IMMUNOREACTIVITY OF HEPATITIS B SURFACE ANTIGEN

by

PHILIP GEORGE ASHTON-RICKARDT (BSc)

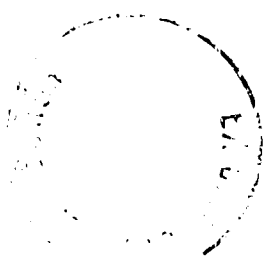
Thesis submitted for the degree of Doctor of Philosophy

Department of Molecular Biology

Faculty of Science

University of Edinburgh

1988



DECLARATION

I declare that, apart from the assistance acknowledged, the research described was performed by myself and this thesis is of my own composition.

July 1988

To my parents and Manuelle.

CONTENTS

ABSTRACT

The nature of the immunoreactivity of hepatitis B surface antigen (HBsAg) has been examined by mutation of specific amino acid residues. The replacement of arginine at 122 for lysine, or replacement of tyrosine at 134 for phenylalanine, when combined with the substitution at 113 (serine → threonine) switched the antigenic subtype of HBsAg from $\underline{y}^+\underline{d}^-$ to $\underline{y}^+\underline{d}^+$. However, differences in antigenic subtype were not attributable to any single amino acid substitution. The effect of these mutations (most notably at 122) on the subtype of anti-HBsAg sera was more dramatic, a combination of all three mutations facilitated the complete switch of immunogenic subtype from $\underline{y}$ to $\underline{d}$. Proline 142 was required for the exhibition of full antigenicity of the immunodominant $\underline{a}$-region. Replacement of cysteines 124 and 147 by serines drastically reduced or eliminated antigenicity, which implicates these two residues in stabilising the conformation of the antigen. The $\underline{a}$-region of HBsAg has also been shown to influence both the immunoreactivity of the adjacent subtype antigenic region, despite being immunologically distinct from it, and the ability of the antigen to interact with a subtype-specific monoclonal antibody. These results emphasise the importance of the polypeptide region between the cysteine residues 124 and 147 in determining $\underline{a}$-antigenicity as well as manifestation of the subtype of HBsAg.

ACKNOWLEDGEMENTS

I wish to thank Professor Kenneth Murray for his guidance and support throughout my PhD. I am also indebted to the members of his laboratory - Sandra Bruce, Heather Houston, Elaine McGlynn, Ron Jackson and Michael Rossner who have helped me in every possible way. Special thanks go to Tam Bruce for his fermentation of yeast; David Tannahill, Nik Somia, Steve Jackson and Michael Sullivan for their invaluable discussion and help. I thank all members of the media room for their help, and especially Margaret Anderson for her patience in preparing media for me.

For assistance with the immunological aspects of my PhD, I am grateful for the help provided by Dr J Peutherer, Stuart Black and Glynis Leadbetter of the Department of Bacteriology, Edinburgh University Medical School. I also thank Fiona Govan for typing my thesis in the face of my appalling writing, Mairead Blair and Annie Wilson for the drawing of figures, and Graham Brown for photographs.

I am also grateful to the Green Cross Corporation and Biogen Inc for their gifts of antisera and antigen.

Financial support for myself was provided by the Science and Engineering Research Council and the Roseanne-Campbell Hepatitis Research Trust.

ABBREVIATIONS

ATP	adenosine triphosphate
Bisacrylamide	N,N'-methylene bisacrylamide
bp	base pair
BSA	bovine serum albumin
CP	core gene promoter
CTP	cytosine triphosphate
cpm	counts per minute
ddNTP	dideoxynucleotide triphosphate
dH_2O	distilled H_2O
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
DR1	direct repeat one
DR2	direct repeat two
dsDNA	double-stranded DNA
EDTA	ethylene diaminetetra-acetate
g	acceleration due to gravity ($9.8 m/s^2$)
GTP	guanine triphosphate
kb	kilo base
kD	kilo dalton
LTB	low Tris buffer
mRNA	messenger RNA
NF-1	nuclear factor 1
NP40	nonidet P40
ORF	open reading frame

PBS	phosphate-buffered saline
RIP	radioimmunoprecipitation
RF	replicative form
RNA	ribonucleic acid
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
SSC	standard saline citrate
TCA	Trichloroacetic acid
Tm	melting temperature
Tris.HCl	tris (hydroxymethyl) aminomethane hydrochloride
Triton X-100	octylphenolpoly (ethyleneglycolethan)
TTP	thymidine triphosphate
XP	X gene promoter

All other abbreviations deined in the text.

Amino acid symbols:

Amino acid	Three-letter symbol	One-letter symbol
Alanine	Ala	A
Arginine	Arg	R
Aspargine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E

Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

1. INTRODUCTION

1. INTRODUCTION

1.1 General

The association of a viral particle with the disease state of hepatitis B was made by Blumberg in 1967, when the so-called Australia antigen (HBsAg) (Blumberg, 1965) was implicated with acute hepatitis B in a number of patients. Since the discovery of the viral pathogen responsible for the human liver disease of hepatitis B analogous agents have been discovered to cause similar conditions in the woodchuck (Summers et al., 1978), the beechy ground squirrel (Marion et al., 1980) and the Pekin duck (Mason et al., 1980). Along with the Hepatitis B virus (HBV) of humans these animal viruses comprise the Hepadna group of viruses, their taxonomy being based on similarities in viral particle structure, genome organisation, and disease characteristics.

1.2 Biology of Hepatitis B Virus

1.2.1 Hepatitis B virus particles

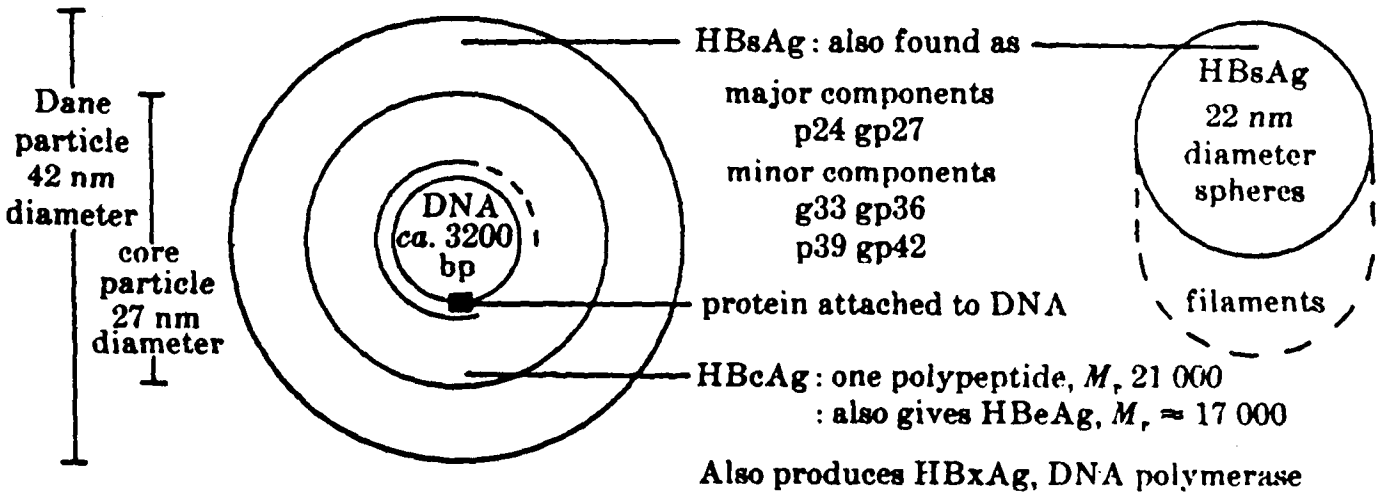
The HBV virion, known as the Dane particle, appears in infected sera as a spherical 42nm diameter particle (Dane et al., 1970). It has an inner core of 27nm (containing virus genome and associated core antigen), and an outer envelope of 7nm thickness containing the surface antigen (HBsAg) (Cossart, 1971; figure 1.2.1). This complex glycosylated lipoprotein is also present in the serum of infected

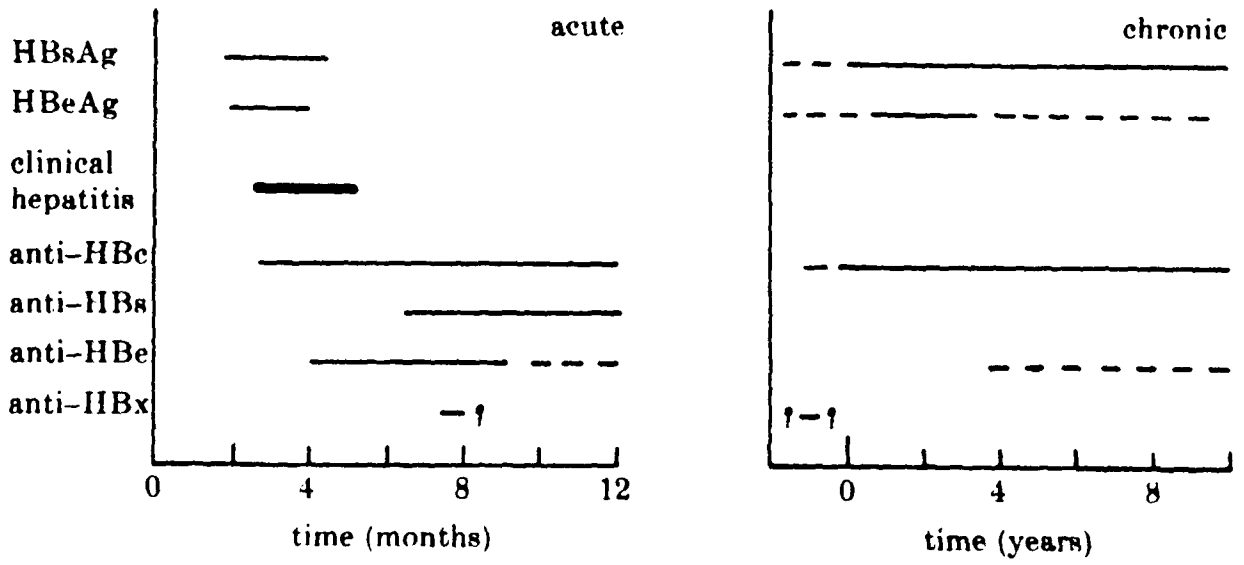
Figure 1.2.1 Structure of the HBV virion (from Murray, 1987)

The Dane particle (42nm in diameter) is composed of an outer envelope and a nucleocapsid core (27nm in diameter). HBsAg is located in the outer envelope along with lipid; present in the serum of infected individuals are spherical particles composed of HBsAg of diameter 22nm and filaments of similar diameter. The nucleocapsid is composed of the core antigen (HBcAg) and the partially double stranded DNA genome of the virus. Also associated with the core particle are the X antigen (HBxAg), the e antigen (HBeAg) and a DNA polymerase activity.

Figure 1.2.3 Immunological profiles of hepatitis B disease states (from Murray, 1987)

The temporal progression of acute and chronic hepatitis B indicated by serological markers is shown.


