector deletions arose due to late termination of the strong stop plus-strand DNA (see Figure 4A, step 7). This resulted in the inclusion of additional nucleotide sequence at the 3' terminus of the plus-strand DNA. This sequence was complementary to the tRNA molecule used as a primer for minus-strand DNA synthesis. Following transfer of the plus-strand DNA to the 3' terminus of the minus-strand DNA, the tRNA sequences at its 3' terminus could not hybridise to nucleotides adjacent to the primer binding site (PBS). Instead, these tRNA sequences hybridised to nucleotides further 3' than the PBS, with the deletion of intervening minus-strand DNA sequences.

The deletions described using this experimental system are of interest for two reasons: (i) they are generated during a replication cycle which is similar to that of HBV and (ii) the deletions are positioned in a region which is the genomic homologue of the core gene: 3' to the binding site for the primer of plus-strand DNA synthesis.

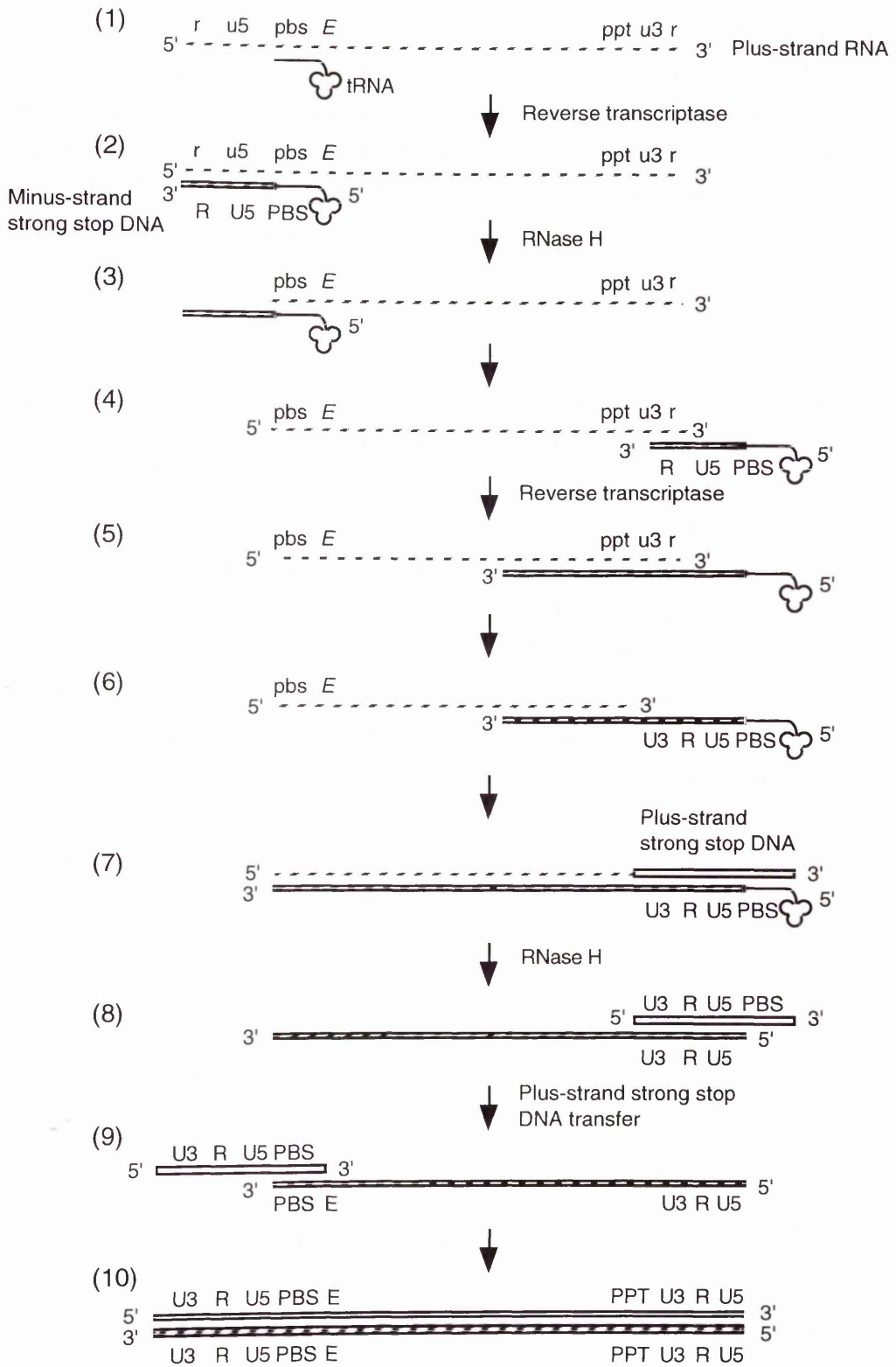


Figure 4A: Summary diagram of retroviral reverse transcription and DNA replication
(1) Retroviral plus-strand RNA (dashed line) is the template for the synthesis of minus-strand strong stop DNA (hatched rectangle). (2) This process occurs by reverse transcription and utilises a tRNA primer. (3) Removal of the bound RNA from the minus-

strand strong stop DNA by RNase H. (4) Minus-strand strong stop DNA transfer. (5) Minus-strand DNA synthesis by reverse transcriptase. (6) Nicking of RNA at the polypurine tract (*ppt*) and removal of 3' sequence by RNase H. (7) Minus-strand DNA synthesis continues and plus-strand strong stop DNA (clear rectangle) synthesis is initiated using the RNA primer. (8) Removal of RNA by RNase H. (9) Plus-strand strong stop DNA transfer. (10) Completion of minus-strand and plus-strand DNA synthesis. (Adapted from Pulsinelli and Temin, 1991)

However, the mechanism of generation of the second class of retroviral-based vector deletion mutants may not be relevant to the generation of the HBV core gene deletions. This is due to the differences that exist between the strategies for DNA replication employed by HBV and the retroviruses. In the second class of retroviral-based vector deletion mutants, the plus-strand strong stop DNA contained tRNA sequences at its 3' terminus. As a result, the DNA sequences 3' to the PBS sequence were missing in the final double-stranded retroviral DNA and the additional tRNA sequences included in their place. However, in HBV DNA replication, the inclusion of extra nucleotides at the 3' terminus of the primer for plus-strand DNA would be impossible. This is because the primer for synthesis of HBV minus-strand DNA is a protein molecule (pol/RTase), rather than a nucleic acid. As a result, extra nucleotides complementary to this primer cannot be added to the primer for plus-strand DNA synthesis. Also, no additional nucleotides are inserted in place of the deleted regions in the HBV core gene deletions. The locations of the deletions in this second class of retroviral-based vector deletion mutants also makes their mechanism of generation less relevant to the HBV core gene deletions. The retroviral deletions were immediately adjacent to the PBS sequence. This removed the E sequence, a positional and functional homologue of the ε sequence of HBV. As ε sequences are present in all HBV core gene deletions, a different mechanism must exist for their generation.

The mechanism of generation of the first class of retroviral-based vector deletions may be similar to that used to generate HBV core gene deletions: while the 3' terminus of the plus-strand DNA primer is being extended, it could misalign with downstream sequences similar to those being copied. This would delete the intervening regions of the core gene and leave the ε sequences intact. However, a closer examination of the sequences surrounding the 5' and 3' limits of the core gene deletions shows that this mechanism cannot explain the observed deletions. Even the two nucleotide homology between the 3' end of the plus-strand strong stop DNA and the location of the misalignment observed for a retroviral-based vector deletion was not present for the equivalent sequences in the HBV core gene deletions. Unless the site of misalignment of

the HBV plus-strand DNA primer is entirely random, with no requirement for nucleotide homology, a novel mechanism must exist for the generation of the HBV core gene deletions.

As chronic carriers of HBV represent a significant proportion of the population it is important that both the viral and host factor/s which determine the outcome of HBV infection are more fully investigated. Despite the number of reports of patients with core gene deletion variants, no investigations to date have examined possible functions they may have, or the effects they may have on the co-existing wild-type viral population. The results presented in this thesis have shown that the core deletion proteins encoded by these variants cannot functionally substitute for the wild-type protein. Therefore, further studies should concentrate on the effects of the interactions between the two types of core protein and any consequences these may have on the respective viral populations. The possibility that the deletions affect other genes, in particular the polymerase gene, also deserves further investigation.

REFERENCES

Ackrill, A.M., Naoumov, N.V., Eddleston, A.L.W.F. and Williams, R. (1993) Specific Deletions in the Hepatitis B Virus Core Open Reading Frame in Patients With Chronic Active Hepatitis B. J. Med. Virol., **41**, 165-169.

Akarca, U.S., Greene, S. and Lok, A.S.F. (1994) Detection of Precore Hepatitis B Virus Mutants in Asymptomatic HBsAg-positive Family Members. Hepatology, **19**, 1366-1370.

Akarca, U.S. and Lok, A.S.F. (1995a) Naturally Occurring Hepatitis B Virus Core Gene Mutations. Hepatology, **22**, 50-60.

Akarca, U.S. and Lok, A.S.F. (1995b) Naturally occurring core-gene-defective hepatitis B viruses. J. Gen. Virol., **76**, 1821-1826.

Alexander, G.J.M. and Eddleston, A.L.W.F. (1986) Does maternal antibody to core antigen prevent recognition of transplacental transmission of hepatitis-B-virus infection? Lancet, **i**, 296-297.

Ando, K., Moriyama, T., Guidotti, L., Wirth, S., Schreiber, R.D., Schlicht, H.J., Huang, S. and Chisari, F.V. (1993) Mechanisms of Class I Restricted Immunopathology. A Transgenic Mouse Model of Fulminant Hepatitis. J. Exp. Med., **178**, 1541-1554.

Antonucci, T.K. and Rutter, W.J. (1989) Hepatitis B Virus (HBV) Promoters Are Regulated by the HBV Enhancer in a Tissue-Specific Manner. J. Virol., **63**, 579-583.

Araki, K., Miyazaki, J-I., Hino, O., Tomita, N., Chisaka, O., Matsubara, K. and Yamamura, K-I. (1989) Expression and replication of hepatitis B virus genome in transgenic mice. Proc. Natl. Acad. Sci. USA, **86**, 207-211.

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

Ashton-Rickardt, P.G. and Murray, K. (1989a) Mutants of the Hepatitis B Virus Surface Antigen That Define Some Antigenically Essential Residues in the Immunodominant *a* Region. J. Med. Virol., **29**, 196-203.

Ashton-Rickardt, P.G. and Murray, K. (1989b) Mutants That Change the Immunological Subtype of Hepatitis B Virus Surface Antigen and Distinguish Between Antigenic and Immunogenic Determination. J. Med. Virol., **29**, 204-214.

Bancroft, W.H., Mundon, F.K. and Russell, P.K. (1972) Detection of Additional Antigenic Determinants of Hepatitis B Antigen. J. Immunol., **109**, 842-848.

Barnaba, V., Franco, A., Alberti, A., Balsano, C., Benvenuto, R. and Balsano, F. (1989) Recognition of hepatitis B virus envelope proteins by liver-infiltrating T lymphocytes in chronic hepatitis B virus infection. J. Immunol., **143**, 2650-2655.

Barnaba, V., Franco, A., Alberti, A., Benvenuto, R. and Balsano, F. (1990) Selective killing of hepatitis B envelope antigen-specific B cells by class I-restricted, exogenous antigen-specific T lymphocytes. Nature, **345**, 258-260.

Barnaba, V., Franco, A., Paroli, M., Benvenuto, R., de Petrillo, G., Burgio, V.L., Santilio, I., Balsano, C., Bonavita, M.S., Cappelli, G., Colizzi, V., Cutrona, G. and Ferrarini, M. (1994) Selective Expansion of Cytotoxic T Lymphocytes with a $CD4^{+}$ CD^{56+} Surface Phenotype and a Type 1 Profile of Cytokine Secretion in the Liver of Patients Chronically Infected with Hepatitis B Virus. J. Immunol., **152**, 3074-3087.

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

Bartenschlager, R. and Schaller, H. (1992) Hepadnaviral assembly is initiated by polymerase binding to the encapsidation signal in the viral RNA genome. EMBO J., **11**, 3413-3420.

Bartenschlager, R., Junker-Niepmann, M. and Schaller, H. (1990) The P Gene Product Of Hepatitis B Virus Is Required as a Structural Component for Genomic RNA Encapsidation. J. Virol., **64**, 5324-5332.

Bavand, M., Feitelson, M. and Laub, O. (1989) The Hepatitis B Virus-Associated Reverse Transcriptase Is Encoded by the Viral *pol* Gene. J. Virol., **63**, 1019-1021.

Bayer, M.E., Blumberg, B.S. and Werner, B. (1968) Particles associated with Australia Antigen in the Sera of Patients with Leukaemia, Down's Syndrome and Hepatitis. Nature, **218**, 1057-1059.

Beames, B. and Lanford, R.E. (1993) Carboxy-Terminal Truncations of the HBV Core Protein Affect Capsid Formation and the Apparent Size of Encapsidated HBV RNA. Virology, **194**, 597-607.

Beames, B. and Lanford, R.E. (1995) Insertions within the Hepatitis B Virus Capsid Protein Influence Capsid Formation and RNA Encapsidation. J. Virol., **69**, 6833-6838.

Beasley, R.P. and Hwang, L-Y. (1983) Postnatal Infectivity of Hepatitis B Surface Antigen-Carrier Mothers. J. Infect. Dis., **147**, 185-190.

Beesley, K.M., Francis, M.J., Clarke, B.E., Beesley, J.E., Dopping-Hepenstal, P.J.C., Clare, J.J., Brown, F. and Romanos, M.A. (1990) Expression in yeast of amino-terminal peptide fusions to hepatitis B core antigen and their immunological properties. Bio/Technology, **8**, 644-649.

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.

Benn, J. and Schneider, R.J. (1995) Hepatitis B virus HBx protein deregulates cell cycle checkpoint controls. Proc. Natl. Acad. Sci. USA, **92**, 11215-11219.

Benn, J., Su , F., Doria, M. and Schneider, R.J. (1996) Hepatitis B Virus HBx Protein Induces Transcription Factor AP-1 by Activation of Extracellular Signal-Regulated and c-Jun N-Terminal Mitogen-Activated Protein Kinases. J.Virol., **70**, 4978-4985.

Berting, A., Hahnen, J., Kröger, M. and Gerlich, W.H. (1995) Computer-Aided Studies on the Spatial Structure of the Small Hepatitis B Surface Protein. Intervirology, **38**, 8-15.