Dane particle 42 nm diameter
core particle 27 nm diameter
DNA ca. 3200 bp
HBsAg : also found as
major components
p24 gp27
minor components
g33 gp36
p39 gp42
protein attached to DNA
HBcAg : one polypeptide, M_r 21 000
: also gives HBeAg, $M_r \approx$ 17 000
Also produces HBxAg, DNA polymerase
HBsAg
22 nm diameter spheres
filaments

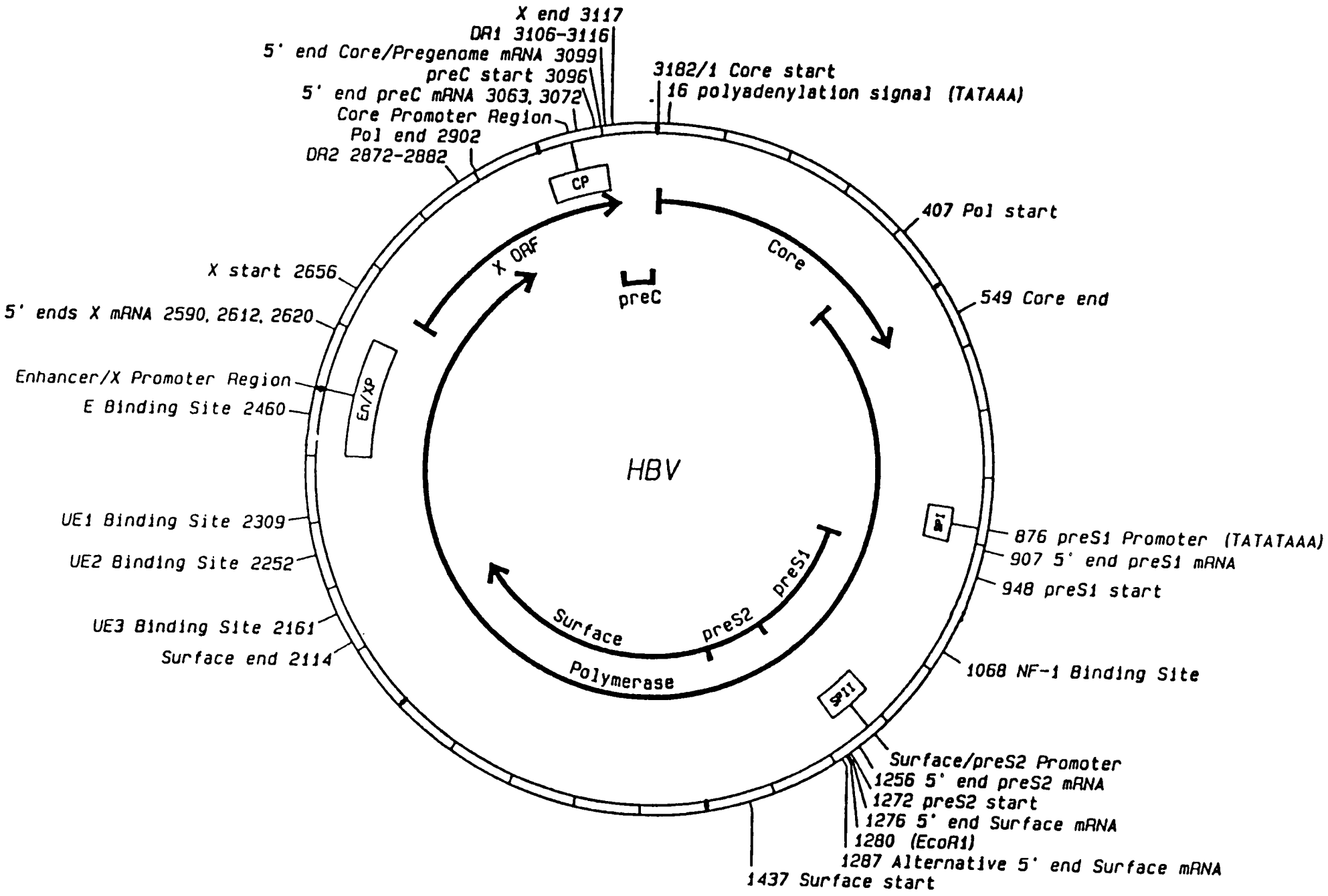

acute
HBsAg
HBeAg
clinical hepatitis
anti-HBc
anti-HBs
anti-HBe
anti-HBx
0
4
8
12
time (months)
chronic
0
4
8
time (years)

individuals as spherical particles with a diameter of 22nm or filaments of similar diameter. HBsAg exists mainly as a 24kD (p24) or 27kD (gp27) glycosylated derivative in 22nm particles and Dane particles (Peterson et al., 1977). There are however larger less abundant forms, namely the 33kD glycoprotein, gp33, and the 36kD glycoprotein, gp36, present in 22nm particles and Dane particles, as well as other larger forms such as gp42 and p39, and minor, very high molecular weight polypeptides (Heermann et al., 1984). The inner nucleocapsid contains the core antigen (HBcAg), which is composed of a major polypeptide of molecular weight approximately 23kD. Also found in infected sera is the e antigen (HBeAg), which is a soluble component of the nucleocapsid and can be produced in vitro by proteolytic cleavage of HBcAg when dissociated (MacKay et al., 1981a). The nucleocapsid contains the virus DNA, a DNA polymerase activity (Kaplan et al., 1973; Robinson and Greenman, 1974) and a protein kinase which can phosphorylate the major core protein. On the basis of amino acid sequence deduced from the nucleotide sequence of its gene, HBcAg has a protamine-like C-terminus and since it is located in the nucleocapsid is likely to be a DNA binding protein.

The HBV genome has a characteristic and unusual partially double stranded form (for a review see Robinson, 1977) (figure 1.2.2). The single stranded portion varies in length from approximately 15 - 50% of the complete strand with a preferred minimum length of 650 - 700 nucleotides. The short strand has a variable 3' and fixed 5' position relative to the long strand. The long strand is 3.2kb long

Figure 1.2.2 HBV genome structure and transcriptional signals (drawn by R Jackson)

Open reading frames (in three translational phases) are indicated by heavy lines. Promoter and enhancer sequences are denoted by rectangular boxes. The numbering system is based on position 1 being the start of the core ORF.


X end 3117
DR1 3106-3116
5' end Core/Pregenome mRNA 3099
preC start 3096
5' end preC mRNA 3063, 3072
Core Promoter Region
Pol end 2902
DR2 2872-2882
3182/1 Core start
16 polyadenylation signal (TATAAA)
CP
407 Pol start
X start 2656
5' ends X mRNA 2590, 2612, 2620
X ORF
preC
Core
549 Core end
Enhancer/X Promoter Region
E Binding Site 2460
En/Xp
HBV
UE1 Binding Site 2309
UE2 Binding Site 2252
UE3 Binding Site 2161
Surface end 2114
Surface
preS2
preS1
Polymerase
876 preS1 Promoter (TATATAAA)
907 5' end preS1 mRNA
948 preS1 start
1068 NF-1 Binding Site
SP II
Surface/preS2 Promoter
1256 5' end preS2 mRNA
1272 preS2 start
1276 5' end Surface mRNA
1280 (EcoR1)
1287 Alternative 5' end Surface mRNA
1437 Surface start

and has a nick located at a unique site which is situated approximately 200 nucleotides from the 5' end of the short strand, where a protein of unknown function is covalently attached (Gerlich and Robinson, 1980).

1.2.2 Clinical aspects of hepatitis B

Hepatitis B is a worldwide public health problem of immense importance, with some 2 x 10^8 carriers of the virus in the world. The HBV virus can be transmitted by saliva, semen and blood in a horizontal or vertical fashion and exerts it pathogenic effects in humans and chimpanzees by infection of the liver with concomitant hepatocyte damage. A particular feature of HBV, and of all hepadna viruses, is that infection exhibits great tissue and host specificity occurring only in the liver of humans and chimpanzees. Infection may result in the acute form of the disease or develop into longer-lived chronic hepatitis B. The latter form can give rise to substantial liver cirrhosis and is believed to represent 5 - 10% of all HBV infections; chronic infection may also lead to a carrier state characterised by the persistence of viral antigens in the patient's serum in the absence of other pathological features. This may ultimately lead to viral DNA integration into the host chromosome and even to the development of hepatocellular carcinoma or primary liver cancer (Vyas et al., 1978; Szmuness et al., 1982).

The serological profiles of acute and chronic disease states are represented in figure 1.2.3. The initial component detectable in

serum is HBsAg, which is usually produced in large amounts, followed by HBeAg; the other nucleocapsid antigen HBcAg is not normally detected in infected sera. Both of the serum antigens appear just before the onset of detectable hepatocyte necrosis, manifested by serum transaminase activity, and patient jaundice. However, antibodies to viral components including those against HBcAg appear at the same time as the manifestation of tissue damage. If the individual develops the chronic form both HBsAg and HBeAg may persist in sera, furthermore upon integration of HBV DNA HBsAg may be produced by active transcription of integrated viral genes long after the disappearance of Dane particles from the serum.

Vaccines against HBV have been developed from HBsAg, either purified from infected sera in the 22nm particle form (for review Maupas and Guesry, 1981; Szmuness et al., 1982) or produced by recombinant DNA methods (Murray et al., 1984). The vaccine based on serum derived HBsAg has been proven safe and effective and is used widely (Szmuness et al., 1980). Due to the problems associated with the high cost of the serum derived vaccine plus the necessary purification and testing procedures, and fear of HIV contamination, the rationale for producing HBsAg by expression of the gene in microbial systems (S. cerevisiae) by genetic engineering was born. The antigen produced behaves immunologically similarly to the native protein and also forms 22nm-like particles on expression in yeast (Murray et al., 1984; Miyanohara et al., 1983; Valenzuela et al., 1982; Hitzeman et al., 1983). The recombinant protein also protects chimpanzees challenged with HBV, and has now been approved for general use. By

using HBsAg as a vaccine we assume that prophylaxis is afforded through priming of the immune system by the antigen, before viral challenge. Recent work has revealed the importance of pre S protein domains in facilitating virus uptake into hepatocytes (Neurath et al., 1986b). Therefore a new generation of candidate pre S/S antigen vaccines is being produced by genetic engineering techniques (Valenzuela et al., 1985a, b).

1.2.3 HBV genome organisation

The complete nucleotide sequence of the HBV genome was established following its cloning in Escherichia coli (Pasek et al., 1979; Galibert et al., 1979) (Appendix A). It shows that the organisation of genetic information is compact, with the presence of 4 open reading frames in all 3 translational phases of the 3.2kb long strand (minus strand) and one small potential gene (700HBV) in the short strand (positive strand). The predicted locations of the genes for HBV components in terms of the 4 open reading frames (ORF) were established by cloning, sequencing and expressing restriction fragments of the genome in E. coli and comparing nucleotide sequence and expression product with the native HBV protein (figure 1.2.2).

HBsAg: The HBsAg gene was mapped to nucleotide positions 1437 - 2114 by comparing the predicted amino acid sequence of the ORF with the partial amino acid sequence of serum derived HBsAg (Peterson et al., 1977). The S gene gives rise to the p24 protein and its glycosylated form gp27 (Peterson et al., 1977; Peterson, 1981; Skelly et al.,

1978). The larger forms, gp33 and gp36, represent S gene and pre S2 translation products, the largest forms,present in consistently small amounts, are the p39 and gp42 proteins which represent pre S1-pre S2-S gene translation products. The coding basis of these forms of HBsAg protein was deduced using monoclonal antibody probes, specific for synthetic peptides representing the different domains, to detect the presence of these regions in the various forms of the protein (Heermann et al., 1984). Other evidence implicating the pre S region in HBsAg proteins derives from the observation that secretion of gp33 and gp36 from HBV DNA transfected, L-cells is abolished by altering the reading frame of the HBV DNA 10 nucleotides downstream of the pre S2 ATG, whereas p24 and gp27 secretion is unaffected (Persing et al., 1985).

HBcAg: The core gene was mapped to nucleotide positions 1 - 549 by expression of the antigen in E. coli followed by immunological characterisation of the recombinant protein (Pasek et al., 1979) and electromicroscopic characterisation of the core particles produced (Cohen and Richmond, 1982).

Ninety nucleotides upstream of the HBcAg gene start there is an in-phase initiation codon. The region between the 2 initiation points, denoted the pre-core region, encodes largely hydrophobic residues and is conserved between all the members of the Hepdna viruses (Moroy et al., 1985; Enders et al., 1985). HBeAg represents a proteolytic cleavage product of HBcAg (MacKay et al., 1984a), with the loss of between 34 and 36 amino acids from the C-terminus of core during e-

antigen biosynthesis (Takahashi et al., 1983). It has been shown that the pre-core sequence is necessary for HBeAg production and secretion but not for HBcAg (McLachlan et al., 1987; Roosinicck et al., 1986; Ou et al., 1986; Vy et al., 1986). The mRNA utilised is the low abundance core promoter transcript initiating 5' to pre-core giving rise to pre-core/core messages (McLachlan et al., 1987). The processing of pre-core/core protein into secretable HBeAg involves the endoplasmic reticulum (ER) association of nascent protein via the pre-core signal sequence (Schlicht et al., 1987; Vy et al., 1986; Ou et al., 1986) and subsequent cleavage of signal when it enters the lumen of the ER. The C-terminal cleavage of core is also membrane associated allowing its conversion to HBeAg and secretion (Schlicht et al., 1987). Based on protein sequence homology it is argued that the N-terminus of core contains a protease domain, and that this protease catalyses the self cleavage of membrane associated core protein into HBeAg (Miller, 1987). As well as C-terminus removal the other post-translational modifications carried out on core by the secretory mechanisms of infected cells are membrane associated phosphorylation (Roossinick and Siddiqui, 1987) and glycosylation (Schlicht et al., 1987). The cryptic protease activity of core is activated, it is suggested, by post-translational modifications inferring that activity is dependent on the glycosylation and/or phosphorylation of membrane associated core then allowing full e-antigen conversion (Miller, 1987). Consideration of membrane associated core peptides as viral proteases has allowed the postulation of a retroviral like scheme for the maturation of core and polymerase gene products by proteolysis of core-pol polyproteins

produced from the translation of genomic length HBV messages. Such fusion proteins have been detected in cell culture systems (Junker et al., 1987) and primary liver carcinomas (Will et al., 1986), and are thought to be analogous to the gag-pol fusions found in retrovirally infected cells which are resolved by self cleavage by the protease region of gag (Ganem and Varmus, 1987).

Several investigators claim that as well as pre-core allowing HBeAg production, the signal also targets HBcAg destined for nucleocapsid formation to the ER, so facilitating core-envelope interactions in the assembling virus (Ou et al., 1986; Vy et al., 1986). However, in the light of the finding that mutations in the pre-core of DHBV abolishes HBeAg secretion in a Hep-G2 cell culture system but not virus assembly, it seems that pre-core is not needed for nucleocapsid formation (Schlicht et al., 1987). This result is supported by the observation that similarly mutated DHBV is still infective when injected intrahepatically into new born ducks, whereas HBeAg secretion is diminished (Chang et al., 1987). It seems therefore that nucleocapsid formation and association with HBsAg derived envelope does not require the secretory pathway, but that cores form spontaneously as they do when produced in E. coli (Stahl et al., 1982; Cohen and Richmond, 1982) or other expression systems (Murray, 1987).

HBxAg: The other putative gene present in the pre-core/core translational phase is the so-called X ORF (Tiollais et al., 1981), which was not until recently assigned a protein coding function. It

had been shown that a mRNA species of the size expected for the X message existed in a rat cell line transformed with HBV DNA, rat 2/130.4/TK4 (Gough, 1983). The X ORF has now been expressed in E. coli (Pugh et al., 1986; Meyers et al., 1986; Kay et al., 1985) and a protein of expected size and amino acid sequence produced. The recombinant protein identified X specific antibodies in HBV infected sera and produced antibodies which reacted against HBV DNA transfected cell lines (Pugh et al., 1986; Meyers et al., 1986; Kay et al., 1985). Furthermore, X specific synthetic peptides gave rise to antisera which detected HBxAg in HBV infected liver and hepatocellular carcinomas (Moriarty et al., 1985). There is evidence to support the view that X protein functions as a transcriptional activator via a trans mechanism (Twu and Schloemer, 1987; Spandau and Lee, 1988). X protein can apparently trans-activate other viral enhancers and the X ORF has been shown necessary for S gene expression from the SPII promotor in a transient cell line assay system (R Jackson, personal communication).

pol ORF: The large open reading frame occupying 80% of a single translational phase is thought to code for the DNA polymerase/ reverse transcriptase enzyme associated with Dane particles (Kaplan et al., 1973). The indirect evidence for this proposal is based on the fact that the ORF has the capacity to code for a protein of size 93kD, comparable with the size of other enzymes of similar function and that the predicted amino acid sequence bears similarities to the other reverse transcriptases (Toh et al., 1983). The possibility that the polymerase activity associated with Dane particles is due to

a host cell mammalian DNA polymerase encapsulated by the virion seems slight, since the HBV activity has different ionic requirements to that of mammalian DNA polymerases (Hess et al., 1981). Direct evidence for the pol ORF encoding such a protein comes from the expression of the protein in E. coli in the form of a β-galactosidase fusion. The recombinant protein was found to have reverse transcriptase activity and anti-sera raised against it cross reacted with sera from HBV infected chimpanzees (McGlynn and Murray, 1988).

1.2.4 Transcription of the HBV genome

Until recently the study of gene expression of HBV during its normal life cycle has been made difficult by the lack of a laboratory host or cell culture system. However, now several groups have succeeded in propagating the virus in cell culture by the transfection of Hep-G2 hepatoma cells with cloned HBV DNA (Sureau et al., 1986; Yaginuma et al., 1987). Before this system was available other approaches were used including the study of transcription of the genes of other Hepadna viruses after host infection, and HBV transcripts from infected chimpanzee livers. The establishment of cell lines expressing some HBV genes, by transfection with cloned HBV DNA in a mammalian virus expression vector or co-tranfection with a selectable marker and the use of hepatomas such as PLC/PRF/5 (Alexander et al., 1976) bearing integrated HBV sequences, has led to progress in the study of HBV gene expression. Using vector mediated transfected cell lines (Moriarty et al., 1981; Christman et al., 1982) and the PLC/PRF/5 system (Chakrabarty et al., 1980; Edman et al., 1980) the

S gene mRNA was identified as a 2.1 - 2.5kb species by hybridisation of total mRNA to radioactively labelled genome segments. An HBV DNA transformed mouse cell line was made by co-transfection with the Herpes Simplex Thymidine Kinase gene (TK) into TK⁻ recipient cells. The HBV DNA became integrated but the only viral gene expressed was that of HBsAg as a 2.1 - 2.5kb message (Dubois et al., 1980; Pourcel et al., 1982). A more useful system was created by co-transfection of TK⁻ rat 2 cells by cloned HBV DNA present as four tandemly arranged copies of the genome and TK to produce rat 2/130.4/TK4 (Gough and Murray, 1982). In this system HBsAg, HBeAg and HBxAg (Pugh et al., 1986) were detected; also by Northern blot analysis of total rat 2/130.4/TK4 mRNA four distinct HBV specific mRNAs were identified (Gough, 1983). The most abundant species was the previously detected HBsAg message, there were also two greater than genome length messages of approximate sizes 4.2 and 4.5kb, the fourth message was a transcript hybridising to the X region of the genome.

The area of sequence necessary for S gene transcription was initially defined by transfection of L-cells with DNA fragments from which deletions had been made upstream of S (Pourcel, 1982) and by mapping of the S transcript by hybridisation with a range of labelled genome fragments with total message (Gough, 1983). Between the two approaches the region 523 - 900 was identified as necessary for S transcription and a putative promoter, TATATAA identified at 876. This TATA-like promotor has also been identified by in vitro transcription studies (Rall et al., 1983). The initiation point of S

gene transcription was further mapped by S1 nuclease digestion and primer extension experiments on S message from HBV infected chimpanzee livers and rat 2 cells transformed with HBV DNA showed that transcription started at position 1256, 185 nucleotides upstream of the start of S and in the pre S1 gene (Cattaneo et al., 1983, 1984). The proposed promotor sequence at 1215 is atypical of the TATA-like RNA polymerase II promoters, but is similar to the SV40 late promotor (Brady et al., 1982), sharing the common GAAGGNA-CCTANCC sequence; also in vitro transcription studies identify the 1215 rather than the 876 sequence as the S promotor (Standring et al., 1983).

These findings are difficult to rationalise in the light of earlier work establishing the TATA-like sequence at 876 as the S gene promotor. Quantitation of transcription from SP1 (876) and SPII (1215) promoters by assaying the levels of Chloramphenicol Acetyl Transferase (CAT) activity from hybrids of SPI or SPII and the bacterial enzyme gene (Siddiqui et al., 1986) has revealed that approximately 98% of transcription occurs from SPII and 2% from SPI. Thus, S transcripts of varying lengths (2.1 - 2.5kb) are transcribed from the dominant SPII and weaker SPI promoters with only SPI transcripts coding for gp42, and p39 forms of HBsAg - the abundance of SPII derived transcripts is consistent with the fact that p24, gp27, gp36 and gp33 are the major forms of HBsAg in Dane and 22nm particles (Heermann et al., 1984).

The HBV genome contains a single polyadenylation site TATAAA at position 16 where all transcripts are eventually terminated. The element was identified by S1 nuclease and primer extension experiments (Cattaneo et al., 1983, 1984).

The greater than genome length transcripts identified as 4.2kb and 4.5kb species in the rat 2/130.4/TK4 system (Gough, 1983) and 3.5kb messages from infected chimpanzee livers (Cattaneo et al., 1983, 1984) must be produced from covalently closed HBV genomes. To be produced, RNA polymerase II must read through the polyadenylation signal at position 16 on its first round of transcription and terminate on the second cycle. The multiple initiation points of these genomic transcripts were mapped to around position 3063 by S1 nuclease and primer extension studies (Cattaneo et al., 1983) with transcription being driven by the core promotor at about position 3050. This promotor was identified by in vitro and in vivo transcription studies (Standring et al., 1983; Rall et al., 1983) as a TATA box-like sequence. It is 10 times less strong than the SPII promotor on the basis of relative message levels (Cattaneo et al., 1983). These transcripts allow therefore for pre-core/core, core, and polymerase gene translation since separate messages for each have not been found and the genome messages are not spliced (Cattaneo et al., 1983). They may also act as pre-genome RNA molecules for HBV genome replication (Summers and Mason, 1982).

The inference that the X gene has its own message (Gough, 1983), and that in the absence of HBV mRNA splicing must therefore have its own promoter, was substantiated by the experiments of Treinin and Laub in 1987. By testing elements of DNA 5' to the X gene, using the CAT gene system for assessing promoter activity and mapping the initiation of transcription by primer extension, a TATA like element at 2500 was identified. Like other mammalian viruses, HBV has an enhancer element (for review see Serfling et al., 1985), which has been mapped to around position 2500 by CAT assays on HBV DNA (Shaul et al., 1985) and by "expression selection" of HBV genome elements (Tognani et al., 1985). The presence of this enhancer element is necessary for the activity of the core and X promoters, so it therefore must play a key role in viral genome replication and virion assembly.