Bertoletti, A., Ferrari, C., Fiaccadori, F., Penna, A., Morgolskee, R., Schlicht, H.J., Fowler, P., Guilhot, S. and Chisari, F.V. (1991) HLA class I-restricted human cytotoxic T cells recognize endogenously synthesized hepatitis B virus nucleocapsid antigen. Proc. Natl. Acad. Sci. USA, **88**, 10445-10449.

Bertoletti, A., Chisari, F.V., Penna, A., Guilhot, S., Galati, L., Missale, G., Fowler, P., Schlicht, H-J., Vitiello, A., Chesnut, R.C., Fiaccadori, F. and Ferrari, C. (1993) Definition of a Minimal Cytotoxic T-Cell Epitope within the Hepatitis B Virus Nucleocapsid Protein. J. Virol., **67**, 2376-2380.

Bertoletti, A., Sette, A., Chisari, F.V., Penna, A., Levrero, M., De Carli, M., Fiaccadori, F. and Ferrari, C. (1994a) Natural variants of cytotoxic epitopes are T-cell receptor antagonists for antiviral cytotoxic T cells. Nature, **369**, 407-410.

Bertoletti, A., Constanzo, A., Chisari, F.V., Levrero, M., Artini, M., Sette, A., Penna, A., Giuberti, T., Fiaccadori, F. and Ferrari, C. (1994b) Cytotoxic T Lymphocyte Response to a Wild Type Hepatitis B Virus Epitope in Patients Chronically Infected by Variant Viruses Carrying Substitutions within the Epitope. J. Exp. Med., **80**, 933-943.

Bichko, V., Schödel, F., Nassal, M., Gren, E., Berzinsh, I., Borisova, G., Miska, S., Peterson, D.L., Gren, E., Pushko, P. and Will, H. (1993) Epitopes recognized by antibodies to denatured core protein of hepatitis B virus. Mol. Immunol., **30**, 221-231.

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

Bishop, D.H.L. (1992) Baculovirus expression vectors. Semin. in Virol., **3**, 253-264.

Blum, H.E., Liang, T.J., Galun, E. and Wands, J.R. (1991) Persistence of Hepatitis B Viral DNA After Serological Recovery from Hepatitis B Virus Infection. Hepatology, **14**, 56-63.

Blumberg, B.S., Alter, H.J. and Visnich, S. (1965) A 'New' Antigen in Leukemia Sera. J. Amer. Med. Assoc., **191**, 541-546.

Blumberg, B.S., Gerstley, B.J.S., Hungerford, D.A., London, W.T. and Sutnick, A.I. (1967) A Serum Antigen (Australia Antigen) in Down's Syndrome, Leukemia and Hepatitis. Ann. Intern. Med., **66**, 924-931.

Bock, C-T., Schwinn, S., Schröder, C.H., Velhagen, I. and Zentgraf, H. (1996) Localization of Hepatitis B Virus Core Protein and Viral DNA at the Nuclear Membrane. Virus Genes, **12**, 53-63.

Boner, W., Schlicht, H-J., Hanreider, K., Holmes, E.C. and Carman, W.F. (1995) Further Characterization of 2 Types of Precore Variant Hepatitis B Virus Isolates from Hong Kong. J. Infect. Dis., **171**, 1461-1467.

Borisova, G.P., Berzino, 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 the 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.

Brown, C.S., van Lent, J.W.M., Vlak, J.M. and Spaan, W.J.M. (1991) Assembly of Empty Capsids by Using Baculovirus Recombinants Expressing Human Parvovirus B19 Structural Proteins. J. Virol., **65**, 2702-2706.

Brown, P.M., Tagari, P., Rowan, K.R., Yu, V.L., O'Niell, G.P., Vader, D., Middaugh, C.R., Sanyal, G., Ford-Hutchinson, A.W. and Nicholson, D.W. (1995) Epitope-labeled Soluble Human Interleukin-5 (IL-5) Receptors: Affinity cross-link labeling, IL-5 binding, and biological activity. J. Biol. Chem., **270**, 29236-29243.

Bruce, S.A. and Murray, K. (1995) Mutations of Some Critical Amino Acid Residues in the Hepatitis B Virus Surface Antigen. J. Med. Virol., **46**, 157-161.

Brunetto, M.R., Giarin, M.M., Oliveri, F., Chiaberge, E. Baldi, M., Alfarano, A., Serra, A., Saraccco, G., Verme, G., Will, H. and Bonino, F. (1991) Wild-type and e antigen-minus hepatitis B viruses and course of chronic hepatitis. Proc. Natl. Acad. Sci. USA, **88**, 4186-4190.

Bruss, V. and Gerlich, W.H. (1988) Formation of Transmembraneous Hepatitis B e-Antigen by Cotranslational *in Vitro* Processing of the Viral Precore Protein. Virology, **163**, 268-275.

Bruss, V. and Ganem, 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 the envelope proteins in hepatitis B virus assembly. Proc. Natl. Acad. Sci. USA, **88**, 1059-1063.

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., 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.

Bulla, G.A. and Siddiqui, A. (1989) Negative Regulation of the Hepatitis B Virus Pre-S1 Promoter by Internal DNA Sequences. Virology, **170**, 252-260.

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

Carman, W.F., Jacyna, M.R., Hadziyannis, S., Karayiannis, P., McGarvey, M.J., Makris, A. and Thomas, H.C. (1989) Mutation preventing formation of hepatitis B e antigen in patients with chronic hepatitis B infection. Lancet, **ii**, 588-591.

Carman, W.F., Ferrao, M., Lok, A.S.F., Ma, O.C.K., Lai, C.L. and Thomas, H.C. (1992) Precore Sequence Variation in Chinese Isolates of Hepatitis B Virus. J. Inf. Dis., **165**, 127-133.

Carman, W.F. (1995a) Variation in the Core and X Genes of Hepatitis B Virus. Intervirology, **38**, 75-88.

Carman, W.F., Thursz, M., Hadziyannis, S., McIntyre, G., Colman, K., Gioustoz, A., Fattovich, G., Alberti, A. and Thomas, H.C. (1995b) Hepatitis B e antigen negative chronic active hepatitis: hepatitis B virus core mutations occur predominantly in known antigenic determinants. J. Viral Hepatitis, **2**, 77-84.

Cattaneo, R., Will, H., Hernandez, N. and Schaller, H. (1983) Signals regulating hepatitis B surface antigen transcription. Nature, **305**, 336-338.

Cattaneo, R., Will, H. and Schaller, H. (1984) Hepatitis B virus transcription in the infected liver. EMBO J., **3**, 2191-2196.

Chang, H-K., Wang, B-Y., Yuh, C-H., Wei, C-L. and Ting, L-P. (1989) A Liver-Specific Nuclear Factor Interacts with the Promoter Region of the Large Surface Protein of Human Hepatitis B Virus. Mol. Cell. Biol., **9**, 5189-5197.

Chang, C., Zhou, S., Ganem, D. and Standring, D.N. (1994) Phenotypic Mixing between Different Hepadnavirus Nucleocapsid Proteins Reveals C Protein Dimerization To Be *cis* Preferential. J. Virol., **68**, 5225-5231.

Chattopadhyay, S.K., Morse III, H.C., Makino, M., Ruscetti, S.K. and Hartley, J.W. (1989) Defective virus is associated with induction of murine retrovirus-induced immunodeficiency syndrome. Proc. Natl. Acad. Sci. USA, **86**, 3862-3866.

Chen, P-J., Chen, C-R., Sung, J-L. and Chen, D-S. (1989) Identification of a Doubly Spliced Viral Transcript Joining the Separated Domains for Putative Protease and Reverse Transcriptase of Hepatitis B Virus. J. Virol., **63**, 4165-4171.

Chen, M., Hieng, S., Qian, X., Costa, R., and Ou, J-H. (1994) Regulation of Hepatitis B Virus ENI Enhancer Activity by Hepatocyte-Enriched Transcription Factor HNF3. Virology, **205**, 127-132.

Chen, Y-C. J., Delbrook, K., Dealwis, C., Mimms, L., Mushahwar, I.K. and Mandecki, W. (1996) Discontinuous epitopes of the hepatitis B surface antigen derived from a filamentous phage peptide library. Proc. Natl. Acad. Sci. USA, **93**, 1997-2001.

Cheong , J., Yi, M., 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., Filippi, P., McLachlan, A., Milich, 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.

Chu, C-M. and Liaw, Y-F. (1987) Intrahepatic Distribution of Hepatitis B Surface and Core Antigens in Chronic Hepatitis B Virus Infection. Gastroenterology, **92**, 220-225.

Chu, C-M. and Liaw, Y-F. (1992) Immunohistological study of intrahepatic expression of hepatitis B core and E antigens in chronic type B hepatitis. J. Clin. Pathol., **45**, 791-795.

Chu, C-M., Karayiannis, P., Fowler, M.J.F., Monjardino, J., Liaw, Y-F. and Thomas, H.C. (1985) Natural History of Chronic Hepatitis B Virus infection in Taiwan: Studies of Hepatitis B Virus DNA in Serum. Hepatology, **5**, 431-434.

Chu, C-M., Yeh, C-T., Sheen, I-S. and Liaw, Y-F. (1995) Subcellular Localization of Hepatitis B Core Antigen in Relation to Hepatocyte Regeneration in Chronic Hepatitis B. Gastoenterology, **109**, 1926-1932.

Chuang, W-L., Omata, M., Ehata, T., Yokosuka, O., Ito, Y., Imazeki, F., Lu, S-N., Chang, W-Y. and Ohto, M. (1993) Precore Mutations and Core Clustering Mutations in Chronic Hepatitis B Virus Infection. Gastroenterology, **104**, 263-271.

Clark, J.M. (1988) Novel non-templated nucleotide addition reactions catalysed by procaryotic and eucaryotic DNA polymerases. Nucleic Acids Res., **16**, 9677-9686.

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

Colgrove, R., Simon, G. and Ganem, D. (1989) Transcriptional Activation of Homologous and Heterologous Genes by the Hepatitis B Virus X Gene Product in Cells Permissive for Viral Replication. J. Virol., **63**, 4019-4026.

Colucci, G., Beazer, Y., Cantaluppi, C. and Tackney, C. (1988) Identification of a major hepatitis B core antigen (HBcAg) determinant by using synthetic peptides and monoclonal antibodies. J. Immunol., **141**, 4376-4380.

Condreay, L.D. Wu, T-T., Aldrich, C.E., Delaney, M.A., Summers, J., Seeger, C. and Mason, W.S. (1992) Replication of DHBV Genomes with Mutations at the Sites of Initiation of Minus- and Plus-Strand DNA Synthesis. Virology, **188**, 208-216.

Conway, J.F., Cheng, N., Zlotnick, A., Wingfield, P.T., Stahl, S.J. and Steven, A.C. (1997) Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. Nature, **386**, 91-94.

Cote, P.J. and Gerin, J.L. (1983) Nonoverlapping Antigenic Sites of Woodchuck Hepatitis Virus Surface Antigen and Their Cross-Reactivity with Ground Squirrel Hepatitis Virus and Hepatitis B Virus Surface Antigens. J. Virol., **47**, 15-23.

Cross, J.C., Wen, P. and Rutter, W.J. (1993) Transactivation by hepatitis B virus X protein is promiscuous and dependent on mitogen-activated cellular serine/threonine kinases. 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.

Dane, D.S., Cameron, C.H. and Briggs, M. (1970) Virus-Like Particles in the Serum of Patients with Australia-Antigen-Associated Hepatitis. Lancet, **i**, 695-698.

de Bruin, W.C.C., Hertogs, K., Leenders, W.P.J., Depla, E. and Yap, S.H. (1995) Hepatitis B virus: specific binding and internalisation of small HBsAg by human hepatocytes. J. Gen. Virol., **76**, 1047-1050.

Dejean, A., Sonigo, P., Wain-Hobson, S. and Tiollais, P. (1984) Specific hepatitis B virus integration through a viral 11-base pair direct repeat. Proc. Natl. Acad. Sci. USA, **81**, 5350-5354.

Delius, H., Gough, N.M., Cameron, C.H. and Murray, K. (1983) Structure of the Hepatitis B Virus Genome. J. Virol., **47**, 337-343.

del Val, M., Schlicht, H-J., Volkmer, H., Messerk, M., Reddehase, M.J. and Koszinowski, V.H. (1991) Protection against Lethal Cytomegalovirus Infection by a Recombinant Vaccine Containing a Single Nonameric T-Cell Epitope. J. Virol., **65**, 3641-3646.

dePolo, N.J., Giachetti, C. and Holland, J.J. (1987) Continuing Coevolution of Virus and Defective Interfering Particles and of Viral Genome Sequences during Undiluted Passages: Virus Mutants Exhibiting Nearly Complete Resistance to Formerly Dominant Defective Particles. J. Virol., **61**, 454-464.

Diepolder, H.M., Jung, M-C., Wierenga, E., Hiffman, R.M., Zachoval, R., Gerlach, T.J., Scholz, S., Heavner, G., Riethmüller, G. and Pape, G.R. (1996) Anergic TH1

Clones Specific for Hepatitis B Virus (HBV) Core Peptides Are Inhibitory to Other HBV Core-Specific $CD4^+$ Cells In Vitro. J. Virol., **70**, 7540-7548.

Dingwall, C., Robbins, J., Dilworth, S.M., Roberts, D. and Richardson, W.D. (1988) The Nucleoplasmin Nuclear Location Sequence Is Larger and More Complex than That of SV-40 Large T Antigen. J. Cell Biol., **107**, 841-849.

Dopheide, T.A.A. and Azad, A.A. (1996) The hepatitis B virus X protein is a potent AMP kinase. J. Gen. Virol., **77**, 173-176.

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 factors. EMBO J., **14**, 4747-4757.

Dornan, E., Boner, W., Wakeling, M., Hadziyannis, S., Fattovich, G. and Carman, W.F. (1996) Shift in cellular distribution by confocal microscopy of an in vitro expressed HBcAg from sequential patient samples. Abstract A195, IXth Triennial International Symposium on Viral Hepatitis and Liver Disease.

Dreesman, G.R., Sanchez, Y., Ionescu-Matiu, I., Sparrow, J.T., Six, H.R., Peterson, D.L., Hollinger, F.B. and Melnick, J.L. (1982) Antibody to hepatitis B surface antigen after a single inoculation of uncoupled synthetic HBsAg peptides. Nature, **295**, 158-160.

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, D.E., Lingappa, V.R. and Ganem, D. (1986) Hepatitis B Surface Antigen: an Unusual Secreted Protein Initially Synthesized as a Transmembrane Polypeptide. Mol. Cell. Biol., **6**, 1454-1463.

Eble, D.E., MacRae, D.R., Lingappa, V.R. and Ganem, D. (1987) Multiple Topogenic Sequences Determine the Transmembrane Orientation of Hepatitis B Surface Antigen. Mol. Cell. Biol., **7**, 3591-3601.