The enhancer element has also been found to act in a tissue specific manner, as judged by CAT assays, with maximal activity in Alexander cells as opposed to non-hepatocyte derived lines such as HeLa cells (Shaul et al., 1985). The basis of this tissue specific nature of enhancer function has been investigated, and a glucocorticoid-responsive element has been identified at approximately 2015 using CAT assays to measure dexamethasone stimulation of enhancer activity by various HBV genome fragments (Tur-Kaspa et al. (1986). The proposed element shares the consensus TGTTCT with other glucocorticoid-responsive elements such as that present in Mouse Mammary Tumour Virus (Geisse et al., 1982) and may help to account for the tissue specific regulation of HBV transcription. Also it

has been found that several liver-specific proteins bind to the HBV enhancer and upstream sequences (Shaul and Ben-Levy, 1987). The presence of some of the binding proteins is dependent on the differentiation stage of the hepatocyte, a core octamer sequence in the HBV enhancer appears also in the promotor of several liver-specific genes suggesting that the binding site and binding proteins contribute in some part to the tissue-specific expression of the HBV enhancer and hence to HBV gene expression. Furthermore, studies on transgenic mice harbouring integrated HBV sequences (Babinet et al., 1985) indicate that S gene expression varies with hormonally driven liver tissue differentiation and that HBsAg levels were 5 - 10 times higher in male mice than female mice (Farza et al., 1987). These sex or other hormone mediated regulation mechanisms may account for the epidemiological data which indicates that HBV infected males are more likely to become carriers than females (Beasley and Huiang, 1984). A summary of transcriptional signals operating in the HBV genome can be found in figure 1.2.2.

1.2.5 Replication of the HBV genome

The process by which HBV and the other Hepadna viruses replicate their genomes can be considered as involving four major steps: (i) conversion of the asymmetric DNA of Dane particles to covalently closed circular DNA (cccDNA) within the nucleus of infected hepatocytes, (ii) transcription of cccDNA into the RNA pregenome by host RNA polymerase in virus cores, (iii) the reverse transcription of the mRNA pregenome using the virally encoded enzyme from a protein

primer to yield the minus strand DNA, (iv) synthesis of the plus strand of DNA by copying the minus strand using an oligomer of viral RNA as primer (figure 1.2.4). Studies of intrahepatic DNA have revealed the presence of cccDNA forms of the viral genome (Mason et al., 1982; Weiser et al., 1983; Blum et al., 1983; Miller and Robinson, 1984) which are located in the nucleus. Upon infection of ducks with DHBV, cccDNA is the first virus specific nucleic acid to be detected (Mason et al., 1983; Tagawa et al., 1986), preceding the accumulation of viral RNA. Direct evidence implicating cccDNA as the first intermediate in replication is afforded by the observation that when DHBV cccDNA is density labelled in cultures of infected duck hepatocytes, the pattern of labelling is compatible with synthesis by reverse transcription but not by semiconservative replication (Tuttleman et al., 1986), in that the density of the product is consistent with that of a RNA-DNA hybrid as opposed to a DNA-DNA duplex.

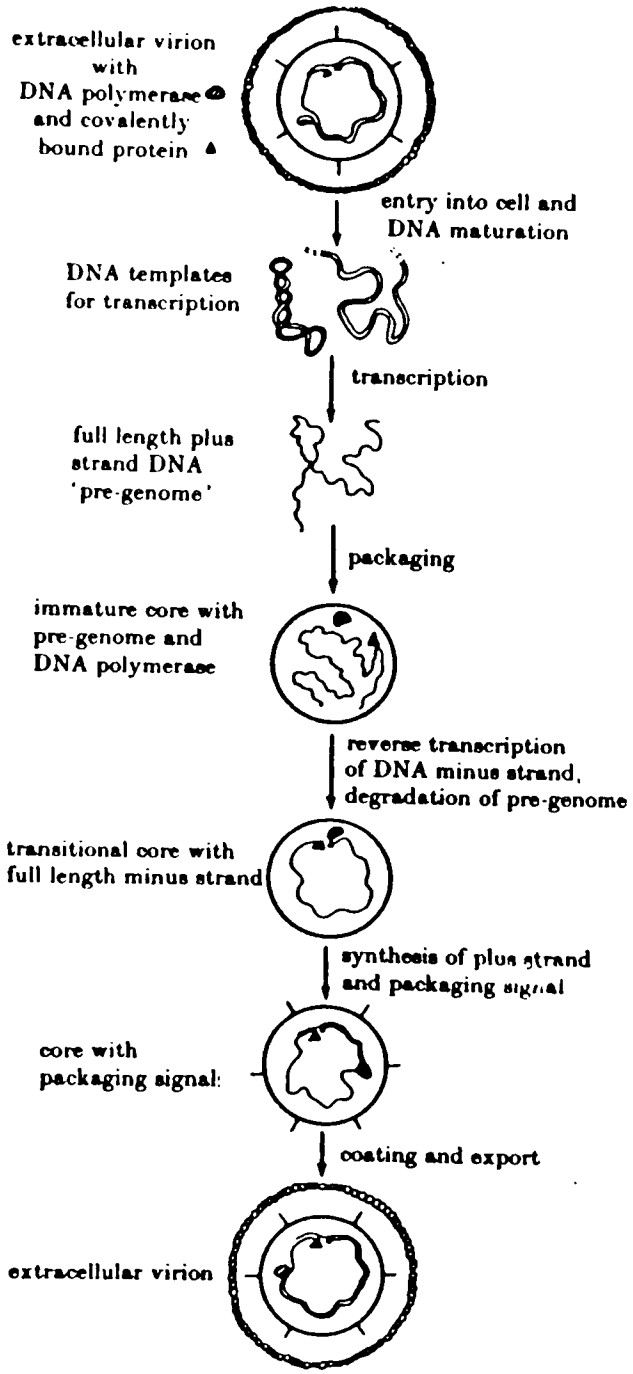
The 3.5kb genomic RNA transcriped from the core antigen promotor by RNA polymerase II serves as the template for reverse transcription. The 3' end of the transcript bears a terminal redundancy (R) which varies from 130 to 270 nucleotides among the hepadnaviruses (Ganem and Varmus, 1987), arising from the fact that the polyadenylation signal for HBV transcription is ignored for the first round of transcription, but obeyed the second time. Only one form of pre-genomic RNA is encapsulated into viral cores; the shortest form of 3.5kb RNA is found exclusively in cores separated from other 5' end derivatives of the message (Enders et al., 1985). The mechanism by

Figure 1.2.4 Replication of the HBV genome

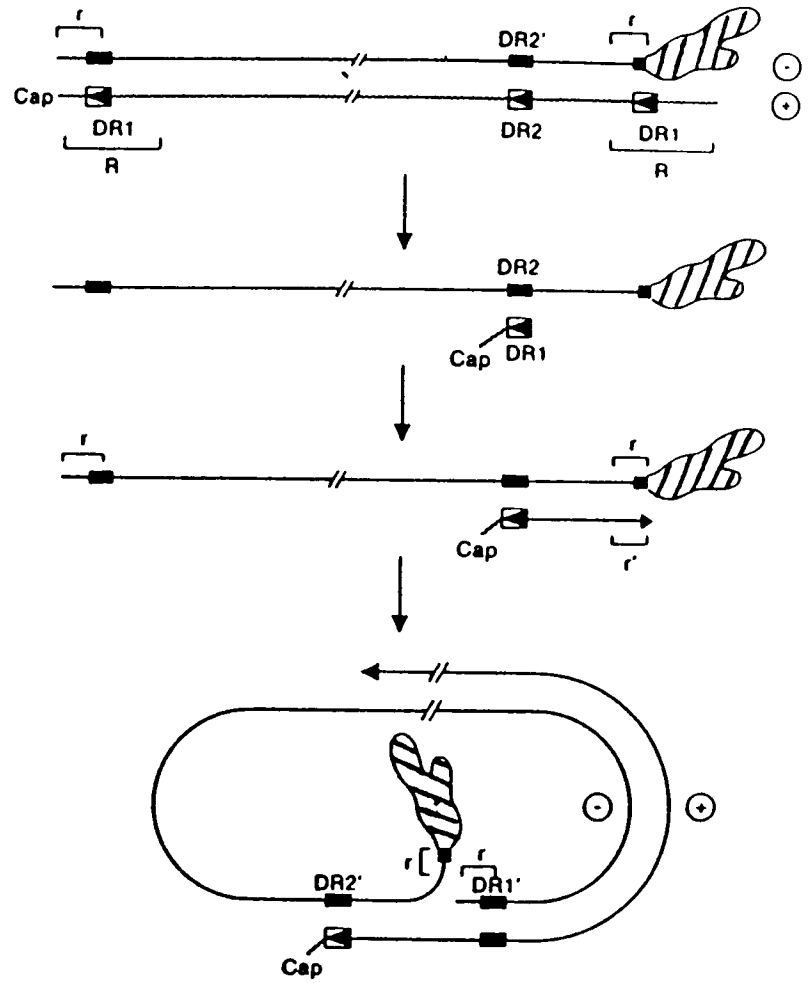
A. Overall view of the process (from Murray, 1987).

B. (From Ganem and Varmus, 1987) Model for synthesis of the second (plus) strand of Hepadnavirus DNA, showing translocation of an oligoribonucleotide containing DR1 to the position of DR2 to initiate DNA synthesis. Boxed arrowheads represent DR elements, solid rectangles represent copies of the DR elements, thin wavy lines denote RNA, straight lines denote DNA, R indicates the terminal repeat in pregenomic RNA, r indicates the short terminal repeat in full-length minus DNA. The shaded eccentrically shaped symbol represents the putative protein primer for the minus strand.

A


extracellular virion with DNA polymerase and covalently bound protein
entry into cell and DNA maturation
DNA templates for transcription
transcription
full length plus strand DNA 'pre-genome'
packaging
immature core with pre-genome and DNA polymerase
reverse transcription of DNA minus strand, degradation of pre-genome
transitional core with full length minus strand
synthesis of plus strand and packaging signal
core with packaging signal
coating and export
extracellular virion

B


r
DR2'
r
Cap
DR1
DR2
DR1
R
R
DR2
Cap
DR1
r
r
Cap
r'
DR2'
r
r
DR1'
Cap

which the sequence specific sorting of the whole 3.5kb message population is performed is unknown.

Reverse transcription as part of the replicative process is inferred by the observation that minus-strand DNA synthesis proceeds via an actinomycin D-insensitive process implying that the template is not DNA (Summers and Mason, 1982). Furthermore, duplexes produced under these conditions in sub-viral particles have a density in equilibrium centrifugation equivalent to RNA-DNA hybrids. Also, the fact that infected hepatocytes contain large quantities of unpaired minus strand DNA in cytoplasmic viral nucleocapsids suggests that HBV does not replicate via the classical semiconservative method (Summers and Mason, 1982).

The initiation site for minus strand DNA synthesis has been mapped to within the DR1 sequence by determining the position of the 5' deproteinated DNA strand (Molnar-Kimber et al., 1984; Seeger et al., 1986). It would appear likely that the 3' DR1 sequence, and not the 5' DR1, of the pregenome contains the initiation point, since elongation of (-) DNA from the 3' DR1 allows uninterrupted synthesis across the entire genomic sequence giving rise to the minus DNA with a similar terminal redundancy called r. This predicted structure has been observed in virion DNA (Seeger et al., 1986) (figure 1.2.4).

Mapping of the 5' end of plus strand DNA placed the initiation site at the 3' end of DR2 (Seeger et al., 1986; Lien et al., 1986). There is a short oligomer of viral RNA covalently attached to the 5'

end of (+) DNA, which when sequenced, revealed that the sequence was homologous to DR2 permitting priming by hybridisation to minus-strand DNA at this point. However, it was observed that the DR sequence was flanked by nucleotides from around DR1 rather than DR2 (Lein et al., 1986). By mutagenising DR1 to make it distinguishable from DR2, it was found in the GSHV system that the RNA oligomer at the 5' end of (+) strand DNA bore the DR1 sequence not DR2 (Seeger et al., 1986). To account for these observations it has been proposed that an oligomer containing DR1 from either 5' or 3' copy of R is transferred to basepair with the DR2 site in the minus strand.

Extension of the plus strand continues from DR2 of the minus strand until the protein primer at the 5' end of the template is reached. At this point the 3' terminally redundant sequence r of the minus strand displaces the 5' r sequence template allowing continuation of the plus strand elongation from the 3' end of the transferred minus strand. The resulting plus strands are of incomplete genome length and have variable end points. This model for HBV DNA replication is illustrated in figure 1.2.4.

1.2.6 HBV DNA integration and hepatoma

Although integration of HBV DNA in the livers of chronically infected patients correlates with the occurrence of hepatocellular carcinoma, any causal link between the former and the latter is yet to be established. Indeed, HBV infection is only one of several chronic inflammatory conditions associated with an increased risk of

hepatoma, others include shistosomiasis and alcoholic cirrhosis. Therefore, the fact that HBV infection can cause chronic liver inflammation may be sufficient to explain the correlation between integration and hepatoma.

The most obvious molecular feature (in the light of findings with other mammalian viruses), which would allow HBV to act directly as an oncogenic virus would be if it possessed an oncogene. However, this seems unlikely since HBV DNA integrated in tumour cells does not have one universally uninterrupted reading frame (for review see Ganem and Varmus, 1987), which one would expect of an oncogene carrying virus. Furthermore HBV, or for that matter any other Hepadna virus, does not have genes that show any homology with normal cellular genes or known oncogenes (Varmus and Swanstrom, 1985).

Integration of HBV DNA may activate cellular proto-oncogenes by insertional mutation or deregulation of expression via its enhancer in an analogous manner to c-myc activation by certain tumour viruses (Varmus, 1984). One such report involves an HBV DNA integration event such that the pre S region is fused in frame to a cellular sequence encoding a protein related to a steroid receptor and to the product of the v-erb-A oncogene (DeJean et al., 1986). The integration event characterised (denoted hap) correlated with the presence of several hepatomas and importantly gave rise to a 2.5kb transcript in the neoplastic tissues studied (de Thé et al., 1987). This suggests that the hap locus may relate in a causative manner to hepatocellular carcinogenesis.

The possibility that HBV DNA integration may contribute to oncogenesis by causing a recessive mutation in a specific region of the host genome has also been investigated. Although an integration event may cause gross mutations, for example, in the case of a 12kb deletion in chromosome 11 near p13 (Rogler et al., 1985), no mutation in a hepatoma has yet been mapped to a locus already implicated in other neoplasia (Glaser et al., 1986). Other mutations in host DNA have been detected in hepatocellular carcinomas harbouring HBV sequences, but the integrated DNA is not adjacent to the mutations, and some of the mutations are not even neoplastic. (Ganem and Varmus, 1987).

The structure and nature of integrated HBV DNA in tumours and chronically infected livers has been studied extensively. A consensus has been reached as to the structure of viral and host DNA sequences involved in the integration and in the sequence characteristics of flanking host regions. There appears to be no specificity in host DNA sequence for integration (Ganem and Varmus, 1987). However, the virus DNA appears to integrate preferentially at the sequences in or near the direct-repeats (DR), presumably using the free ends of the DNA molecules as recombination sites (DeJean et al., 1984).

1.3 HBV Envelope

1.3.1 Structure of HBsAg and outer envelope

Amino acid sequence studies of the 226 residue protein product of the S gene by Peterson (Peterson, 1977; 1981) have led to a rationalisation of the higher order structure of the protein in terms of amino acid sequence. The amino acid sequence of the S gene HBsAg product is presented in Appendix B.

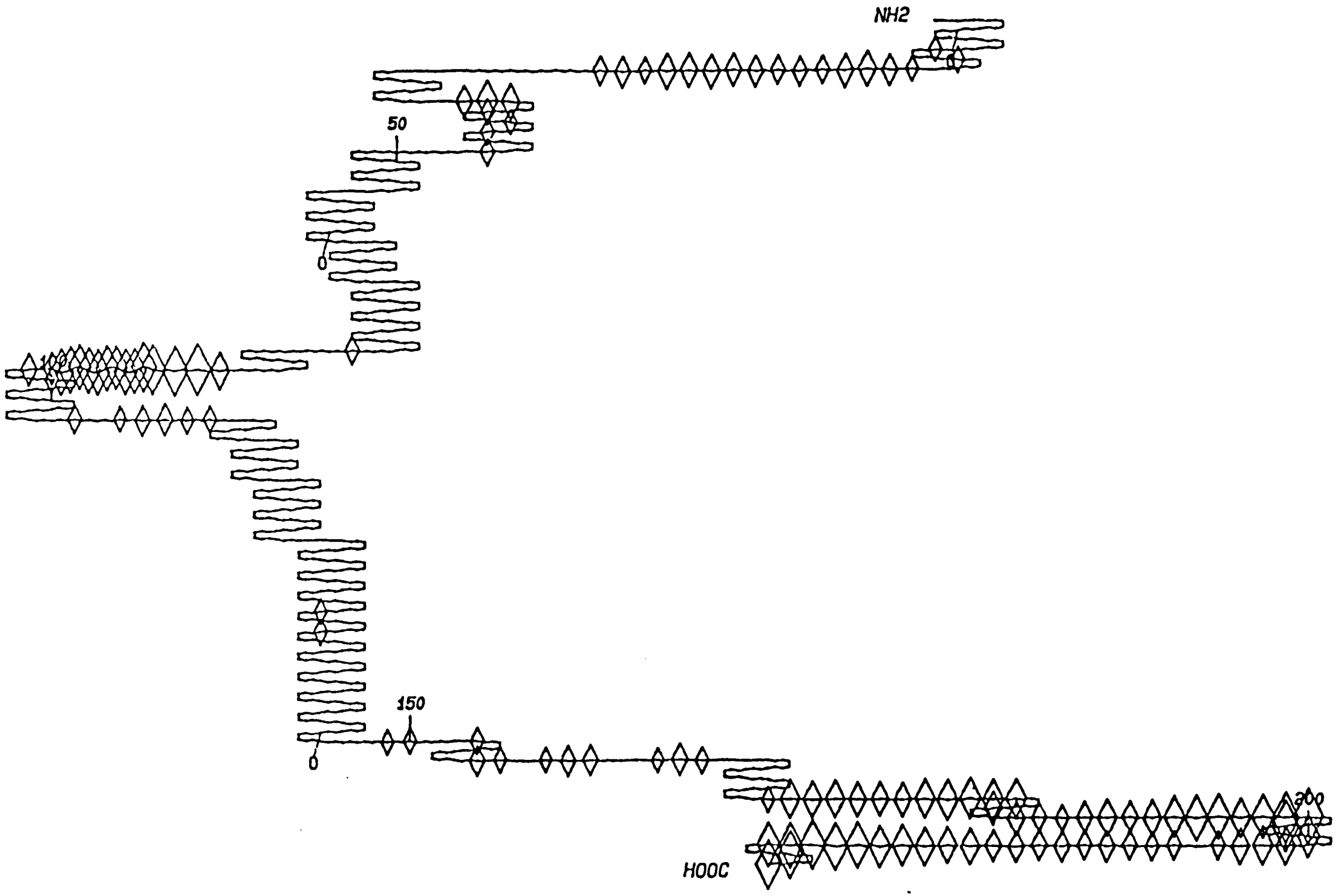
From the analysis of amino acid composition, it is apparent that there exists extensive hydrophobicity and a high proportion of proline and cysteine residues. These are consistent with early observations that HBsAg exists as a lipoprotein with complex tertiary structure (Sukeno et al., 1972; Vyas et al., 1972; Steiner et al., 1974). There are three possible glycosylation sites at aspargines number 3, 59, and 146; however, in the gp27 form of HBsAg only the 146 residue is glycosylated (Peterson, 1981).

If one analyses the secondary structure of HBsAg using computer based predictions for local hydrophilicity (Hopp and Woods, 1981) and probability of α-helix or β-sheet formation (Garnier et al., 1978) certain features are revealed (figure 1.3.1) (Appendix C). The major regions of hydrophobicity are the N- (position 1 - 45) and C-terminal domains (position 150 - 226), and in an internal domain (position 80 - 110). The relatively hydrophilic domains exist either side of the internal hydrophobic domain in positions 45 - 80 and 110 - 150. The

Figure 1.3.1 Secondary structure prediction of HBsAg

The prediction for secondary structure indicated is based on the Garnier-Osguthorpe-Robson method (Garnier et al., 1978), the prediction for hydrophobicity on the method of Hopp and Woods (1981). Both predictions were plotted using the University of Wisconsin Genetics Computer group (UWGCG) peptide structure program.

Predicted turns are represented by 180° turns in the plotted peptide sequence; the degree of hydrophobicity is indicated by the size of diamond positioned on the peptide sequence. The predicted sites of oligosaccharide attachment at 3, 59 and 146 are indicated by circles. Helices are shown with a sine wave, β-sheets with a sharp saw-tooth wave and coils with a dull saw-tooth wave.


NH2
50
0
100
150
0
200
HOOC

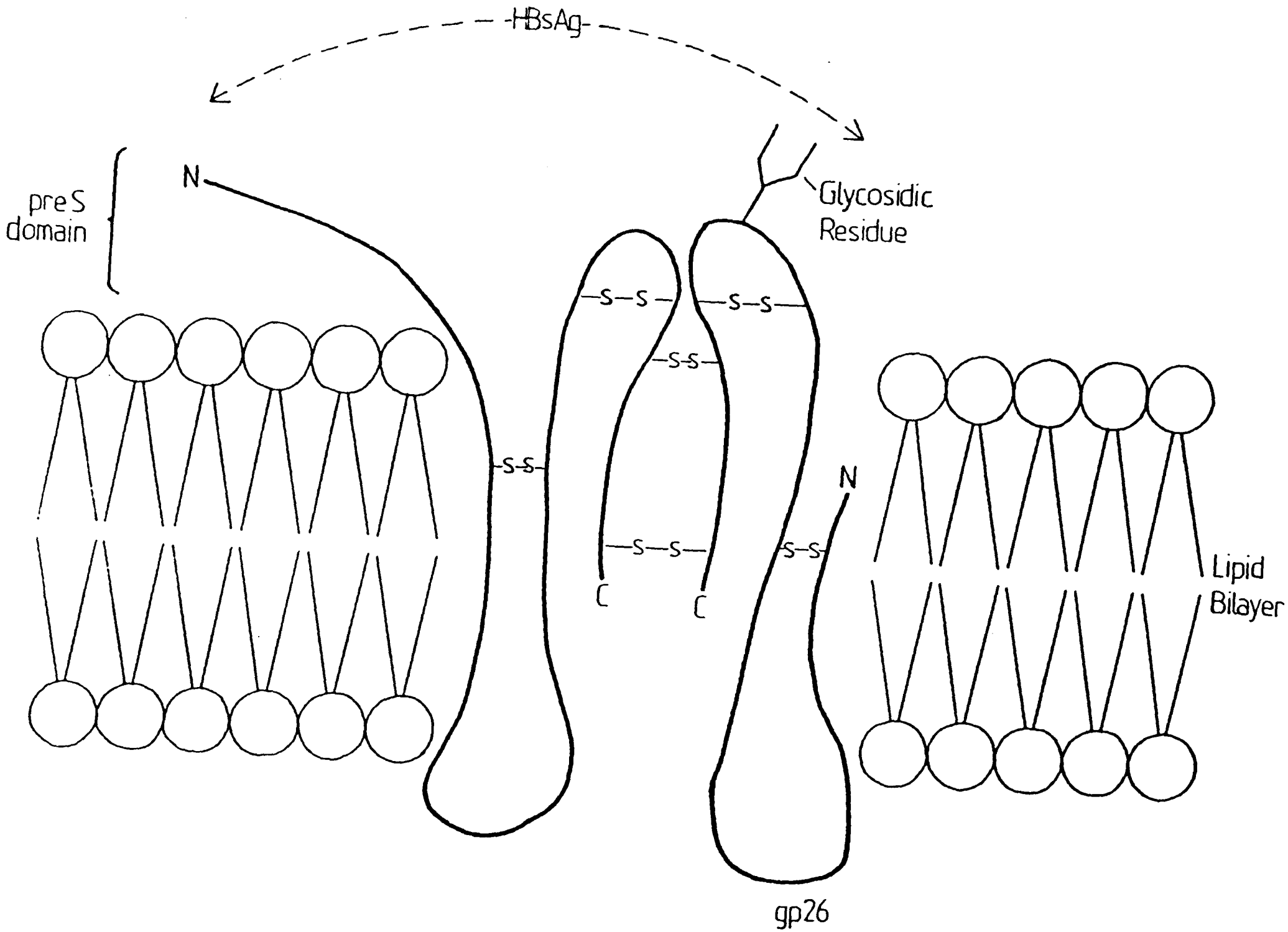
second hydrophilic domain (110 - 150) contains the glycosylation site at 146 as well as the majority of the cysteine and proline residues. It is predicted that the N- and C-termini domains contain regions disposed to α-helix and β-pleated sheet formation, whereas the two hydrophilic regions are areas of low secondary structure featuring multiple peptide chain turns dictated by prolines and glycines. If one evaluates propensity to form secondary structure on the basis of protein homology-based predictions (Coulson et al., 1987) (Appendix D) a similar picture emerges to that generated by physically based methods, in that the regions of greatest flexibility correlate with the two hydrophilic regions located either side of a hydrophobic α-helix, with the molecule well ordered in the hydrophobic N- and C-terminal domains. Thus from these various theoretical approaches it is possible to postulate that the extreme N- and C- regions of HBsAg associate with the lipid of Dane envelope or 22nm particles such that the two hydrophilic regions are situated internally or externally of the lipid envelope and in an aqueous environment, separated by a lipophilic region, making it necessary to postulate that the molecule must loop back into the membrane at this point (figure 1.3.2).