Eble, D.E., Lingappa, V.R. and Ganem, D. (1990) The N-Terminal (pre-S2) Domain of a Hepatitis B Virus Surface Glycoprotein Is Translocated across Membranes by Downstream Signal Sequences. J. Virol., **64**, 1414-1419.

Eckhardt, S.G., Milich, D.R. and McLachlan, A. (1991) Hepatitis B Virus Core Antigen Has Two Nuclear Localization Sequences in the Arginine-Rich Carboxyl Terminus. J. Virol., **65**, 575-582.

Ehata, T., Omata, M., Yokosuka, O., Hosoda, K. and Ohto, M. (1992) Variations in Codons 84-101 in the Core Nucleotide Sequence Correlate with Hepatocellular Injury in Chronic Hepatitis. J. Clin. Invest., **89**, 332-338.

Ehata, T., Omata, M., Chuang, W-L, Yokosuka, O., Ito, Y., Hosoda, K. and Ohto, M. (1993) Mutations in Core Nucleotide Sequence of Hepatitis B Virus Correlate with Fulminant and Severe Hepatitis. J. Clin. Invest., **91**, 1206-1213.

Fallows, D.A. and Goff, S.P. (1995) Mutations in the ε Sequences of Human Hepatitis B Virus Affect both RNA Encapsidation and Reverse Transcription. J. Virol., **69**, 3067-3073.

Farza, H., Hadchouel, M., Scotto, J., Tiollais, P., Babin, C. and Pourcel, C. (1988) Replication and Gene Expression of Hepatitis B Virus in a Transgenic Mouse That Contains the Complete Viral Genome. J. Virol., **62**, 4144-4152.

Feitelson, M.A., Marion, P.L. and Robinson, W.S. (1982) Core Particles of Hepatitis B Virus and Ground Squirrel Hepatitis Virus I. Relationship Between Hepatitis B Core Antigen and Ground Squirrel Hepatitis Core Antigen-Associated Polypeptides by Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis and Tryptic Peptide Mapping. J. Virol., **43**, 687-696.

Feitelson, M.A. and Clayton, M.M. (1990) X Antigen Polypeptides in the Sera of Hepatitis B Virus-Infected Patients. Virology, **177**, 367-371.

Feitelson, M.A., Zhu, M., Duan, L-X. and London, W.T. (1993) Hepatitis B x antigen and p53 are associated *in vitro* and in liver tissues from patients with primary hepatocellular carcinoma. Oncogene, **8**, 1109-1117.

Feitelson, M.A., Duan, L-X., Guo, J., Horiike, N., McIntyre, G., Blumberg, B.S., Thomas, H.C. and Carman, W. (1995) Precore and X region mutants in hepatitis B virus infections among renal dialysis patients. J. Viral Hepatitis, **2**, 19-31.

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. Virology, **194**, 137-148.

Ferns, R.B. and Tedder, R.S. (1986) Human and Monoclonal Antibodies to Hepatitis B Core Antigen Recognise a Single Immunodominant Epitope. J. Med. Virol., **19**, 193-203.

Ferrari, C., Mondelli, M., Penna, A., Fiaccadori, F. and Chisari, F.V. (1987a) Functional characterization of cloned intrahepatic, hepatitis B virus nucleoprotein-specific helper T cell lines. J. Immunol., **139**, 539-544.

Ferrari, C., Penna, A., Giuberti, T., Tong, M.J., Ribera, E., Fiaccadori, F. and Chisari, F.V. (1987b) Intrahepatic, nucleocapsid antigen-specific T cells in chronic active hepatitis B. J. Immunol., **139**, 2050-2058.

Ferrari, C., Penna, A., Bertoletti, A., Valli, A., Antoni, A.D., Giuberti, T., Cavalli, A., Petit, M-A. and Fiaccadori, F. (1990) Cellular immune response to hepatitis B virus-encoded antigens in acute and chronic hepatitis B virus infection. J. Immunol., **145**, 3442-3449.

Ferrari, C., Bertoletti, A., Penna, A., Cavalli, A., Missale, G., Pilli, M., Giuberti, T., Chisari, F.V. and Fiaccadori, F. (1991) Identification of Immunodominant T Cell Epitopes of the Hepatitis B Virus Nucleocapsid Antigen. J. Clin. Invest., **88**, 214-222.

Ferrari, C., Cavalli, A., Penna, A., Valli, A., Bertoletti, A., Pedretti, G., Pilli, M., Vitali, P., Neri, T.M., Giuberti, T. and Fiaccadori, F. (1992) Fine Specificity of the Human T-Cell Response to the Hepatitis B Virus preS1 Antigen. Gastoenterology, **103**, 255-263.

Fiordalisi, G., Cariani, E., Mantero, G., Zanetti, A., Tanzi, E., Chiaramante, M. and Primi, D. (1990) High Genomic Variability in the Pre-C Region of Hepatitis B Virus in Anti-HBe, HBV DNA-Positive Chronic Hepatitis. J. Med. Vir., **31**, 297-300.

Fiordalisi, G., Primi, D., Tanzi, E., Magni, E., Incarbone, C., Zanetti, A.R. and Cariani, E. (1994) Hepatitis B Virus C Gene Heterogeneity in a Familial Cluster of Anti-HBc Negative Chronic Carriers. J. Med. Virol., **42**, 109-114.

Foster, G.R., Ackrill, A.M., Goldin, R.D., Kerr, I.M., Thomas, H.C. and Stark, G.R. (1991) Expression of the terminal protein region of hepatitis B virus inhibits cellular responses to interferons α and β and double-stranded RNA. Proc. Natl. Acad. Sci. USA, **88**, 2888-2892. *Erratum* (1995) Proc. Natl. Acad. Sci. USA, **92**, 3632.

Foster, G.R., Goldin, R.D., Hay, A., McGarvey, M.J., Stark, G.R. and Thomas, H.C. (1993) Expression of the Terminal Protein of Hepatitis B Virus Is Associated with Failure to Respond to Interferon Therapy. Hepatology, **17**, 757-762.

Fouillot, N. and Rossignol, J-M. (1996) Translational stop codons in the precore sequence of hepatitis B virus pre-C RNA allow translation reinitiation at downstream AUGs. J. Gen. Virol., **77**, 1123-1127.

Fouillot, N., Tlouzeau, S., Rossignol, J-M. and Jean-Jean, O. (1993) Translation of the Hepatitis B Virus P Gene by Ribosomal Scanning as an Alternative to Internal Initiation. J. Virol., **67**, 4886-4895.

Galibert, F., Mandart, E., Fitoussi, F., Tiollais, P. and Charnay, P. (1979) Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in *E.coli.* Nature, **281**, 646-650.

Galibert, F., Chen, T.N. and Mandart, E. (1982) Nucleotide sequence of a cloned Woodchuck Hepatitis Virus Genome: Comparison with the Hepatitis B Virus Sequence. J. Virol., **41**, 51-65.

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., Weiser, B., Barchuk, A., Broen, R.J. and Varmus, H.E. (1982a) Biological Characterization of Acute Infection with Ground Squirrel Hepatitis Virus. J. Virol., **44**, 366-373.

Ganem, D., Greenbaum, L. and Varmus, H.E. (1982b) Virion DNA of Ground Squirrel Hepatitis Virus: Structural Analysis and Molecular Cloning. J. Virol., **44**, 374-383.

Garcia, P.D., Ou, J-H., Rutter, W.J. and Walter, P. (1988) Targeting of the Hepatitis B Virus Precore Protein to the Endoplasmic Reticulum Membrane: After Signal Peptide Cleavage Translocation Can Be Aborted and the Product Released into the Cytoplasm. J. Cell Biol., **106**, 1093-1104.

Gerber, M.A., Hadziyannis, S., Vissoulis, C., Schaffner, F., Paronetto, F. and Popper, H. (1974) Electron Microscopy and Immunoelectronmicroscopy of Cytoplasmic Hepatitis B Antigen in Hepatocytes. Am. J. Path., **75**, 489-502.

Gerelsaikhan, T., Tavis, J.E. and Bruss, V. (1996) Hepatitis B Virus Nucleocapsid Envelopment Does Not Occur without Genomic DNA Synthesis. J. Virol., **70**, 4269-4274.

Gerlich, W.H. and Robinson, W.S. (1980) Hepatitis B Virus Contains Protein Attached to the 5' Terminus of Its Complete DNA Strand. Cell, **21**, 801-809.

Gerlich, W.H., Goldmann, U., Müller, R., Stibbe, W. and Wolff, W. (1982) Specificity and Localization of the Hepatitis B Virus-Associated Protein Kinase. J. Virol., **42**, 761-766.

Giles, J.P., McCollum, R.W., Berndtson, L.W. and Krugman, S. (1969) Relation of Australia/SH Antigen to the Willowbrook MS-2 Strain. New. Engl. Jnl. Med., **281**, 119-122.

Gilles, P.N., Fey, G. and Chisari, F.V. (1992) Tumor Necrosis Factor Alpha Negatively Regulates Hepatitis B Virus Gene Expression in Transgenic Mice. J. Virol., **66**, 3955-3960.

Girones, R. and Miller, R.H. (1989) Mutation Rate of the Hepadnavirus Genome. Virology, **170**, 595-597.

Gluzman, Y. (1981) SV40-Transformed Simian Cells Support the Replication of Early SV40 Mutants. Cell, **23**, 175-182.

Gocke, D.J. and Kavey, N.B. (1969) Hepatitis Antigen: Correlation with Disease and Infectivity of Blood-Donors. Lancet, **i**, 1055-1059.

Gowans, E.J., Burrell, C.J., Jilbert, A.R. and Marmion, B.P. (1985) Cytoplasmic (but not Nuclear) Hepatitis B Virus (HBV) Core Antigen Reflects HBV DNA Synthesis at the Level of the Infected Hepatocyte. Intervirology, **24**, 220-225.

Greenspan, D., Kenobi, O.W., Palese, P. and Krystal, M. (1988) Two Nuclear Location Signals in the Influenza Virus NS1 Nonstructural Protein. J. Virol., **62**, 3020-3026.

Gudat, F. and Bianchi, L. (1977) Evidence for phasic sequences in nuclear HBcAg formation and cell membrane-directed flow of core particles in chronic hepatitis B. Gastroenterology, **73**, 1194-1197.

Guidotti, L.G., Martinez, V., Loh, Y-T., Rogler, C.E. and Chisari, F.V. (1994a) Hepatitis B Virus Nucleocapsid Particles Do Not Cross the Hepatocyte Nuclear Membrane in Transgenic Mice. J. Virol., **68**, 5469-5475.

Guidotti, L.G., Ando, K., Hobbs, M.V., Ishikawa, T., Runkel, L., Schreiber, R.D. and Chisari, F.V. (1994b) Cytotoxic T lymphocytes inhibit hepatitis B virus gene expression by a noncytolytic mechanism in transgenic mice. Proc. Natl. Acad. Sci. USA, **91**, 3764-3768.

Guidotti, L.G., Guilot, S. and Chisari, F.V. (1994c) Interleukin-2 and Alpha/Beta Interferon Down-Regulate Hepatitis B Virus Gene Expression In Vivo by Tumor Necrosis Factor-Dependent and -Independent Pathways. J. Virol., **68**, 1265-1270.

Guidotti, L.G., Matzke, B., Schaller, H. and Chisari, F.V. (1995) High-Level Hepatitis B Virus Replication in Transgenic Mice. J. Virol., **69**, 6158-6169.

Guidotti, L.G., Matzke, B., Pasquinelli, C., Shoenberger, J.M., Rogler, C.E. and Chisari, F.V. (1996) The Hepatitis B Virus (HBV) Precore Protein Inhibits HBV Replication in Transgenic Mice. J. Virol., **70**, 7056-7061.

Guilhot, S., Guidotti, L.G. and Chisari, F.V. (1993) Interleukin-2 Downregulates Hepatitis B Virus Gene Expression in Transgenic Mice by a Posttranscriptional Mechanism. J. Virol., **67**, 7444-7449.

Günther, S., Meisel, H., Reip, A., Miska, S., Krüger, D.H. and Will, H. (1992) Frequent and Rapid Emergence of Mutated Pre-C Sequences in HBV from e-Antigen

Positive Carriers Who Seroconvert to Anti-HBe during Interferon Treatment. Virology, **187**, 271-279.

Günther, S., Li, B-C., Miska, S., Krüger, D.H., Meisel, H. and Will, H. (1995) A Novel Method for Efficient Amplification of Whole Hepatitis B Virus Genomes Permits Rapid Functional Analysis and Reveals Deletion Mutants in Immunosuppressed Patients. J. Virol., **69**, 5437-5444.

Günther, S., Piwon, N., Iwanska, A., Schilling, R., Meisel, H. and Will, H. (1996a) Type, Prevalence, and Significance of Core Promoter/Enhancer II Mutations in Hepatitis B Viruses from Immunosuppressed Patients with Severe Liver Disease. J. Virol., **70**, 8318-8331.

Günther, S., Baginski, S., Kissel, H., Reinke, P., Krüger, D.H., Will, H. and Meisel, H. (1996b) Accumulation and Persistence of Hepatitis B Virus Core Gene Deletion Mutants in Renal Transplant Patients Are Associated With End-Stage Liver Disease. Hepatology, **24**, 751-758.

Guo, W., Chen, M., Yen, T.S.B. and Ou, J-H. (1993) Hepatocyte-Specific Expression of the Hepatitis B Virus Core Promoter Depends on Both Positive and Negative Regulation. Mol. Cell. Biol., **13**, 443-448.

Hadziyannis, S., Gerber, M.A., Vissoulis, C. and Popper, H. (1973) Cytoplasmic Hepatitis B Antigen in 'Ground-Glass' Hepatocytes of Carriers. Arch. Pathol., **96**, 327-330.

Handley-Gearhart, P.M., Stephen, A.G., Trausch-Azar, A.S., Cienchanover, A. and Schwartz A.L. (1994) Human Ubiquitin-activating Enzyme E1: Indication of potential nuclear and cytoplasmic subpopulations using epitope-tagged cDNA constructs. J. Biol. Chem., **269**, 33171-33178.

Haruna, Y., Hayashi, N., Katayama, K., Yuki, N., Kasahara, A., Sasaki, Y., Fusamoto, H. and Kamada, T. (1991) Expression of X Protein and Hepatitis B Virus Replication in Chronic Hepatitis. Hepatology, **13**, 417-421.

Hasegawa, K., Huang, J., Regers, S.A., Blum, H.E. and Liang, T.J. (1994) Enhanced Replication of a Hepatitis B Virus Mutant Associated with an Epidemic of Fulminant Hepatitis. J. Virol., **68**, 1651-1659.