Evidence in support of these predicted features does exist. By trypsin protease mapping regions of exposed HBsAg in intact 22nm particles it was revealed that the region from 122 - 150 was exposed (Peterson et al., 1982). Also the major antigenic region, as determined by studies using synthetic peptides to locate epitopes, was proposed as mapping to this exposed area (Lerner et al., 1981).

Figure 1.3.2 Proposed orientation of HBsAg in outer envelope of HBV

The HBsAg molecule indicated is a dimer between the glycosylated S derived peptide (gp26) and a pre S containing peptide. The hypothetical cystine bridges indicated maintain the integrity of the dimer as well as the peptides themselves.

The N-terminus of gp26 is associated with the hydrophobic lipid bilayer whereas pre S domains appear to be located externally. Of the two hydrophilic internal domains, the one situated nearest the C-terminus exists in the external aqueous environment. It harbours the glycosidic residue at position 146, B-cell antigenic determinants and most of the cysteine residues. The N-terminal hydrophilic domain is situated such that it is associated with the hydrophilic interior of the virus. The C-terminus is situated with the hydrophobic lipid bilayer.


HBsAg
preS
domain
N
Glycosidic
Residue
S–S
S–S
S–S
S–S
S–S
S–S
C
C
N
Lipid
Bilayer
gp26

It would seem most logical therefore to suggest that the hydrophilic domain 45 - 80 interacts internally with hydrophilic species rather than in an exposed external fashion.

Evidence in support of the lipid association of the N- and C-terminal domains comes from studies using a photoreactive probe for lipophilic peptide sequences. It was found that both the N- and C-termini were buried in lipid but the antigenically important region 122 - 150 was exposed to the aqueous environment (Gavilanes et al., 1982).

The importance of cysteine residues in maintaining higher order structure is indicated by their importance in maintaining the antigenicity of the protein. Upon reduction and detergent treatment, HBsAg in 22nm particle becomes 1000 times less antigenic (Yyas et al., 1972; Sukeno et al., 1972; Dreesman et al., 1973) than just detergent-disrupted aggregates. Furthermore it appears as though disulphide bonds allow HBsAg to exist as a dimer of p24 and gp27 in 22nm particles which can be isolated antigenically intact and resolved into its protein components by reduction (Mishiro et al., 1980). Although the 14 cysteine residues of HBsAg do appear to play an important role in the maintenance of inter- and intra-chain conformational integrity, the fact that HBsAg produced by S expression in yeast, can form 22nm-like particles prior to the formation of cysteine mediated interactions, suggests that protein-lipid interactions are of significance also (Wampler et al., 1985). The exact nature of protein-lipid interaction has not been extensively studied, however the proportions of constituent lipids in

the envelope have been deduced: phosphotidylcholine (~60%), cholesterolester (~4%), cholesterol (~5%) and triglycerides (~35%) (Gavilanes et al., 1982).

In the forms of HBsAg with pre S domains the hydrophilic pre S 2 and pre S1 regions appear to be located in an exposed fashion in Dane particle envelopes and 22nm particles (Stibble and Gerlich, 1983; Heermann et al., 1984). The one possible glycosylation site in pre S at position 108 is occupied in the pre S2 containing gp33 and gp36 containing proteins, but not in the pre S1 - pre S2 containing gp42, whereas glycan attachment at the S gene encoded residue 146 exists in all pre S containing proteins (Heermann et al., 1984). The pre S2 domain is highly immunogenic, with antibodies occurring during natural infections against disulphide-bond independent epitopes (Neurath et al., 1984; Persing et al., 1986a). Antibodies against pre S1 domains have also been detected in the sera of HBV infected individuals (Theilmann et al., 1987) and in WHV infection (Shaeffer et al., 1986). The immunogenicity of pre S1 and pre S2 point to their relatively exposed location in the Dane particle envelope. A model for HBsAg location in the lipid bilayer of the HBV virion is presented in figure 1.3.2.

Relatively little is known about the arrangement of HBsAg dimer subunits in the outer envelope. The Dane particle envelope is believed to be composed of approximately 400 dimer subunits of diameter about 7.5nm, separated by about 10nm (Heermann et al., 1984; Aggerbeck and Peterson, 1985). The most abundant components

comprising dimers are the p24 and gp27 proteins present in similar amounts (about 130 of each) in the virion, there are about 80 gp33 or gp36 proteins and only about 60 pre S1 containing gp42 or p39 proteins per virion (Heermann et al., 1984). In the serum located HBsAg aggregates such as 22nm particles, these relative levels of protein forms are maintained with the exception of pre S1 containing peptides which are present at a very low level (Heermann et al., 1984).

1.3.2 Function and biosynthesis of HBV outer envelope

It has been proposed that the means by which HBV and the other viruses of the Hepadna group achieve their tissue and host specificity of infection is by the action of the pre S region of the envelope (Neurath et al., 1986b; Machida et al., 1984).

Imai et al. in 1979 first indicated that gp36 HBsAg had the ability to bind in vitro glutaraldehyde-polymerised human serum albumin (GpHSA), whereas p24 and gp27 forms did not. This suggested that the 55 amino acids of pre S2 had the ability to bind GpHSA. It has also been observed that only hepatocytes from HBV susceptible animals (humans and chimpanzees) have cell surfaces capable of binding GpHSA (Thung and Gerber, 1981). A pre S2 peptide of 8kd (p8) released by cleavage with cyanogen bromide has been shown to bind GpHSA, but not non-human derived forms of the polymer, and monoclonal antibodies to p8 aggregate Dane particles and also compete with binding between gp33 and GpHSA (Machida et al., 1984). A model was proposed based on

these results postulating that GpHSA acts as a bridge between HBV and target hepatocyte via specific receptors located on the hepatocyte and in the pre S2 region of the virus envelope. However, when similar experiments were repeated with naturally occurring pHSA, unlike glutaraldehyde cross-linked HSA, there was no HBsAg binding (Yu et al., 1985), casting doubt on Machida's model. There is evidence that binding of HBV to hepatocytes is direct, involving the pre S region of the envelope (Neurath et al., 1985). Using a series of non-overlapping synthetic peptides, representing the entire pre S region, to inhibit HBsAg-cellulose:Hep G2 cell interaction it was found that both pre S1 and pre S2 regions are involved in virus association, with the pre S2 region playing an auxiliary role in pre S1 mediated attachment (Neurath et al., 1986b). Given the role of pre S in virus targeting the use of anti-pre S antisera has proved successful in neutralising HBV (Neurath et al., 1986a).

It is quite evident that the S gene product itself contains all the information necessary for the correct assembly and secretion of HBsAg in the form of 22nm particles, from studies with eukaryotic expression systems (for review see Murray, 1987). From electronmicroscopic evidence it appears as though 22nm particles are secreted via the constitutive secretory pathway (Gerber et al., 1974; Patzer et al., 1986) which is also involved in Dane particle assembly (Yamada et al., 1982). The use of HBV DNA transformed cell lines producing 22nm particles has confirmed this observation (Patzer et al., 1984) and has indicated that the secretion of HBsAg is less efficient than the secretion of host proteins (Persing et al.,

1986b), with relatively slow glycan processing occurring at the rate determining Golgi apparatus stage (Patzer et al., 1984). The initial event in HBsAg secretion is insertion into the endoplasmic reticulum (ER) via an N-terminal signal sequence with nascent peptide emerging from associated ribosome into the lumen of the ER (Eble et al., 1986; Patzer et al., 1986). Soon after this an emerging hydrophobic domain is thought to anchor the peptide in the membrane with nascent protein now extending on the cytoplasmic side of the ER. The occurrence of a second internal signal sequence allows the cytoplasmically located peptide to re-enter the lumen of the ER (Eble et al., 1986; 1987) (figure 1.3.3). In order that the proposed biosynthetic mechanism be consistent with HBsAg membrane topology (figure 1.3.1), it is necessary to postulate that after the second signal sequence another anchoring domain allows the C-terminus to be associated with the ER membrane. Transmembrane HBsAg molecules are then proposed to undergo aggregation in the lipid bilayer then bud off from the membrane encapsulating HBV core-DNA particles and DNA polymerase molecules, and enter the lumen (Eble et al., 1986). Here the lipid composition of the envelope is altered from that of the original ER membrane, and envelope glycoprotein is modified in the golgi.