Hatton, T., Zhou, S. and Standring, D.N. (1992) RNA- and DNA-Binding Activities in Hepatitis B Virus Capsid Protein: a Model for Their Roles in Viral Replication. J. Virol., **66**, 5232-5241.

Haviv, I., Vaizel, D. and Shaul, Y. (1996) pX, the HBV-encoded coactivator, interacts with components of the transcription machinery and stimulates transcription in a TAF-independent manner. EMBO J., **15**, 3413-3420.

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.

Hertogs, K., Leenders, W.P.J., Depla, E., de Bruin, W.C.C., Meheus, L., Raymackers, J., Moshage, H. and Yap, S.H. (1993) Endonexin II, Present on Human Liver Plasma Membranes, Is a Specific Binding Protein of Small Hepatitis B Virus (HBV) Envelope Protein. Virology, **197**, 549-557.

Hertogs, K., Depla, E., Crabbé, T., de Bruin, W., Leenders, W., Moshage, H. and Yap, S.H. (1994) Spontaneous Development of Anti-Hepatitis B Virus Envelope (Anti-Idiotypic) Antibodies in Animals Immunized with Human Liver Endonexin II or with the $F(ab')_2$ Fragment of Anti-Human Liver Endonexin II Immunoglobulin G: Evidence for a Receptor-Ligand-Like Relationship between Small Hepatitis B Surface Antigen and Endonexin II. J. Virol., **68**, 1516-1521.

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.

Holland, J.J. (1990) Defective Viral Genomes. *In* Fields, B.N. and Knipe, D.M. (ed.), Virology, Second Edition (Raven Press, New York), 151-165.

Honigwachs, J., Faktor, O., Dikstein, R., Shaul, Y. and Laub, O. (1989) Liver-Specific Expression of Hepatitis B Virus Is Determined by the Combined Action of the Core Gene Promoter and the Enhancer. J. Virol., **63**, 919-924.

Hoofnagle, J.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. Ann. Int. Med., **94**, 744-748.

Horwich, A.L., Furtak, K., Pugh, J. and Summers, J. (1990) Synthesis of Hepadnavirus Particles That Contain Replication-Defective Duck Hepatitis B Virus Genomes in Cultured HuH7 Cells. J. Virol., **64**, 642-650.

Hou, J., Karayiannis, P., Waters, J., Luo, K., Liang, C. and Thomas, H.C. (1995) A Unique Insertion in the *S* Gene of Surface Antigen-Negative Hepatitis B Virus Chinese Carriers. Hepatology, **21**, 273-278.

Hruska, J.F., Clayton, D.A., Rubenstein, J.L.R. and Robinson, W.S. (1977) Structure of Hepatitis B Dane Particle DNA Before and After the Dane Particle Polymerase Reaction. J. Virol., **21**, 666-672.

Hsu, H-C., Su, I-J., Lai, M-Y., Chen, D-S., Chang, M-H., Chuang, S-M. and Sung, J-L. (1987) Biologic and prognostic significance of hepatocyte hepatitis B core antigen expressions in the natural course of chronic hepatitis B virus infection. J. Hepatol., **5**, 45-50.

Hsu, H-Y., Chang, M-H., Hsieh, K-H., Lee, C-Y., Lin, H-H., Hwang, L-H., Chen, P-J. and Chen, D-S. (1992) Cellular Immune Response to HBcAg in Mother-to-Infant Transmission of Hepatitis B Virus. Hepatology, **15**, 770-776.

Hu, J. and Seeger, C. (1996) Hsp90 is required for the activity of a hepatitis B virus reverse transcriptase. Proc. Natl. Acad. Sci. USA, **93**, 1060-1064.

Hu, J., Toft, D.O. and Seeger, C. (1997) Hepadnavirus assembly and reverse transcription require a multi-component chaperone complex which is incorporated into nucleocapsids. EMBO J., **16**, 59-68.

Huang, J., Kwong, J., Sun, E.C-Y. and Liang, T.J. (1996) Proteasome Complex as a Potential Cellular Target of Hepatitis B Virus X Protein. J. Virol., **70**, 5582-5591.

Huovila, A-P.J., Eder, A.M. and Fuller, S.D. (1992) Hepatitis B Surface Antigen Assembles in a Post-ER, Pre-Golgi Compartment. J. Cell Biol., **118**, 1305-1320.

Ikeda, T., Pignatelli, M., Lever, A.M.L. and Thomas, H.C. (1986) Relationship of HLA protein display to activation of 2-5A synthetase in HBe antigen or anti-HBe positive chronic HBV infection. Gut, **27**, 1498-1501.

Itoh, Y., Takai, E., Ohnuma, H., Kitajima, K., Tsuda, F., Machida, A., Mishiro, S., Nakamura, T., Miyakawa, Y. and Mayumi, M. (1986) A synthetic peptide vaccine involving the product of the pre-*S*(2) region of hepatitis B virus DNA: Protective efficacy in chimpanzees. Proc. Natl. Acad. Sci. USA, **83**, 9174-9178.

Iwarson, I., Tabor, E., Thomas, H.C., Goodall, A., Waters, J., Snoy, P., Shih, J.W-K. and Gerety, R.J. (1985) Neutralization of Hepatitis B Virus Infectivity by a Murine Monoclonal Antibody: An Experimental Study in the Chimpanzee. J. Med. Virol., **16**, 89-96.

Jean-Jean, O., Levrero, M., Will, H., Perricaudet, M. and Rossignol, J-M. (1989a) Expression Mechanism of the Hepatitis B Virus (HBV) C Gene and Biosynthesis of HBe Antigen. Virology, **170**, 99-106.

Jean-Jean, O., Salhi, S., Carlier, D., Elie, C., de Recondo, A-M. and Rossignol, J-M. (1989b) Biosynthesis of Hepatitis B Virus e Antigen: Directed Mutagenesis of the Putative Aspartyl Protease Site. J. Virol., **63**, 5497-5500.

Jin, Y., Shih, W-K. and Berkower, I. (1988) Human T cell response to the surface antigen of hepatitis B virus (HBsAg). J. Exp. Med., **168**, 293-306.

Johnson, J.L., Raney, A.K. and McLachlan, A. (1995) Characterization of a Functional Hepatocyte Nuclear Factor 3 Binding Site in the Hepatitis B Virus Nucleocapsid Promoter. Virology, **208**, 147-158.

Jokelainen, P.T., Krohn, K., Prince, A.M. and Finlayson, N.D.C. (1970) Electron Microscopic Observations on Virus-Like Particles Associated with SH Antigen. J. Virol., **6**, 685-689.

Jung, M-C., Spengler, U., Schraut, W., Hoffmann, R., Zachoval., R., Eisenburg, J., Eichenlaub, D., Riethmüller, G., Paumgartner, G., Ziegler-Heitbrock, H.W.L., Will, H. and Pape, G.R. (1991) Hepatitis B virus antigen-specific T-cell activation in patients with acute and chronic hepatitis B. J. Hepatol., **13**, 310-317.

Jung, M-C., Diepolder, H.M., Spengler, U., Wierenga, E.A., Zachoval, R., Hoffmann, R.M., Eichenlaub, D., Frösner, G., Will, H. and Pape, G.R. (1995) Activation of a Heterogeneous Hepatitis B (HB) Core and e Antigen Specific $CD4^+$ T-Cell Population during Seroconversion to Anti-HBe and Anti-HBs in Hepatitis B Virus Infection. J. Virol., **69**, 3358-3368.

Junker-Niepmann, M., Bartenschlager, R. and Schaller, H. (1990) A short *cis*-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. EMBO J., **9**, 3389-3396.

Kalderon, D., Richardson, W.D., Markham, A.F. and Smith, A.E. (1984) Sequence requirements for nuclear location of simian virus 40 large-T antigen. Nature, **311**, 33-38.

Kann, M. and Gerlich, W.H. (1994) Effect of Core Protein Phosphorylation by Protein Kinase C on Encapsidation of RNA within Core Particles of Hepatitis B Virus. J. Virol., **68**, 7993-8000.

Kann, M., Bischof, A. and Gerlich, W.G. (1997) In Vitro Model for the Nuclear Transport of the Hepadanvirus Genome. J. Virol., **71**, 1310-1316.

Kann, M., Thomssen, R., Köchel, H.G. and Gerlich, W.H. (1993) Characterization of the endogenous protein kinase activity of the hepatitis B virus. Arch. Virol. (Suppl.), **8**, 53-62.

Kaplan, P.M., Greenman, R.L., Gerin, J.L., Purcell, R.H. and Robinson, W.S. (1973) DNA Polymerase Associated with Human Hepatitis B Antigen. J. Virol., **12**, 995-1005.

Kekule, A.S., Lauer, U., Weiss, L., Luber, B. and Hofschneider, P.H. (1993) Hepatitis B virus transactivator HBx uses a tumour promoter signalling pathway. Nature, **361**, 742-745.

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

Khudyakov, Y.E., Kalinina, T.I., Neplyueva, V.S., Gazina, E.V., Kadoshnikov, Y.P., Bogdanova, S.L. and Smirnov, V.D. (1991) The effect of the structure of the terminal regions of the hepatitis B virus gene *C* polypeptide on the formation of core antigen (HBcAg) particles. Biomed. Sci., **2**, 257-265.

Kidd, A.H. and Kidd-Ljunggren, K. (1996) A revised secondary structure model for the 3'-end of hepatitis B virus pregenomic RNA. Nucleic Acids Res., **24**, 3295-3301.

Kim, C-M., Koike, K., Saito, I., Mujamura, T. and Jay, G. (1991) *HBx* gene of hepatitis B virus induces liver cancer in transgenic mice. Nature, **351**, 317-320.

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. Nucleic Acids Res., **21**, 3967-3975.

Knowles, B.B., Howe, C.C. and Aden, D.P. (1980) Human Hepatocellular Carcinoma Cell Lines Secrete the Major Plasma Proteins and Hepatitis B Surface Antigen. Science, **209**, 497-499.

Köck, J. and Schlicht, H-J. (1993) Analysis of the Earliest Steps of Hepadnavirus Replication: Genome Repair after Infectious Entry into Hepatocytes Does Not Depend on Viral Polymerase Activity. J. Virol., **67**, 4867-4874.

Koike, K., Moriya, K., Iino, S., Yotsuyanagi, H., Endo, Y., Miyamura, T. and Kurokawa, K. (1994a) High-level Expression of Hepatitis B Virus HBx Gene and Hepatocarcinogenesis in Transgenic Mice. Hepatology, **19**, 810-819.

Koike, K., Moriya, K., Yotsuyanagi, H., Iino, S. and Kurokawa, K. (1994b) Induction of Cell Cycle Progression by Hepatitis B Virus HBx Gene Expression in Quiescent Mouse Fibroblasts. J. Clin. Invest., **94**, 44-49.

Koike, K. and Takada, S. (1995) Biochemistry and Functions of Hepatitis B Virus X Protein. Intervirology, **38**, 89-99.

Kojima, T., Bloemen, J. and Desmet, V.J. (1987) Immune electron microscopic determination of hepatitis B core antigen (HBcAg) in liver cell plasma membranes. Liver, **7**, 191-200.

Korba, B., Wells, F., Tennant, B., Yoakum, G., Purcell, R. and Gerin, J. (1986) Hepadnavirus Infection of Peripheral Blood Lymphocytes In Vivo: Woodchuck and Chimpanzee Models of Viral Hepatitis. J. Virol., **58**, 1-8.

Kwee, L., Lucito, R., Aufiero, B. and Schneider, R.J. (1992) Alternate Translation Initiation on Hepatitis B Virus X mRNA Produces Multiple Polypeptides That Differentially Transactivate Class II and III Promoters. J. Virol., **66**, 4382-4389.

Laemmli, U.K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature, **227**, 680-685.

Lai, M.E., Solinas, A., Mazzoleni, A.P., Deplano, A., Farci, P., Lisci, V., Porru, A., Tocco, A. and Balestrieri, A. (1994) The role of pre-core hepatitis B virus mutants on the long-term outcome of chronic hepatitis B virus hepatitis. A longitudinal study. J. Hepatol., **20**, 773-781.

Lamberts, C., Nassal, M., Velhagen, I., Zentgraf, H. and Schroder, C.H. (1993) Precore-Mediated Inhibition of Hepatitis B Virus Progeny DNA Synthesis. J. Virol., **67**, 3756-3762.

Lanford, R.E. and Notvall, L. (1990) Expression of Hepatitis B Virus Core and Precore Antigens in Insect Cells and Characterization of a Core-Associated Kinase Activity. Virology, **176**, 222-233.

Lanford, R.E., Notvall, L. and Beames, B. (1995) Nucleotide Priming and Reverse Transcription Activity of Hepatitis B Virus Polymerase Expressed in Insect Cells. J. Virol., **69**, 4431-4439.

Lanford, R.E., Notvall, L., Lee, H. and Beames, B. (1997) Transcomplementation of Nucleotide Priming and Reverse Transcription between Independently Expressed TP and RT Domains of the Hepatitis B Virus Reverse Transcriptase. J. Virol., **71**, 2996-3004.

Laskus, T., Rakela, J. and Persing, D.H. (1994) The Stem-Loop Structure of the *cis*-Encapsidation Signal Is Highly Conserved in Naturally Occurring Hepatitis B Virus Variants. Virology, **200**, 809-812.

Laub, O. and Bavand, M.R. (1988) Two Proteins with Reverse Transcriptase Activities Associated with Hepatitis B Virus-Like Particles. J. Virol., **62**, 626-628.

le Bouvier, G.L. (1971) The Heterogeneity of Australia Antigen. J. Infect. Dis., **123**, 671-675.

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.

Leenders, W.P.J., Glansbeek, H.L., de Bruin, W.C.C. and Yap, S-H. (1990) Binding of the Major and Large HBsAg to Human Hepatocytes and Liver Plasma Membranes: Putative External and Internal Receptors for Infection and Secretion of Hepatitis B Virus. Hepatology, **12**, 141-147.

Li, L., Sheng, M-H., Tong, S-P., Chen, H-Z. and Wen, Y-M. (1986) Transplacental transmission of hepatitis B virus. Lancet, **ii**, 872.

Li, J-S., Tong, S-P., Wen, Y-M., Vitvitski, L., Zhang, Q. and Trépo, C. (1993) Hepatitis B Virus Genotype A Rarely Circulates as an HBe-Minus Mutant. Possible Contribution of a Single Nucleotide in the Precore Region. J. Virol., **67**, 5402-5410.

Lien, J-M., Aldrich, C.E. and Mason, W.S. (1986) Evidence that a Capped Oligoribonucleotide Is the Primer for Duck Hepatitis B Virus Plus-Strand DNA Synthesis. J. Virol., **57**, 229-236.

Lin, C-G. and Lo, S.J. (1992) Evidence for Involvement of a Ribosomal Leaky Scanning Mechanism in the Translation of the Hepatitis B Virus *Pol* Gene from the Viral Pregenome RNA. Virology, **188**, 342-352.

Lingappa, J.R., Martin, R.L., Wang, M.L., Ganem, D., Welch, W.J. and Lingappa, V.R. (1994) A Eukaryotic Cytosolic Chaperonin Is Associated with a High Molecular Weight Intermediate in the Assembly of Hepatitis B Virus Capsid, a Multimeric Particle. J. Cell Biol., **125**, 99-111.

Lo, W-Y. and Ting, L-P. (1994) Repression of Enhancer II Activity by a Negative Regulatory Element in the Hepatitis B Virus Genome. J. Virol., **68**, 1758-1764.

Loeb, D.D. and Tian, R. (1995) Transfer of the Minus Strand of DNA during Hepadnavirus Replication is not Invariable but Prefers a Specific Location. J. Virol., **69**, 6886-6891.

Loeb, D.D., Gulya, K.G. and Tian, R. (1997) Sequence Identity of the Terminal Redundancies on the Minus-Strand DNA Template Is Necessary but Not Sufficient for the Template Switch during Hepadnavirus Plus-Strand DNA Synthesis. J. Virol., **71**, 152-160.

Löhr, H.F., Gerken, G., Schlicht, H-J., zum Büschenfelde, K-H.M. and Fleischer, B. (1993) Low Frequency of Cytotoxic Liver-Infiltrating T Lymphocytes Specific for

Endogenous Processed Surface and Core Proteins in Chronic Hepatitis B. J. Infect. Dis., **168**, 1133-1139.

Löhr, H.F., Weber, W., Schlaak, J., Goergen, B., zum Büschenfelde, K-H.M. and Gerken, G. (1995) Proliferative Response of $CD4^+$ T Cells and Hepatitis B Virus Clearance in Chronic Hepatitis With or Without Hepatitis B e-Minus Hepatitis B Virus Variants. Hepatology, **22**, 61-68.

Lok, A.S.F., Akarca, U. and Greene, S. (1994) Mutations in the pre-core region of hepatitis B virus serve to enhance the stability of the secondary structure of the pre-genome encapsidation signal. Proc. Natl. Acad. Sci. USA, **91**, 4077-4081.

Lu, X., Block, T.M. and Gerlich, W.H. (1996) Protease-Induced Infectivity of Hepatitis B Virus for a Human Hepatoblastoma Cell Line. J. Virol., **70**, 2277-2285.

Lucito, R. and Schneider R.J. (1992) Hepatitis B Virus X Protein Activates Transcription Factor NF-κB without a Requirement for Protein Kinase C. J. Virol., **66**, 983-991.

Lyons, R.H., Ferguson, B.Q. and Rosenberg, M. (1987) Pentapeptide Nuclear Localization Signal in Adenovirus E1a. Mol. Cell. Biol., **7**, 2451-2456.

Machida, A., Kishimoto, S., Ohnuma, H., Baba, K., Oda, K., Nakamura, T., Miyakawa, Y. and Mayumi, M. (1983) A Hepatitis B Surface Antigen Polypeptide (P31) With the Receptor for Polymerised Human as Well as Chimpanzee Albumins. Gastoenterology, **85**, 268-274.

Machida, A., Ohnuma, H., Takai, E., Tsuda, F., Tanaka, T., Naito, M., Munekata, E., Miyakawa, Y. and Mayumi, M. (1989) Antigenic sites on the arginine-rich carboxyl-terminal domain of the capsid protein of hepatitis B virus distinct from hepatitis B core or e antigen. Mol. Immunol., **26**, 413-421.

Machida, A., Ohnuma, H., Tsuda, F., Yoshikawa, A., Itoshi, Y., Tanaka, T., Kishimoto, 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.

Mackay, P., Lees, J. and Murray, K. (1981) The Conversion of Hepatitis B Core Antigen Synthesized in E.coli Into e Antigen. J. Med. Virol., **8**, 237-243.

Magnius, L.O. and Espmark, J.A. (1972) New specificities in Australia antigen positive sera distinct from the le Bouvier determinants. J. Immunol., **109**, 1017-1021.

Maguire, H.F., Hoeffler, J.P. and Siddiqui, A. (1991) HBV X Protein Alters the DNA Binding Specificity of CREB and ATF-2 by Protein-Protein Interactions. Science, **252**, 842-844.

Mandart, E., Kay, A. and Galibert, F. (1984) Nucleotide Sequence of a Cloned Duck Hepatitis B Virus Genome: Comparison with Woodchuck and Human Hepatitis B Virus Sequences. J. Virol., **49**, 782-792.

Mangold, C.M.T. and Streeck, R.E. (1993) Mutational Analysis of the Cysteine Residues in the Hepatitis B Virus Small Envelope Protein. J. Virol., **67**, 4588-4597.

Marinos, G., Torre, F., Günther, S., Thomas, M.G., Will, H., Williams, R. and Naoumov, N.V. (1996) Hepatitis B Virus Variants With Core Gene Deletions in the Evolution of Chronic Hepatitis B Infection. Gastroenterology, **111**, 183-192.

Marion, P.L., Oshiro, L.S., Regnery, D.C., Scullard, G.H. and Robinson, W.S. (1980) A virus in Beechey ground sqirrels that is related to hepatitis B virus of humans. Proc. Natl. Acad. Sci. USA, **77**, 2941-2945.

Maruyama, T., McLachlan, A., Iino, S., Koike, K., Kurokawa, K. and Milich, D.R. (1993a) The Serology of Chronic Hepatitis B Infection Revisited. J. Clin. Invest., **91**, 2586-2595.

Maruyama, T., Iino, S., Koike, K., Yasuda, K. and Milich, D.R. (1993b) Serology of Acute Exacerbation in Chronic Hepatitis B Virus Infection. Gastroenterology, **105**, 1141-1151.

Mason, W.S., Seal, G. and Summers, J. (1980) Virus of Pekin Ducks with Structural and Biological Relatedness to Human Hepatitis B Virus. J. Virol., **36**, 829-836.

McLauchlan, J., Liefkins, K. and Stow, N.D. (1994) The herpes simplex virus type 1 UL37 gene product is a component of virus particles. J. Gen. Virol., **75**, 2047-2052.

Michalak, T. and Nowoslawski, A. (1982) Crystalline Aggregates of Hepatitis B Core Particles in Cytoplasm of Hepatocytes. Intervirology, **17**, 247-252.

Milich, D.R. and McLachlan, A. (1986) The Nucleocapsid of Hepatitis B Virus Is Both a T-Cell-Independent and a T-Cell-Dependent Antigen. Science, **234**, 1398-1401.

Milich, D.R., McLachlan, A., Moriarty, A. and Thornton, G.B. (1987a) Immune response to hepatitis B virus core antigen (HBcAg): Localization of T cell recognition sites within HBcAg/HBeAg. J. Immunol., **139**, 1223-1231.

Milich, D.R., McLachlan, A., Thornton, G.B. and Hughes, J.L. (1987b) Antibody production to the nucleocapsid and envelope of the hepatitis B virus primed by a single synthetic T cell site. Nature, **329**, 547-549.

Milich, D.R., McLachlan, A., Stahl, S., Wingfield, P., Thornton, G.B., Hughes, J.L. and Jones, J.E. (1988) Comparative immunogenicity of hepatitis B virus core and e antigens. J. Immunol., **141**, 3617-3624.

Milich, D.R., Jones, J.E., Hughes, J.L., Price, J., Raney, A.K. and McLachlan, A. (1990) Is a function of the secreted hepatitis B e antigen to induce immunologic tolerance *in utero*? Proc. Natl. Acad. Sci. USA, **87**, 6599-6603.

Milich, D.R., Peterson, D.L., Schödel, F., Jones, J.E. and Hughes, J.L. (1995a) Preferential Recognition of Hepatitis B Nucleocapsid Antigens by Th1 or Th2 Cells Is Epitope and Major Histocompatibility Complex Dependent. J. Virol., **69**, 2776-2785.

Milich, D.R., Schödel, F., Peterson, D.L., Jones, J.E. and Hughes, J.L. (1995b) Characterization of self-reactive T cells that evade tolerance in hepatitis B e antigen transgenic mice. Eur. J. Immunol., **25**, 1663-1672.

Milich, D.R., Schödel, F., Hughes, J.L., Jones, J.E. and Peterson, D.L. (1997) The Hepatitis B Virus Core and e Antigens Elicit Different Th Cell Subsets: Antigen Structure Can Affect Th Cell Phenotype. J. Virol., **71**, 2192-2201.

Miller, R.H. and Robinson, W.S. (1984) Hepatitis B Virus DNA Forms in Nuclear and Cytoplasmic Fractions of Infected Human Liver. Virology, **137**, 390-399.

Missale, G., Redeker, A., Person, J., Fowler, P., Guilhot, S., Schlicht, H-J, Ferrari, C. and Chisari, F.V. (1993) HLA-A31- and HLA-Aw68-restricted Cytotoxic T Cell Responses to a Single Hepatitis B Virus Nucleocapsid Epitope during Acute Viral Hepatitis. J. Exp. Med., **177**, 751-762.

Miyanohara, A., Imamura, T., Araki, M., Sugawara, K., Ohtomo, N. and Matsubara, K. (1986) Expression of Hepatitis B Virus Core Antigen Gene in *Saccharomyces cerevisiae*: Synthesis of Two Polypeptides Translated from Different Initiation Codons. J. Virol., **59**, 176-180.

Molnar-Kimber, K.L., Summers, J., Taylor, J.M. and Mason, W.S. (1983) Protein Covalently Bound to Minus-Strand DNA Intermediates of Duck Hepatitis B Virus. J. Virol., **45**, 165-172.

Montaño, L., Miescher, G.C., Goodall, A.H., Wiedmann, K.H., Janossy, G. and Thomas, H.C. (1982) Hepatitis B Virus and HLA Antigen Display in the Liver During Chronic Hepatitis B Virus Infection. Hepatology, **2**, 557-561.

Montaño, L., Aranguibel, F., Boffill, M., Goodall, A.H., Janossy, G. and Thomas, H.C. (1983) An Analysis of the Composition of the Inflammatory Infiltrate in Autoimmune and Hepatitis B Virus-Induced Chronic Liver Disease. Hepatology, **3**, 292-296.

Montross, L., Watkins, S., Moreland, R.B., Mamon, H., Caspar, D.L.D. and Garcea, R.L. (1991) Nuclear Assembly of Polyomavirus Capsids in Insect Cells Expressing the Major Capsid Protein VP1. J. Virol., **65**, 4991-1998.

Moreland, R.B., Lagevin, G.L., Singer, R.H., Garcea, R.L. and Hereford, L.M. (1987) Amino Acid Sequences That Determine the Nuclear Localization of Yeast Histone 2B. Mol. Cell. Biol., **7**, 4048-4057.

Morgan, D.J. and Dimmock, N.J. (1992) Defective Interfering Influenza Virus Inhibits Immunopathological Effects of Infectious Virus in the Mouse. J. Virol., **66**, 1188-1192.

Moriyama, T., Guilhot, S., Klopchin, K., Moss, B., Pinkert, C.A., Palmiter, R.D., Brinster, R.L., Kanagawa, O. and Chisari, F.V. (1990) Immunobiology and Pathogenesis of Hepatocellular Injury in Hepatitis B Virus Transgenic Mice. Science, **268**, 361-364.

Naoumov, N.V., Schneider, R., Grötzinger, T., Jung, M.C., Miska, S., Pape, G.R. and Will, H. (1992) Precore Mutant Hepatitis B Virus Infection and Liver Disease. Gastroenterology, **102**, 538-543.

Nassal, M. (1992a) The Arginine-Rich Domain of the Hepatitis B Virus Core Protein Is Required for Pregenome Encapsidation and Productive Viral Positive-Strand DNA Synthesis but Not for Virus Assembly. J. Virol., **66**, 4107-4116.

Nassal, M. (1992b) Conserved Cysteines of the Hepatitis B Virus Core Protein Are Not Required for Assembly of Replication-Competent Core Particles Nor for Their Envelopment. Virology, **190**, 499-505.

Nassal, M. and Rieger, A. (1993) An Intramolecular Disulfide Bridge between Cys-7 and Cys61 Determines the Structure of the Secretory Core Gene Product (e Antigen) of Hepatitis B Virus. J. Virol., **67**, 4307-4315.

Nassal, M. and Rieger, A. (1996) A Bulged Region of the Hepatitis B Virus RNA Encapsidation Signal Contains the Replication Origin for Discontinuous First-Strand DNA Synthesis. J. Virol., **70**, 2764-2773.

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

Nassal, M., Galle, P.R. and Schaller, H. (1989) Proteaselike Sequence in Hepatitis B Virus Core Antigen Is Not Required for e Antigen Generation and May Not Be Part of an Aspartic Acid-Type Protease. J. Virol., **63**, 2598-2604.

Nassal, M., Junker-Niepmann, M. and Schaller, H. (1990) Translational Inactivation of RNA Function: Discrimination against a Subset of Genomic Transcripts during HBV Nucleocapsid Assembly. Cell, **63**, 1357-1363.

Nassal, M., Rieger, A. and Steinau, O. (1992) Topological Analysis of the Hepatitis B Virus Core Particle by Cysteine-Cysteine Cross-Linking. J. Mol. Biol., **225**, 1013-1025.

Natoli, G., Avantaggiati, M.L., Chirillo, P., Costanzo, A., Artini, M., Balsano, C. and Levrero, M. (1994) Induction of the DNA-Binding Activity of c-Jun/c-Fos Heterodimers by the Hepatitis B Virus Transactivator pX. Mol. Cell. Biol., **14**, 989-998.

Nayersina, R., Fowler, P., Guilhot, S., Missale, G., Cerny, A., Schlicht, H-J., Vitiello, A., Chesnut, R., Person, J.L., Redeker, A.G. and Chisari, F.V. (1993) HLA A2

Restricted Cytotoxic T Lymphocyte Responses to Multiple Hepatitis B Surface Antigen Epitopes during Hepatitis B Virus Infection. J. Immunol., **150**, 4659-4671.

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.

Norder, H., Ebert, J.W., Fields, H.A., Mushahwar, I.K. and Magnius L.O. (1996) Complete Sequencing of a Gibbon Hepatitis B Virus Genome Reveals a Unique Genotype Distantly Related to the Chimpanzee Hepatitis B Virus. Virology, **218**, 214-223.

Nouri-Aria, K.T., Magrin, S., Alexander, G.J.M., Anderson, M.G., Williams, R. and Eddleston, A.L.W.F. (1988) Abnormal T-cell activation in chronic hepatitis B viral infection: a consequence of monocyte dysfunction? Imunology, **64**, 733-738.