Although pre S sequences are not required for the membrane association of S gene translational products, it is clear that pre S1 domains inhibit the secretion of pre S1-proteins and non-pre S1 containing forms of HBsAg (Persing et al., 1986b; Standring et al., 1986; Chisari et al., 1986; Cheng et al., 1986). The point at

Figure 1.3.3 HBsAg topagenesis

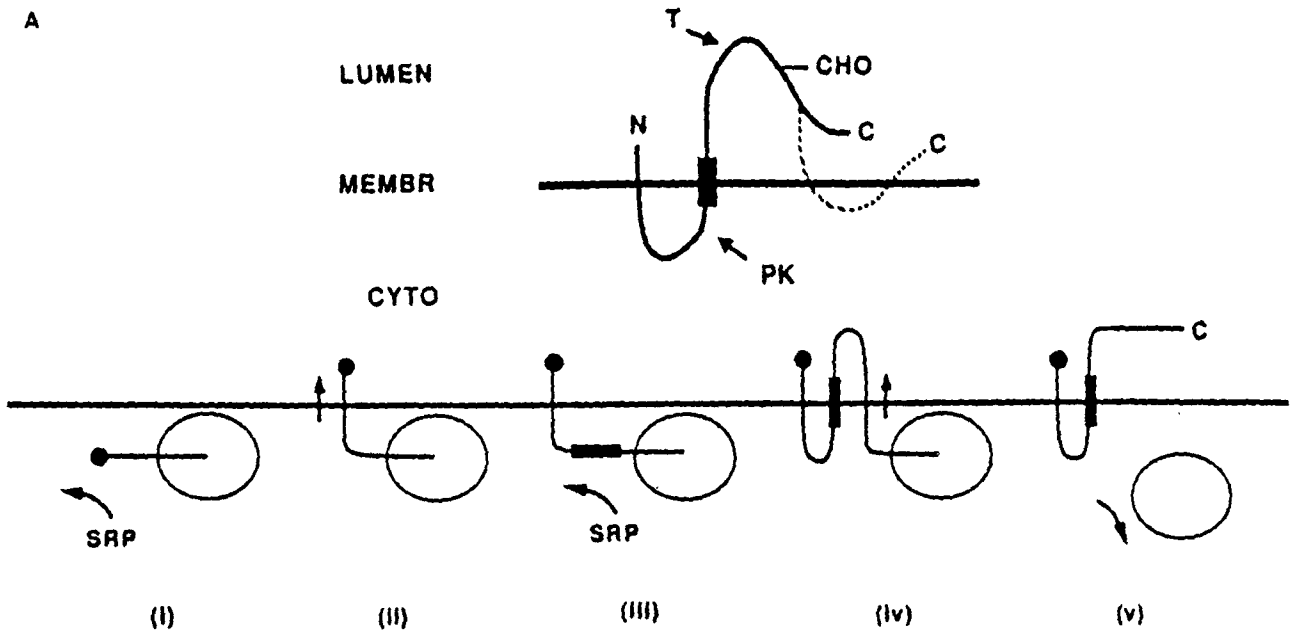
A. Model for the functioning of multiple different signal sequences (from Eble et al., 1987)

In the model, an N-terminal signal (filled circle) emerging from the ribosome (large oval) interacts with SRP (signal receptor protein (i), to initiate translocation of N-terminal domains (ii). An internal signal (filled bar) interacts with SRP (iii) and inserts within the membrane in a C-terminal end-first orientation (iv). Translocation of C-terminal domains by the internal signal (iv) completes the generation of the final tritopic configuration (v).

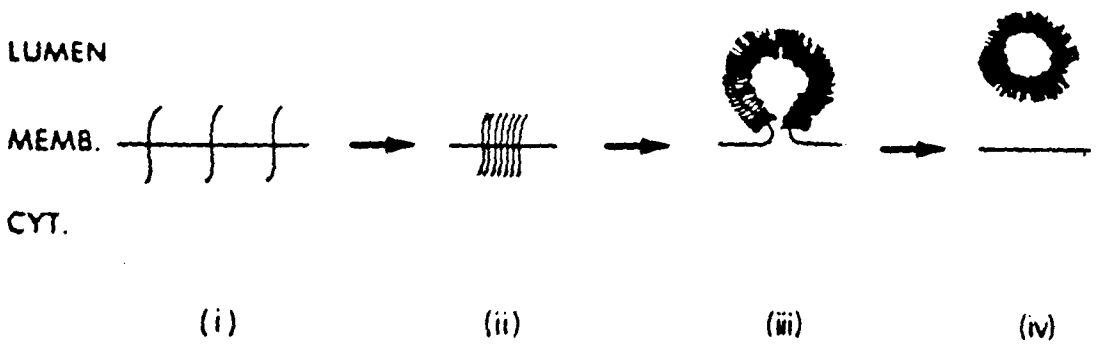
B. Model for HBsAg particle formation (from Eble et al., 1986)

In this model particle formation is envisioned to begin with the transmembrane insertion of HBsAg peptides (i), depicted (for simplicity only) in a monotopic configuration. Subsequently, monomers undergo aggregation in the plane of the bilayer (ii), concomitant with or followed by exclusion of host membrane proteins and reorganisation of host lipid components. Budding (iii) and detachment (iv) from the membrane result in the delivery of the newly formed particle into the ER lumen.

A


T
LUMEN
CHO
N
C
C
MEMBR
PK
CYTO
C
SRP
SRP
(I)
(II)
(III)
(Iv)
(v)

B


LUMEN
MEMB.
CYT.
(i)
(ii)
(iii)
(iv)

which pre S1 inhibits HBsAg secretion appears to be after the membrane insertion step, since pre S1-proteins are glycosylated. Retardation of secretion by pre S1 is dosage dependent suggesting that inhibition occurs during sub-unit mixing during Dane particle formation (Standring et al., 1986; Cheng et al., 1986). The pre S1-proteins are the only surface proteins to be acylated, harbouring covalently linked myristic acid at their N-termini (Persing et al., 1987); such a modification would explain their restricted movement through the secretory machinery with the acyl group either anchoring pre S1 in the membrane or inhibiting movement through the bilayer. Thus such restricted pre S1-protein translocation could regulate envelope (and so virion) formation and play a role in phasing envelope synthesis with other viral morphological events. The fact that HBsAg in 22nm particles is more efficiently secreted than that in Dane particles which possess far higher levels of pre S1 proteins is consistent with the proposal of a regulatory role for pre S1 (Heermann et al., 1984).

1.4 Immunology of HBsAg

1.4.1 Nature of antigen structure

The differences between the humoral (B-cell) and cellular (T-cell) mediated parts of the immune system are reflected in the nature of immunogens that trigger them. For every immunogenic particle, for example a virus, there exist different T-cell and B-cell specific epitopes (Mitchison, 1971). Viral epitopes specific to B-cells are presented to the membrane immunoglobulins directly in a

conformationally intact form (Roitt et al., 1985). In contrast T-cells (T-cytotoxic $CD8^+$ or T-helper $CD4^+$) are only triggered by non-native, intracellularly processed antigens presented on the surface of a cell with the correct Major Histocompatibility Complex (MHC) molecules (Katz et al., 1973; Ziegler and Unanue, 1981). In this way the two populations of responding cells can be triggered by different features of the same virus or even the same protein (Roitt et al., 1985). The expansion of a specific T-helper ($CD4^+$) cell clone after antigen presentation by a macrophage or B-cell (Lanzavacchia, 1985) will maximise the expansion of the B-cell clone that recognised a feature of the intact antigen. This situation has been encountered from studies of immune responses to viruses such as influenza (Townsend et al., 1985), foot and mouth disease virus (Francis et al., 1987) and HBV (Milich et al., 1987a; 1988; Vento et al., 1987). Recently two models have been proposed to identify the amino acid sequences of antigen involved in T-cell recognition. De Lisi and Berzofsky have found that a large proportion (70%) of T-lymphocytes appear to react with α-helical amphipathic structures contained within the amino acid sequences of antigen proteins (Margalit et al., 1987; De Lisi and Berzofsky, 1985). On the other hand Rothbard, McMichael and Townsend point to the existence of a four to five-amino acid sequence motif in a large number of T-cell antigenic determinants (Townsend et al., 1986; Rothbard, 1986). This motif is composed of a glycine or a charged amino acid residue in the first position, followed by 2 or 3 hydrophobic residues which are followed in turn by either a glycine, a polar, or a charged residue in the last position.

It would appear that most B-cell epitopes are conformationally determined and can be imagined as discontinuous epitopes being dependent upon the spatial arrangement of two or more peptide regions. Continuous epitopes are regions of short peptide sequence, therefore less dependent on tertiary and secondary structure, existing at the loops and protruding regions of proteins (Benjamin et al., 1984; Barlow et al., 1986). There have been few antigen-antibody complexes studied directly by X-ray diffraction, examples include epitopes of hen-egg white lysozyme (Amit et al., 1986; Sheriff et al., 1987) and of influenza virus neuraminidase (Colman et al., 1987). Each of the studies revealed discontinuous epitopes all of which covered a large proportion of the surface area of the molecule.

In an attempt to define what chemical features of a peptide region predispose it to be an epitope, Hopp and Woods (1981) analysed protein structure and correlated it to antigenicity. They concluded that the degree of exposure of a region to the external environment, assessed by a preponderance of hydrophilic residues, determines its B-cell antigenicity. Therefore, factors like hydrophilicity, degree of flexibility and probability of situation at the surface of the protein, appear to be important. The testing of these theoretically based predictions has been afforded by representing an entire protein antigen (of known structure) by short synthetic peptides and assessing the proportion of antibodies in the native protein antiserum which recognise a particular peptide. This approach has largely vindicated earlier predictions since the most reactive peptides in studies mapped to positions of maximum mobility,

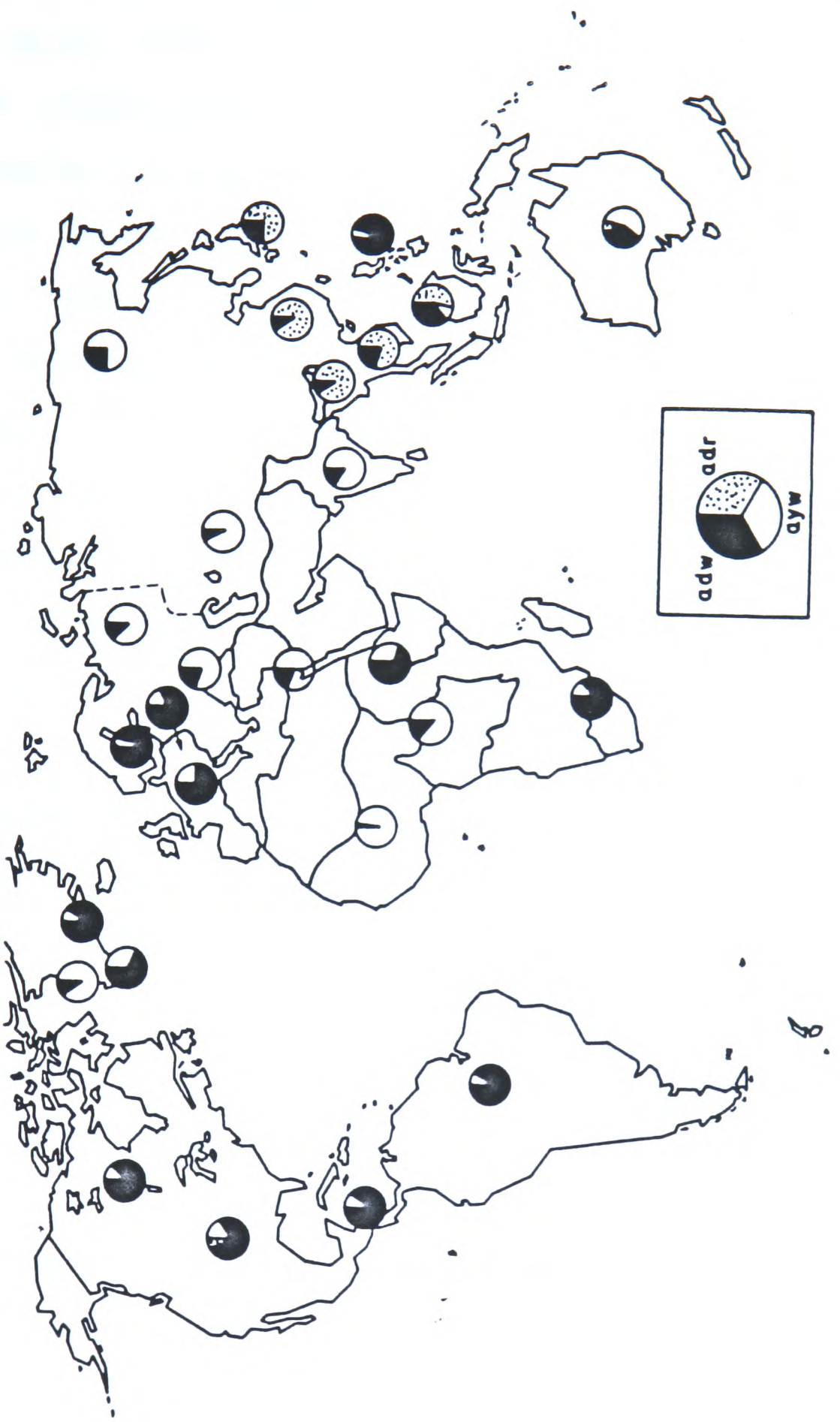
accessibility, convex shape (Geysen et al., 1987; Barlow et al., 1986; Thornton et al., 1986; Novotny et al., 1986). Surfaces which displace bound water, such as those of negative electrostatic charge and convex shape, reduce stearic hindrance to antibody-epitope binding (Geysen et al., 1987). The peptide mapping approach to epitope definition is biased in favour of continuous epitopes, since it is always point specific and unlikely to locate the larger discontinuous epitopes (Benjamin et al., 1984). The use of site-directed mutagenesis (SDM) (Zoller and Smith, 1983) in conjunction with molecular modelling to test changes in predicted structure on antigenicity allows an alternative approach to epitope definition (Rees and de la Paz, 1986). This approach was used to test the predictions of area size and necessary electrostatic charge of an epitope of hen-egg white lysozyme; the results obtained confirmed the size of the epitope and cast doubt on the importance of electrostatic charge in the region, since substitution of amino acids comprising the epitope with ones of different charge had little effect (Roberts et al., 1987).