Ohtsu, T., Sai, T., Oka, M., Sugai, Y. and Tobinai, K. (1991) Activation of Hepatitis B Virus Infection by Chemotherapy Containing Glucocorticoid in Hepatitis B Virus Carriers with Hematologic Malignancies. Jpn. J. Clin. Oncol., **21**, 360-365.

Okamoto, H., Imai, M., Tsuda, F., Tanaka, T., Miyakawa, Y. and Mayumi, M. (1987a) Point Mutation in the *S* Gene of Hepatitis B Virus for a d/y or w/r Subtypic Change in Two Blood Donors Carrying a Surface Antigen of Compound Subtype adyr or adwr. J. Virol., **61**, 3030-3034.

Okamoto, H., Tsuda, F. and Mayumi, M. (1987b) Defective Mutants of Hepatitis B Virus in the Circulation of Symptom-Free Carriers. Jpn. J. Exp. Med., **57**, 217-221.

Okamoto, H., Omi, S., Wang, Y., Itoh, Y., Tsuda, F., Tanaka, T., Akahame, Y., Miyakawa, Y. and Mayumi, M. (1989) The loss of subtypic determinants in alleles, *d/y* or *w/r*, on hepatitis B surface antigen. Mol. Immunol., **26**, 197-205.

Okamoto, H., Yotsumoto, S., Akahane, Y., Yamanaka, T., Miyazaki, Y., Sugai, Y., Tsuda, F., Tanaka, T., Miyakawa, Y. and Mayumi, M. (1990) Hepatitis B Viruses with Precore Region Defects Prevail in Persistently Infected Hosts along with Seroconversion to the Antibody against e Antigen. J. Virol., **64**, 12998-1303.

Okamoto, H., Wang, Y., Tanaka, T., Machida, A., Miyakawa, Y. and Mayumi, M. (1993) Trans-complementation among naturally occurring deletion mutants of hepatitis

B virus and integrated viral DNA for the production of viral particles with mutant genomes in hepatoma cell lines. J. Gen. Virol., **74**, 407-414.

Ono, Y., Onda, H., Sasada, R., Igarashi, K., Sugino, Y. and Nishioka, K. (1983) The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype adr and adw. Nucleic Acids Res., **11**, 1747-1757.

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.

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.

Ou, J-H., Yeh, C-T. and Yen, T.S.B. (1989) Transport of Hepatitis B Virus Precore Protein into the Nucleus after Cleavage of Its Signal Peptide. J. Virol., **63**, 5238-5243.

Ou, J-H., Bao, H., Shih, C. and Tahara, S.M. (1990) Preferred Translation of Human Hepatitis B Virus Polymerase from Core Protein- but Not from Precore Protein-Specific Transcript. J. Virol., **64**, 4578-4581.

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.

Pathak, V.K. and Temin, H.M. (1990) Broad spectrum of *in vivo* forward mutations, hypermutations and mutational hotspots in a retroviral shuttle vector after a single replication cycle: Deletions and deletions with insertions. Proc. Natl. Acad. Sci. USA, **87**, 6024-6028.

Patzer, E.J., Nakamura, G.R., Simonsen, C.C., Levinson, A.D. and Brands, R. (1986) Intracellular Assembly and Packaging of Hepatitis B Surface Antigen Particles Occur in the Endoplasmic Reticulum. J. Virol., **58**, 884-892.

Pedrali-Noy, G., Spadari, S., Miller-Faurès, A., Miller, A.O.A., Kruppa, J. and Koch, G. (1980) Synchronization of HeLa cell cultures by inhibition of DNA polymerase α with aphidicolin. Nucleic Acids Res., **8**, 377-387.

Penna, A., Chisari, F.V., Bertoletti, A., Missale, G., Fowler, P., Giuberti, T., Fiaccadori, F. and Ferrari, C. (1991) Cytotoxic T Lymphocytes Recognize an HLA-A2-restricted Epitope within the Hepatitis B Virus Nucleocapsid Antigen. J. Exp. Med., **174**, 1565-1570.

Penna, A., Fowler, P., Bertoletti, A., Guilhot, S., Moss, B., Margolskee, R.F., Cavalli, A., Valli, A., Fiaccadori, F., Chisari, F.V. and Ferrari, C. (1992) Hepatitis B Virus (HBV)-Specific Cytotoxic T-Cell (CTL) Response in Humans: Characterization of HLA Class II-Restricted CTLs That Recognize Endogenously Synthesized HBV Envelope Antigens. J. Virol., **66**, 1193-1198.

Perri, S. and Ganem, D. (1996) A Host Factor That Binds near the Termini of Hepatitis B Virus Pregenomic RNA. J. Virol., **70**, 6803-6809.

Persing, D.H., Varmus, H.E. and Ganem, D. (1986) Inhibition of Secretion of Hepatitis B Surface Antigen by a Related Presurface Polypeptide. Science, **234**, 1388-1391.

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.

Peterson, D.L., Nath, N. and Gavilanes, F. (1982) Structure of the Hepatitis B Surface Antigen: Correlation of subtype with amino acid sequence and location of the carbohydrate moiety. J. Biol. Chem., **257**, 10414-10420.

Peterson, D.L., Paul, D.A., Lam, J., Tribby, I.I.E. and Achord, D.T. (1984) Antigenic structure of hepatitis B surface antigen: Identification of the 'd' subtype determinant by chemical modification and use of monoclonal antibidies. J. Immunol., **132**, 920-927.

Petit, M-A., Capel, F., Riottot, M.M., Daugnet, C. and Pillot, J. (1987) Antigenic Mapping of the Surface Proteins of Infectious Hepatitis B Virus Particles. J. Gen. Virol., **68**, 2759-2767.

Petit, M-A., Dubanchet, S., Capel, F., Voet, P., Daugnet, C. and Hauser, P. (1991) HepG2 Cell Binding Activities of Different Hepatitis B Virus Isolates: Inhibitory Effect of Anti-HBs and Anti-preS1(21-47). Virology, **180**, 483-491.

Petit, M-A., Capel, F., Dubanchet, S. and Mabit, H. (1992) Pre-S1 Specific Binding Proteins as Potential Receptors for Hepatitis B Virus in Human Hepatocytes. Virology, **187**, 211-222.

Poisson, F., Severac, A., Hourioux, C., Goudeau, A. and Roingeard, P. (1997) Both Pre-S1 and S Domains of Hepatitis B Virus Envelope Proteins Interact with the Core Particle. Virology, **228**, 115-120.

Poitrine, A., Chousterman, S., Chousterman, M., Naveau, S., Thang, M.N. and Chaput, J-C. (1985) Lack of *In Vivo* Activation of the Interferon System in HBsAg-Positive Chronic Active Hepatitis. Hepatology, **5**, 171-174.

Pollack, J.R. and Ganem, D. (1993) An RNA Stem-Loop Structure Directs Hepatitis B Virus Genomic RNA Encapsidation. J. Virol., **67**, 3254-3263.

Pollack, J.R. and Ganem, D. (1994) 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., Ruvoletto, M.G., Gerlich, W.H., Heermann, K-H., Bardini, R. and Alberti, A. (1989a) Identification of an Attachment Site for Human Liver Plasma Membranes on Hepatitis B Virus Particles. Virology, **173**, 522-530.

Pontisso, P., Petit, M-A., Bankowski, M.J. and Peebles, M.E. (1989b) Human Liver Plasma Membranes Contain Receptors for the Hepatitis B Virus Pre-S1 Region and, via Polymerised Human Serum Albumin, for the Pre-S2 Region. J. Virol., **63**, 1981-1988.

Pontisso, P., Ruvoletto, M.G., Tiribelli, C., Gerlich, W.H., Ruol, A. and Alberti, A. (1992) The preS1 domain of hepatitis B virus and IgA cross-react in their binding to the hepatocyte surface. J. Gen. Virol., **73**, 2041-2045.

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

Prince, A.M. (1968) An antigen detected in the blood during the incubation period of serum hepatitis. Proc. Natl. Acad. Sci. USA, **60**, 814-821.

Pulsinelli, G.A. and Temin, H.M. (1991) Characterization of Large Deletions Occurring during a Single Round of Retrovirus Vector Replication: Novel Deletion Mechanism Involving Errors in Strand Transfer. J. Virol., **65**, 4786-4797.

Pumpens, P., Borisova, G.P., Crowther, R.A. and Grens, E. (1995) Hepatitis B Virus Core Particles as Epitope Carriers. Intervirology, **38**, 63-74.

Pushko, P., Sallberg, M., Borisova, G., Ruden, U., Bichko, V., Wahren, B., Pumpens, P. and Magnius, L. (1994) Identification of Hepatitis B Virus Core Protein Regions Exposed or Internalized at the Surface of HBcAg Particles by Scanning with Monoclonal Antibodies. Virology, **202**, 912-920.

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.

Qadri, I., Conaway, J.W., Conaway, R.C., Schaack, J. and Siddiqui, A. (1996) Hepatitis B virus transactivator protein, HBx, associates with the components of TFIIH and stimulates the DNA helicase activity of TFIIH. Proc. Natl. Acad. Sci. USA, **93**, 10578-10583.

Radziwill, G., Tucker, W. and Schaller, H. (1990) Mutational Analysis of the Hepatitis B Virus P Gene Product: Domain Structure and RNase H Activity. J. Virol., **64**, 613-620.

Raimondo, G., Scheider, R., Stemler, M., Smedile, V., Rodino, G. and Will, H. (1990) A New Hepatitis B Virus Variant in a Chronic Carrier with Multiple Episodes of Viral Reactivation and Acute Hepatitis. Virology, **179**, 64-68.

Raney, A.K., Milich, D.R. and McLachlan, A. (1989) Characterization of Hepatitis B Virus Major Surface Antigen Gene Transcriptional Regulatory Elements in Differentiated Hepatoma Cell Lines. J. Virol., **63**, 3919-3925.

Raney, A.K., Milich, D.R., Easton, A.J. and McLachlan, A. (1990) Differentiation-Specific Transcriptional Regulation of the Hepatitis B Virus Large Surface Antigen Gene in Human Hepatoma Cell Lines. J. Virol., **64**, 2360-2368.

Raney, A.K., Zhang, P. and McLachlan, A. (1995) Regulation of Transcription from the Hepatitis B Virus Large Surface Antigen Promoter by Hepatocyte Nuclear Factor 3. J. Virol., **69**, 3265-3272.

Raney, A.K., Johnson, J.L., Palmer, C.A. and McLachlan, A. (1997) Members of the Nuclear Receptor Superfamily Regulate Transcription from the Hepatitis B Virus Nucleocapsid Promoter. J. Virol., **71**, 1058-1071.

Rehermann, R., Fowler, P., Sidney, J., Person, J., Redeker, A., Brown, M., Moss, B., Sette, A. and Chisari, F.V. (1995) The Cytotoxic T Lymphocyte Response to Multiple Hepatitis B Virus Polymerase Epitopes During and After Acute Viral Hepatitis. J. Exp. Med., **181**, 1047-1058.

Rehermann, R., Lau, D., Hoofnagle, J.H. and Chisari, F.V. (1996) Cytotoxic T Lymphocyte Responsiveness after Resolution of Chronic Hepatitis B Virus Infection. J. Clin. Invest., **97**, 1655-1665.

Rieger, 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 revealed by a combined *in vivo* selection/ *in vitro* amplification system. Nucleic Acids. Res., **23**, 3909-3915.

Rieger, A. and Nassal, M. (1996) Specific Hepatitis B Virus Minus-Strand DNA Synthesis Requires Only the 5' Encapsidation Signal and the 3'-Proximal Direct Repeat DR1*. J. Virol., **70**, 585-589.

Roberts, I.M., Bernard, C.C., Vyas, G.N. and Mackay, I.R. (1975) T-cell dependence of immune response to hepatitis B antigen in mice. Nature, **254**, 606-607.

Roberts, J.D., Preston, B.D., Johnston, L.A., Soni, A., Loeb, L.A. and Kunkel, T.A. (1989) Fidelity of Two Retroviral Reverse Transcriptases during DNA-Dependent DNA Synthesis In Vitro. Mol. and Cell. Biol., **9**, 469-476.

Robinson, W.S. and Greenman, R.L. (1974) DNA Polymerase in the Core of the Human Hepatitis B Virus Candidate. J. Virol., **13**, 1231-1236.

Robinson, W.S., Clayton, D.A. and Greenman, R.L. (1974) DNA of a Human Hepatitis B Virus Candidate. J. Virol., **14**, 384-391.

Roingeard, P., Romet-Lemonne, J-L., Leturcq, D., Goudeau, A. and Essex, M. (1990) Hepatitis B Virus Core Antigen (HBc Ag) Accumulation in an HBV Nonproducer Clone of HepG2-Transfected Cells Is Associated with Cytopathic Effect. Virology, **179**, 113-120.

Roossinck, M.J. and Siddiqui, A. (1987) In Vivo Phosphorylation and Protein Analysis of Hepatitis B Virus Core Antigen. J. Virol., **61**, 955-961.

Rosmorduc, O., Petit, M-A., Pol, S., Capel, F., Bortolotti, F., Berthelot, P., Brechot, C. and Kremsdorf, D. (1995) *In Vivo* and *In Vitro* Expression of Defective Hepatitis B Virus Particles Generated by Spliced Hepatitis B Virus RNA. Hepatology, **22**, 10-19.

Rosmorduc, O., Sirma, H., Soussan, P., Lebon, P., Horrisberger, M., Brechot, C. and Kremsdorf, D. (1996) Expression of the interferon-induced antiviral protein MxA is inhibited in vitro by the hepatitis B virus (HBV) capsid protein. Abstract A221, IXth Triennial International Symposium on Viral Hepatitis and Liver Disease.

Rossner, M.T. (1992) Hepatitis B Virus X-Gene Product: A Promiscuous Transcriptional Transactivator. J. Med. Virol., **36**, 101-117.

Roychoudhury, S., Faruqi, A.F. and Shih, C. (1991) Pregenomic RNA Encapsidation Analysis of Eleven Missense and Nonsense Polymerase Mutants of the Human Hepatitis B Virus. J. Virol., **65**, 3617-3624.

Russnak, R. and Ganem, D. (1990) Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. Genes and Dev., **4**, 764-776.

Saïb, A., Périès, J. and de Thé, H. (1993) A defective human foamy provirus generated by pregenome splicing. EMBO J., **12**, 4439-4444.

Sakamoto, Y., Yamada, G., Mizuno, M., Nishihara, T., Kinoyama, S., Kobayashi, T., Takahashi, T. and Nagashima, H. (1983) Full and Empty Particles of Hepatitis B Virus in Hepatocytes from Patients with HBsAg-Postive Chronic Active Hepatitis. Lab. Invest., **48**, 678-682.

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.

Sällberg, M., Rudén, U., Magnius, L.O., Harthus, H.P., Noah, M. and Wahren, B. (1991a) Characterisation of a Linear Binding Site for a Monoclonal Antibody to Hepatitis B Core Antigen. J. Med. Virol., **33**, 248-252.

Sällberg, M., Rudén, U., Wahren, B., Noah, M. and Magnius, L.O. (1991b) Human and murine B-cells recognize the HBeAg/Beta (or HBe2) epitope as a linear determinant. Mol. Immunol., **28**, 719-726.

Sällberg, M., Rudén, U., Wahren, B. and Magnius, L.O. (1994) Immune Recognition of Linear Antigenic Regions Within the Hepatitis B pre-C and Core Translation Products Using Synthetic Peptides. J. Med. Virol., **42**, 7-15.

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

Santantonio, T., Jung, M-C., Miska, S., Pastore, G., Pape, G.R. and Will, H. (1991) Prevalence and Type of Pre-C HBV Mutants in Anti-HBe Positive Carriers with Chronic Liver Disease in a Highly Endemic Area. Virology, **183**, 840-844.

Santantonio, T., Jung, M-C., Schneider, R., Fernholz, D., Milella, M., Monno, L., Pastore, G., Pape, G.R. and Will, H. (1992) Hepatitis B Virus Genomes that Cannot Synthesize Pre-S2 Proteins Occur Frequently and as Dominant Virus Populations in Chronic Carriers in Italy. Virology, **188**, 948-952.

Sattler, F. and Robinson, W.S. (1979) Hepatitis B Viral DNA Molecules Have Cohesive Ends. J. Virol., **32**, 226-233.

Scaglioni, P.P., Melegari, M. and Wands, J.R. (1997) Posttranscriptional Regulation of Hepatitis B Virus Replication by the Precore Protein. J. Virol., **71**, 345-353.

Schaller, H. and Fischer, M. (1991) Transcriptional Control of Hepadnavirus Gene Expression. Curr. Top. Microbiol. Immunol., **168**, 21-39.

Schek, N., Bartenschlager, R., Kuhn, C. and Schaller, H. (1991) Phosphorylation and rapid turnover of hepatitis B virus X-protein expressed in HepG2 cells from a recombinant vaccinia virus. Oncogene, **6**, 1735-1744.

Schlicht, H-J. and Schaller, H. (1989) The Secretory Core Protein of Human Hepatitis B Virus Is Expressed on the Cell Surface. J. Virol., **63**, 5399-5404.

Schödel, F., Moriarty, 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.

Scullard, G.H., Smith, C.I., Merigan, T.C., Robinson, W.S. and Gregory, P.B. (1981) Effects of Immunosuppressive Therapy on Viral Markers in Chronic Active Hepatitis B. Gastroenterology, **61**, 987-991.

Seeger, C. and Maragos, J. (1990) Identification and Characterization of the Woodchuck Hepatitis Virus Origin of DNA Replication. J. Virol., **64**, 16-23.

Seeger, C. and Maragos, J. (1991) Identification of a Signal Necessary for initiation of Reverse Transcription of the Hepadnavirus Genome. J. Virol., **65**, 5190-5195.

Seeger, C., Ganem, D. and Varmus, H.E. (1984) Nucleotide Sequence of an Infectious Molecularly Cloned Genome of Ground Squirrel Hepatitis Virus. J. Virol., **51**, 367-375.

Seeger, C., Ganem, D. and Varmus, H.E. (1986) Biochemical and Genetic Evidence for the Hepatitis B Virus Replication Strategy. Science, **232**, 477-484.

Seifer, M and Standring, D.N. (1993) Stability Governs the Apparent Expression of 'Particulate' Hepatitis B e Antigen by Mutant Hepatitis B Virus Core Particles. Virology, **196**, 70-78.

Seifer, M. and Standring, D.N. (1994) A Protease-Sensitive Hinge Linking the Two Domains of the Hepatitis B Virus Core Protein Is Exposed on the Viral Capsid Surface. J. Virol., **68**, 5548-5555.

Seifer, M. and Standring, D.N. (1995) Ribonucleoprotein Complex Formation by the Human Hepatitis B Virus Polymerase. Intervirology, **38**, 295-303.

Seifer, M., Zhou, S. and Standring, D.N. (1993) A Micromolar Pool of Antigenically Distinct Precursors Is Required To Initiate Cooperative Assembly of Hepatitis B Virus Capsids in *Xenopus* Oocytes. J. Virol., **67**, 249-257.

Seto, E., Mitchell, P. and Yen, T.S.B. (1990) Transactivation by the hepatitis B virus X protein depends on AP-2 and other transcription factors. Nature, **344**, 72-74.

Shaul, Y., Rutter, W.J. and Laub, O. (1985) A human hepatitis B viral enhancer element. EMBO J., **4**, 427-430.

Sheu, S.Y. and Lo, S.J. (1994) Biogenesis of the hepatitis B viral middle (M) surface protein in a human hepatoma cell line: demonstration of an alternative secretion pathway. J. Gen. Virol., **75**, 3031-3039.

Siddiqui, A., Sattler, F. and Robinson, W.S. (1979) Restriction endonuclease cleavage map and location of unique features of the DNA of hepatitis B virus, subtype adw_2. Proc. Natl. Acad. Sci. USA, **76**, 4664-4668.

Siddiqui, A., Marion, P.L. and Robinson, W.S. (1981) Ground Squirrel Hepatitis Virus DNA: Molecular Cloning and Comparison with Hepatitis B Virus DNA. J. Virol., **38**, 393-397.

Siddiqui, A., Gaynor, R., Srinivasan, A., Mapoles, J., Farr, R.W. (1989) *trans*-Activation of Viral Enhancers Including Long Terminal Repeat of the Human Immunodeficiency Virus by the Hepatitis B Virus X protein. Virology, **169**, 479-484.

Simon, K., Lingappa, V.R. and Ganem, D. (1988) Secreted Hepatitis B Surface Antigen Polypeptides Are Derived from a Transmembrane Precursor. J. Cell Biol., **107**, 2163-2168.

Sprengel, R., Kaleta, E.F. and Will, H. (1988) Isolation and Characterisation of a Hepatitis B Virus Endemic in Herons. J. Virol., **62**, 3832-3839.

Stahl, S., MacKay, P., Magazin, M., Bruce, S.A. 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.

Standring, D.N., Rutter, W.J., Varmus, H.E. and Ganem, D. (1984) Transcription of the Hepatitis B Surface Antigen Gene in Cultured Murine Cells Initiates Within the Presurface Region. J. Virol., **50**, 563-571.

Standring, D.N., Ou, J-H., Masiarz, F.R. and Rutter, W.J. (1988) A signal peptide encoded within the precore region of hepatitis B virus directs the secretion of a heterogeneous population of e antigens in *Xenopus* oocytes. Proc. Natl. Acad. Sci. USA, **85**, 8405-8409.

Stemler, M., Weimer, T., Tu, Z-X., Wan, D-F., Levrero, M., Jung, C., Pape, G.R. and Will, H. (1990) Mapping of B-Cell Epitopes of the Human Hepatitis B Virus X Protein. J. Virol., **64**, 2802-2809.

Stibbe, W. and Gerlich, W.H. (1982) Variable Protein Composition of Hepatitis B Surface Antigen from Different Donors. Virology, **123**, 436-442.

Stibbe, W. and Gerlich, W.H. (1983) Structural Relationships Between Minor and Major Proteins of Hepatitis B Surface Antigen. J. Virol., **46**, 626-628.

Stirk, H.J., Thornton, J.M. and Howard, C.R. (1992) A Topological Model for Hepatitis B Surface Antigen. Intervirology, **33**, 148-158.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Use of T7 RNA Polymerase to Direct Expression of Cloned Genes. Meth. Enzymol., **185**, 60-89.

Su, F. and Schneider, R.J. (1996) Hepatitis B Virus HBx Protein Activates Transcription Factor NF-κB by Acting on Multiple Cytoplasmic Inhibitors of *rel*-Related Proteins. J. Virol., **70**, 4558-4566.

Summers, J., O'Connell, A. and Millman, I. (1975) Genome of hepatitis B virus: Restriction enzyme cleavage and structure of DNA extracted from Dane particles. Proc. Natl. Acad. Sci. USA, **72**, 4597-4601.

Summers, J., Smolec, J.M. and Snyder, R. (1978) A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. Proc. Natl. Acad. Sci. USA, **75**, 4533-4537.

Summers, J. and Mason, W.S. (1982) Replication of the Genome of a Hepatitis B-Like Virus by Reverse Transcription of an RNA Intermediate. Cell, **29**, 403-415.

Suzuki, M. (1989) SPXX, a Frequent Sequence Motif in Gene Regulatory Proteins. J. Mol. Biol., **207**, 61-84.

Suzuki, T., Kajino, K., Masui, N., Saito, I. and Miyamura, T. (1990) Alternative Splicing of Hepatitis B Virus RNAs in HepG2 Cells Transfected with the Viral DNA. Virology, **179**, 881-885.

Takada, S., Kido, H., Fukutomi, A., Mori, T. and Koike, K. (1994) Interaction of hepatitis B virus X protein with a serine protease, tryptase TL_2, as an inhibitor. Oncogene, **9**, 341-348.

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

Takayanagi, M., Kakumu, S., Ishikawa, T., Higashi, Y., Yoshioka, K. and Wakita, T. (1993) Comparison of Envelope and Precore/Core Variants of Hepatitis B Virus (HBV) during Chronic HBV Infection. Virology, **196**, 138-145.

Takehara, K., Ireland, D. and Bishop, D.H.L. (1988) Co-expression of the Hepatitis B Surface and Core Antigens Using Baculovirus Multiple Expression Vectors. J. Gen. Virol., **69**, 2763-2777.

Tarn, W-Y. and Steitz, J.A. (1997) Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. Trends in Biochem. Sci., **22**, 132-137.

Tavis, J.E. and Ganem, D. (1993) Expression of functional hepatitis B virus polymerase in yeast reveals it to be the sole viral component required for correct initiation of reverse transcription. Proc. Natl. Acad. Sci. USA, **90**, 4107-4111.

Tavis, J.E. and Ganem, D. (1996) Evidence for Activiation of the Hepatitis B Virus Polymerase by Binding to Its RNA Template. J. Virol., **70**, 5741-5750.

Tavis, J.E., Perri, S. and Ganem, D. (1994) Hepadnavirus Reverse Transcription Initiates within the Stem-Loop of the RNA Packaging Signal and Employs a Novel Strand Transfer. J. Virol., **68**, 3536-3543.

Testut, P., Renard, C-A., Terradillos, O., Vitviski-Trepo, L., Tekaia, F., Degott, C., Blake, J., Boyer, B. and Buendida, M.A. (1996) A New Hepadnavirus Endemic in Arctic Ground Squirrels in Alaska. J. Virol., **70**, 4210-4219.

Thursz, M.R., Kwiattzowski, D., Allsopp, C.E.M., Greenwood, B.M., Thomas, H.C. and Hill, A.V.S. (1995) Association between an MHC Class II allele and clearance of hepatitis B virus in the Gambia. New Engl. Jnl. Med., **332**, 1065-1069.

Toh, H., Hajashida, 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.

Tong, S., Li, J., Vitvitski, L. and Trépo, C. (1990) Active Hepatitis B Virus Replication in the Presence of Anti-HBe Is Associated with Viral Variants Containing an Inactive Pre-C Region. Virology, **176**, 596-603.

Tong, S-P., Li, J-S., Vitvitski, L. and Trépo, C. (1992) Replication Capacities of Natural and Artificial Precore Stop Codon Mutants of Hepatitis B Virus: Relevance of Pregenome Encapsidation Signal. Virology, **191**, 237-245.

Tordjeman, M., Fontan, G., Rabillon, V., Martin, J., Trepo, C., Hoffenbach, A., Mabrouk, K., Sabatier, J.M., van Rietschoten, J. and Somme, G. (1993) Characterization of Minor and Major Antigenic Regions Within the Hepatitis B Virus Nucleocapsid. J. Med. Virol., **41**, 221-229.

Towbin, H., Staehhelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. USA, **76**, 4350-4354.

Treichel, U., zum Büschenfelde, K-H.M., Stockert, R.J., Poralla, T. and Gerken, G. (1994) The asialoglycoprotein receptor mediates hepatic binding and uptake of natural hepatitis B virus particles derived from viraemic carriers. J. Gen. Virol., **75**, 3021-3029.

Treinin, M. and Laub, O. (1987) Identification of a Promoter Element Located Upstream from the Hepatitis B Virus X Gene. Mol. Cell. Biol., **7**, 545-548.

Truant, R., Antunovic, J., Greenblatt, J., Prives, C. and Cromlish, J.A. (1995) Direct Interaction of the Hepatitis B Virus HBx Protein with p53 Leads to Inhibition by HBx of p53 Response Element-Directed Transactivation. J. Virol., **69**, 1851-1859.

Trujillo, M.A., Letovsky, J., Maguire, H.F., Lopez-Cabrera, M. and Siddiqui, A. Functional analysis of a liver-specific enhancer of the hepatitis B virus. Proc. Natl. Acad. Sci. USA, **88**, 3797-3801.

Tsai, S.L., Chen, P.J., Lai, M.Y., Yang, P.M., Sung, J.L., Huang, J.H., Hwang, L.H., Chang, T.H. and Chen, D.S. (1992) Acute Exacerbations of Chronic Type B Hepatitis Are Accompanied by Increased T Cell Responses to Hepatitis B Core And e Antigens. J. Clin. Invest., **89**, 87-96.

Tsai, S.L., Chen, M.H., Yeh, C.T., Chu, C.M., Lin, A.N., Chiou, F.H., Chang, T.H. and Liaw, Y.F. (1996) Purification and Characterization of a Naturally Processed Hepatitis B Virus Peptide Recognized by $CD8^+$ Cytotoxic T Lymphocytes. J. Clin. Invest., **97**, 577-584.

Tsui, L.V., Guidotti, L.G., Ishikawa, T. and Chisari, F.V. (1995) Posttranscriptional clearance of hepatitis B virus RNA by cytotoxic T lymphocyte-activated hepatocytes. Proc. Natl. Acad. Sci. USA, **92**, 12398-12402.

Tur-Kaspa, R., Shaul, Y., Moore, D.D., Burk, R.D., Okret, S., Poellinger, L. and Shafritz, D.A. (1988) The Glucocorticoid Receptor Recognizes a Specific Nucleotide Sequence in Hepatitis B Virus DNA Causing Increased Activity of the HBV Enhancer. Virology, **167**, 630-633.

Tur-Kaspa, R., Klein, A. and Aharonsen, S. (1992) Hepatitis B Virus Precore Mutants Are Identical in Carriers from Various Ethnic Origins and Are Associated with a Range of Liver Disease Severity. Hepatology, **16**, 1338-1342.

Tuttleman, J.S., Pourcel, C. and Summers, J. (1986) Formation of the Pool of Closed Circular Viral DNA in Hepadnavirus-Infected Cells. Cell, **47**, 451-460.

Twu, J-S., Lee, C-H., Lin, P-M. and Schloemer, R.H. (1988) Hepatitis B virus suppresses expression of human β-interferon. Proc. Natl. Acad. Sci. USA, **85**, 252-256.

Uchida, T., Aye, T.T., Shikata, T., Yano, M., Yatsuhashi, H., Koga, M. and Mima, S. (1994) Evolution of the Hepatitis B Virus Gene During Chronic Infection in Seven Patients. J. Med. Virol., **43**, 148-154.

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.

Ueda, H., Ullrich, S.J., Gangemi, J.D., Kappel, C.A., Ngo, L., Feitelson, M.A. and Jay, G. (1995) Functional inactivation but not structural mutation of p53 causes liver cancer. Nature Genetics, **9**, 41-47.

Ulrich, T.R., Anders, K., Layfield, L., Cheng, L. and Lewin, K.J. (1985) Chronic Active Hepatitis of Hepatitis B and Non-A, Non-B Etiology. Arch. Path. Lab. Med., **109**, 403-407.

Ulrich, P.P., Bhat, R.A., Kelly, I., Brunetto, M.R., Bonino, F. and Vyas, G.N. (1990) A Precore-Defective Mutant of Hepatitis B Virus Associated With e Antigen-Negative Chronic Liver Disease. J. Med. Virol., **32**, 109-118.

Unger, T. and Shaul, Y. (1990) The X protein of hepatitis B virus acts as a transcription factor when targetted to its responsive element. EMBO J., **9**, 1889-1895.

Valliammai, T., Echevarría, J.M., León, P., Tusets, C. and Harrison, T. (1995) Amplification and Sequence Analysis of the Precore and Core Region of the HBV Genome From Sera of Spanish Patients With HBV_2-Like Infections. J. Med. Virol., **46**, 375-379.

Van Etten, R.A., Jackson, P. and Baltimore, D. (1989) The Mouse Type IV c-*abl* Gene Product Is a Nuclear Protein, and Activation of Transforming Ability Is Associated with Cytoplasmic Localization. Cell, **58**, 669-678.

Vento, S., Rondanelli, E.G., Ranieri, S., O'Brien, C.J., Williams, R. and Eddleston, A.L.W.F. (1987) Prospective study of cellular immunity to hepatitis B virus antigens from the early incubation phase of acute hepatitis B. Lancet, **ii**, 119-122.

Wakita, T., Kakumu, S., Shibata, M., Yoshioka, K., Ito, Y., Shinagawa, T., Ishikawa, T., Takanayagi, M. and Morishima, T. (1991) Detection of Pre-C and Core Region Mutants of Hepatitis B Virus in Chronic Hepatitis B Virus Carriers. J. Clin. Invest., **88**, 1793-1801.

Wakita, T., Kakumu, S., Yoshioka, K., Ishikawa, T., Ito, Y. and Shinagawa, T. (1992) Cellular Immune Responses of Peripheral Blood Mononuclear Cells to HBV Antigens during Chronic and Acute HBV Infection. Digestion, **52**, 26-33.

Wallace, L.A. and Carman, W.F. (1997) Surface Gene Variation of HBV: Scientific and Medical Relevance. Viral Hep. Reviews, **3**, 5-16.

Wang, G-H. and Seeger, C. (1993) Novel Mechanism for Reverse Transcription in Hepatitis B Viruses. J. Virol., **67**, 6507-6512.

Wang, W., London, W.T., Lega, L. and Feitelson, M.A. (1991a) HBxAg in the Liver from Carrier Patients with Chronic Hepatitis and Cirrhosis. Hepatology, **14**, 29-37.

Wang, J., Lee, A.S. and Ou, J-H. (1991b) Proteolytic Conversion of Hepatitis B Virus e Antigen Precursor to End Product Occurs in a Postendoplasmic Reticulum Compartment. J. Virol., **65**, 5080-5083.

Wang, X.W., Forrester, K., Yeh, H., Feitelson, M.A., Gu, J-R. and Harris, C.C. (1994) Hepatitis B virus X protein inhibits p53 sequence-specific DNA binding, transcriptional activity, and association with transcription factor ERCC3. Proc. Natl. Acad. Sci. USA, **91**, 2230-2234.

Wasenauer, G., Kock, J. and Schlicht, H-J. (1992) A Cysteine and a Hydrophobic Sequence in the Noncleaved Portion of the Pre-C Leader Peptide Determine the Biophysical Properties of the Secretory Core Protein (HBe Protein) of Human Hepatitis B Virus. J. Virol., **66**, 5338-5346.

Waters, J.A., Jowett, T.P. and Thomas, H.C. (1986) Identification of a Dominant Immunogenic Epitope of the Nucleocapsid (HBc) of the Hepatitis B Virus. J. Med. Virol., **19**, 79-86.

Weber, M., Bronsema, V., Bartos, H., Bosserhoff, A., Bartenschlager, R. and Schaller, H. (1994) Hepadnavirus P Protein Utilizes a Tyrosine Residue in the TP Domain To Prime Reverse Transcription. J. Virol., **68**, 2994-2999.

Weimer, T., Schödel, F., Jung, M-C., Pape, G.R., Alberti, A., Fattovich, G., Beljaars, van Eerd, P.C.M.A. and Will, H. (1990) Antibodies to the RNase H Domain of Hepatitis B Virus P Protein Are Associated with Ongoing Viral Replication. J. Virol., **64**, 5665-5668.

Werner, B.G., Smolec, J.M., Snyder, R. and Summers, J. (1979) Serological Relationship of Woodchuck Hepatitis Virus to Human Hepatitis B Virus. J. Virol., **32**, 314-322.

Whitten, T.M., Quets, A.T. and Schloemer, R.H. (1991) Identification of the Hepatitis B Virus Factor That Inhibits Expression of the Beta Interferon Gene. J. Virol., **65**, 4699-4704.

Will, H., Reiser, W., Weimer, T., Pfaff, E., Büscher, 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 Composition, Conformational Analysis, and *In Vitro* Capsid Assembly. Biochemistry, **34**, 4919-1932.

Wu, P.C., Lai, C.L., Lam, K.C. and Ho, J. (1982) Prednisolone in HBsAg-Positive Chronic Active Hepatitis: Histologic Evaluation in a Controlled Prospective Study. Hepatology, **2**, 777-783.

Wu, J.Y., Zhou, Z-Y., Judd, A., Cartwright, C.A. and Robinson, W.S. (1990) The Hepatitis B Virus-Encoded Transcriptional *Trans*-Activator hbx Appears to Be a Novel Protein Serine/Threonine Kinase. Cell, **63**, 687-695.

Wu, H-L., Chen, P-J., Tu, S-J., Lin, M-H., Lai, M-Y. and Chen, D-S. (1991a) Characterization and Genetic Analysis of Alternatively Spliced Transcripts of Hepatitis B Virus in Infected Human Liver and Transfected HepG2 Cells. J. Virol., **65**, 1680-1686.

Wu, H-L., Chen, P-J., Lin, M-H. and Chen, D-S. (1991b) Temporal Aspects of Major Viral Transcript Expression in Hep G2 Cells Transfected with Cloned Hepatitis B Virus DNA: With Emphasis on the X Transcript. Virology, **185**, 644-651.

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

Xu, Z. and Yen, T.S.B. Intracellular Retention of Surface Protein by a Hepatitis B Virus Mutant That Releases Virion Particles. J. Virol., **70**, 133-140.

Yaginuma, K., Shirakata, Y., Kobayashi, M. and Koike, K. (1987) Hepatitis B virus (HBV) particles are produced in a cell culture system by transient expression of transfected HBV DNA. Proc. Natl. Acad. Sci. USA, **84**, 2678-82.

Yamada, G. and Nakahane, P.K. (1977) Hepatitis B Core and Surface Antigens in Liver Tissue. Light and Electron Microscopic Localization by the Peroxidase-Labelled Antibody Method. Lab. Invest., **36**, 649-659.

Yamamoto, K., Horikita, M., Tsuda, F., Itoh, K., Akahane, Y., Yotsumoto, S., Okamoto, H., Miyakawa, Y. and Mayumi, M. (1994) Naturally Occurring Escape Mutants of Hepatitis B Virus with Various Mutations in the S Gene in Carriers Seropositive for Antibody to Hepatitis B Surface Antigen. J. Virol., **68**, 2671-2676.

Yang, S.Q., Walter, M. and Standring, D.N. (1992) Hepatitis B Virus p25 Precore Protein Accumulates in *Xenopus* Oocytes as an Untranslocated Phosphoprotein with an Uncleaved Signal Peptide. J. Virol., **66**, 37-45.

Yang, W., Guo, J., Ying, Z., Hua, S., Dong, W. and Chen, H. (1994) Capsid Assembly and Involved Functional Analysis of Twelve Core Protein Mutants of Duck Hepatitis B Virus. J. Virol., **68**, 338-345.

Yee, J-K. (1989) A Liver-Specific Enhancer in the Core Promoter Region of Human Hepatitis B Virus. Science, **246**, 658-661.

Yeh, C-T., Liaw, Y-F. and Ou, J-H. (1990) The Arginine-Rich Domain of Hepatitis B Virus Precore and Core Proteins Contains a Signal for Nuclear Transport. J. Virol., **64**, 6141-6147.

Yeh, C-T., Wong, S.W., Fung, Y-K. and Ou, J-H. (1993) Cell cycle regulation of nuclear localization of hepatitis B virus core protein. Proc. Natl. Acad. Sci. USA, **90**, 6459-6463.

Yen, T.S.B. (1993) Regulation of hepatitis B virus gene expression. Seminars in Virology, **4**, 33-42.

Yon, J., Rud, E., Corcoran, T., Kent, K., Rowlands, D. and Clarke, B. (1992) Stimulation of specific immune responses to simian immunodeficiency virus using chimeric hepatitis B core antigen particles. J. Gen. Virol., **73**, 2569-2575.

Yoo, J.Y., Howard, R., Waggoner, J.G. and Hoofnagle, J.H. (1987) Peroxidase-Anti-Peroxidase Detection of Hepatitis B Surface and Core Antigen in Liver Biopsy Specimens From Patients With Chronic Type B Hepatitis. J. Med. Virol., **23**, 273-281.

Yoshikawa, A., Tanaka, T., Hoshi, Y., Kato, N., Tachibana, K., Iizuka, H., Machida, A., Okamoto, H., Yamasaki, M., Miyakawa, Y. and Mayumi, M. (1993) Chimeric Hepatitis B Virus Core Particles with Parts or Copies of the Hepatitis C Virus Core Protein. J. Virol., **67**, 6064-6070.

Yu, M.W., Finlayson, J.S. and Shih, J.W-K. (1985) Interaction Between Various Polymerised Human Albumins and Hepatitis B Surface Antigen. J. Virol., **55**, 736-743.

Yu, X. and Mertz, J.E. (1996) Promoters for Synthesis of the Pre-C and Pregenomic mRNAs of Human Hepatitis B Virus Are Genetically Distinct and Differentially Regulated. J. Virol., **70**, 8719-8726.

Yu, M., Miller, R.H., Emerson, S. and Purcell, R.H. (1996) A Hydrophobic Heptad Repeat of the Core Protein of Woodchuck Hepatitis Virus Is Required for Capsid Assembly. J. Virol., **70**, 7085-7091.

Yuan, T.T.T., Faruqi, A., Shih, J.W.K. and Shih, C. (1995) The Mechanism of Natural Occurrence of Two Closely Linked HBV Precore Predominant Mutations. Virology, **211**, 144-156.

Yuh, C-H. and Ting, L-P. (1990) The Genome of Hepatitis B Virus Contains a Second Enhancer: Cooperation of Two Elements within This Enhancer Is Required for Its Function. J. Virol., **64**, 4281-4287.

Yuh, C-H. and Ting, L-P. (1993) Differentiated Liver Cell Specificity of the Second Enhancer of Hepatitis B Virus. J. Virol., **67**, 142-149.

Yuh, C-H., Chang, Y-L. and Ting, L-P. (1992) Transcriptional Regulation of Precore and Pregenomic RNAs of Hepatitis B Virus. J. Virol., **66**, 4073-4084.

Yuki, N., Hayashi, N., Kasahara, K., Ueda, K., Fusamoto, H. and Kamada, T. (1990) Detection of Antibodies against the Polymerase Gene Product in Hepatitis B Virus Infection. Hepatology, **12**, 193-198.

Zhang, P., Raney, A.K. and McLachlan, A. (1992) Characterization of the Hepatitis B Virus X- and Nucleocapsid Gene Transcriptional Regulatory Elements. Virology, **191**, 31-41.

Zheng, J., Schödel, F. and Peterson, D.L. (1992) The Structure of Hepadnaviral Core Antigens. Identification of free thiols and determination of the disulphide bonding pattern. J. Biol. Chem., **267**, 9422-9429.

Zhou, S. and Standring, D.N. (1991) Production of Hepatitis B Virus Nucleocapsidlike Core Particles in *Xenopus* Oocytes: Assembly Occurs Mainly in the Cytoplasm and Does Not Require the Nucleus. J. Virol., **65**, 5457-5464.

Zhou, S. and Standring, D.N. (1992a) Hepatitis B virus capsid particles are assembled from core-protein dimer precursors. Proc. Natl. Acad. Sci. USA, **89**, 10046-10050.

Zhou, S. and Standring, D.N. (1992b) Cys Residues of the Hepatitis B Virus Capsid Protein Are Not Essential for the Assembly of Viral Core Particles but Can Influence Their Stability. J. Virol., **66**, 5393-5398.

Zhou, S., Yang, S.Q. and Standring, D.N. (1992) Characterization of Hepatitis B Virus Capsid Particle Assembly in *Xenopus* Oocytes. J. Virol., **66**, 3086-3092.

Zlotnick, A., Cheng, N., Conway, J.F., Booy, F.P., 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. Biochemistry, **35**, 7412-7421.

Zoulim, F. and Seeger, C. (1994) Reverse Transcription in Hepatitis B Viruses is Primed by a Tyrosine Residue of the Polymerase. J. Virol., **68**, 6-13.

Zoulim, F., Zhang, X., Pichoud, C. and Trepo, C. (1996) Heterogeneity of hepatitis B virus (HBV) core gene in a patient with HBV-associated cirrhosis and serum negativity for anti-HBc. J. Hepatol., **24**, 155-160.

GLASGOW UNIVERSITY LIBRARY