1.4.2 Immunochemistry of HBsAg

Serologically, HBsAg has one group specific determinant a and 2 sets of mutually exclusive determinants d or y (Le Bouvier, 1971) and w or r (Bancroft et al., 1972) giving 4 serotypes adw, ayw, adr and ayr. These strain specific serotypes show a stable geographical distribution with ayr subtype occurring very rarely and only in Asia (Sobeslavsky, 1978) (figure 1.4.1). In a natural infection most of

Figure 1.4.1 Geographic distribution of HBsAg subtypes (from Sobeslavsky, 1978)

The relative proportion of total HBsAg as subtype adw, ayw or adr is indicated by the proportion of pie chart. HBsAg of subtype ayr occurs very rarely and only in Asia.


adr
awp
adw
awp

the antibodies are a-specific with the subtypic population contributing little to protective immunity (Murphy et al., 1974; McAuliffe et al., 1980). Using the antigenicity predictions of Hopp and Woods (figure 1.3.1) in conjunction with synthetic peptide epitope mapping, the a-group area has been located between amino acid residues 138 and 149 (Prince et al., 1982), 139 and 147 (Bhatnagar et al., 1982). The a-group is composed of more than one epitope since panels of non-cross reacting a-specific monoclonal antibodies exist (Tedder et al., 1984; Swenson et al., 1988). The majority of the a epitopes are conformationally sensitive and therefore likely to be discontinuous since perturbations in tertiary structure caused by reduction or detergent treatment drastically reduce antigenicity (Vyas et al., 1972; Sukeno et al., 1972; Dreesman et al., 1973). Furthermore, although the linear peptides that identified the a-region reacted with a-antiserum the antigenicity was far less than that of native HBsAg (Bhatnager et al., 1982). The only attempts that have been made to postulate possible conformations have involved using a-region peptides cyclised about specific cysteine residues to test whether these represent native conformations (Brown et al., 1984; Dreesman et al., 1982). In both cases cyclisation about cysteines 139 - 147 (Brown et al., 1984) and 124 - 137 (Dreesman et al., 1982) increased antigenicity, but not to the level of native antigen. Disulphide bridge contribution to a-epitope structure has also been demonstrated by the production of monoclonal antibodies to the disulphide-bond dependent a-epitopes (Swenson et al., 1988; Antoni and Peterson, 1988). The involvement of the glycan group at position 146 was discounted from contributing to a-antigenicity upon

its removal without loss of reactivity to anti-a-sera (Skelly et al., 1978).

The contribution of the cell mediated immune system to the production of anti-HBsAg antibodies was first demonstrated by Roberts et al. (1975), this T-cell priming has been shown to be MHC class II restricted (Milich et al., 1984). A short peptide region of HBsAg (24 - 27) has been identified as a possible MHC-restricted, T-helper cell and cytotoxic cell determinant (Celis et al., 1988), antibodies primed by the region were specific for HBsAg. The sequence of the peptide described was comparable to those proposed as T-cell epitopes by De Lisi and Berzofsky (1985) and Rothbard and collaborators (Townsend et al., 1986; Rothbard, 1986; section 1.4.1). Also a single 10-residue pre-S1 peptide (12 - 21) can prime T-cell help for antibody production to multiple epitopes within the pre-S1, pre-S2 and S regions of HBsAg under H-2 restriction in mice (Milich et al., 1987c).

The nucleocapsid (HBcAg) can induce antibody responses via a T-cell independent pathway (Milich et al., 1986) and is highly immunogenic during HBV infection. This is consistent with the observation that HBcAg can protect chimpanzees from HBV challenge (Murray et al., 1984; 1987). The core antigen can also prime B-cells for anti-HBsAg antibody production by activation of T-helper cells in a MHC class II restricted fashion (Milich et al., 1987a; 1988). The use of synthetic peptides has revealed that a region from residues 85 to 140 determines the T-cell reactivity of HBcAg with different epitopes

within the region being specific for different haplotypes. One particular epitope, contained in a 120 - 140 peptide, was capable of priming anti-HBcAg, anti-HBsAg and anti-Pre S antibody production in a MHC class II restricted fashion (Milich et al., 1987b). As well as activating T-helper cells HBcAg is also capable of activating T-cytotoxic cells through its expression on the surface of infected hepatocytes and recognition in conjunction with MHC class I proteins (Mondelli et al., 1982)

1.4.3 HBsAg subtype

A compound subtype of an HBV serum represents a situation where both subtypes, of a mutually exclusive pair, are present in the serum of an infected individual (Nonderfelt and Le Bouvier, 1973; Mazzur et al., 1975). These compound subtype sera, which are present at very low levels, may be due to infection by more than one virus of single subtype, giving rise to the phenotypic expression of a compound serotype (Paul et al., 1986; van Kooten Kok-Doorshadt et al., 1972; Okamoto et al., 1987a), or to more than one subtype existing on a single HBsAg particle (Paul et al., 1986). The mechanism that gives rise to a compound subtype particle is proposed to involve recombination between the genomes of different subtype HBV to produce a hybrid S gene which when expressed gives rise to a particle of compound subtype (Nordenfelt and Le Bouvier, 1973).

Determination of the d/y subtype appears to be due to the presence or absence of several specific epitopes, since there are panels of d-

specific and y-specific monoclonal antibodies which do not cross react (Swenson et al., 1988). Interestingly, these monoclonal antibodies are also specific for disulphide dependent or independent epitopes. The region of HBsAg comprising the subtype determining region has been mapped in a similar way to that of the a-region, using Hopp and Woods antigenicity predictions to synthesise peptides representing hydrophilic and variable regions of the protein, raising antisera, and then determining the subtype specificity of the antisera (Gerin et al., 1983). Such an approach has mapped the d/y region to between amino acids 110 - 139. Prince et al (1982) using the same approach found that a peptide representing adw from between 138 - 149 was sufficient to elicit a d-specific antiserum. By expressing various regions of the S gene in E. coli then subtyping antisera to antigen produced (MacKay et al., 1981b) mapped d-specific regions to between 4 and 121. It is apparent therefore that epitopes for the d/y system appear to co-occur, with a epitopes in the hydrophilic domain and, that given their disulphide bond dependence, it is likely that some of these epitopes are discontinuous, and distributed about the whole of the hydrophilic area.

The genomes of HBV of all 4 serotypes ayr (Okamoto et al., 1986), ayw (Galibert et al., 1979; Bichko et al., 1985), adr (Ono et al., 1983; Gan et al., 1984; Fujiyama et al., 1983; Kobayashi and Koike, 1984; K Yokoyama, personal communication), and adw (Valenzuela et al., 1979; Ono et al., 1983) have been sequenced along with that of a virus which was originally thought to have a compound subtype adyw, but has been designated ayw2 (Pasek et al., 1979; Pugh et al.,

1986; this thesis). If one compares the peptide sequence of HBsAg proteins from different subtypes (figure 1.4.2) one observes residue positions that correlate to some degree with the presence of a particular subtype. These correlations have been used as starting points for the investigation of the role of a particular residue in conferring an epitope that determines subtype. The difference in amino acid at position 122 determines the binding of a d-(lysine) or y-(arginine) specific monoclonal antibody to adr and ayr proteins (Okamoto et al., 1986). Also differences at position 160 can dictate the binding of a w and a r specific monoclonal antibody to adr and ayw proteins (Okamoto et al., 1987b). If the lysine of adw HBsAg at position 122 is subjected to reductive methylation, reactivity to a d-specific monoclonal antibody is decreased (Peterson et al., 1984). Replacement of the cysteine at position 121 (present in all serotypes of HBsAg) with a serine by site-directed mutagenesis, results in a decrease in reactivity to an a-specific monoclonal antibody as well as to a y-specific monoclonal antibody (Antoni and Peterson, 1988). These experiments infer that the area around position 122 is important in determining the nature of d/y epitopes, and of at least one discontinuous a-epitope. These experiments therefore represent an initial investigation into the structural basis of HBsAg serotypes.

1.5 Aims and Rationale of Project

The molecular basis of serotypic variation in HBV is at least partly due to differences in HBsAg structure. This proposal has been arrived at by indirect and direct means. It is possible to correlate

Figure 1.4.2 Comparison of peptide sequences of HBsAg of different subtypes

The genomes of HBV of all serotypes have been sequenced and the peptide sequence of HBsAg deduced. The consensus peptide sequence is represented by capital letters and amino acids at variable positions by small letters.

ayw2 (adyw) (Pasek et al., 1979; Pugh et al., 1986), ayw (Galibert et al., 1979; Bichko et al., 1985), ayr (Okamoto et al., 1986), adr (Ono et al., 1983; Gan et al., 1984), adr4 (Fujiyama et al., 1983), adw (Valenzuela et al., 1980), adw2 (Ono et al., 1983), adr2 (K Yokayama, personal communication)

```
          1                                                    50
      ayw2   ni                                          gttv l
      ayw    ni                                          gttv l
      ayr    st                                          gapt p
      adr    nt                                          gapt p
      adr4   nt                                          gapt p
      adw    ni                                          gspv l
      adw2   ni                                          gspv l
      adr2   st                                          eapt p
Consensus  ME--TSGFLG PLLVLQAGFF LLTRILTIPQ SLDSWWTSLN FLG----C-G

          51                                                   100
      ayw2       i           t
      ayw        t           t
      ayr        t           t
      adr        t           i
      adr4       t           i
      adw        t           i
      adw2       t           i
      adr2       t           i
Consensus  QNSQSP-SNH SPTSCPP-CP GYRWMCLRRF IIFLFILLLC LIFLLVLLDY

                        *          *            *
          101     110 113/4     122/25-27 131/33/34          150
      ayw2          i   ss        r  ttp    i my         s
      ayw           i   ss        r  mtt    t my         s
      ayr           l   ts        r  tip    t mf         s
      adr           l   ts        k  tip    t mf         s
      adr4          l   ts        k  tip    t mf         s
      adw           i   st        k  ttp    n kf         t
      adw2          i   st        k  ttp    n mf         t
      adr2          l   ts        k  tip    t mf         s
Consensus  QGMLPVCPL- PG--TTSTGP C-TC---AQG -S--PSCCCT KP-DGNCTCI

          151                                                  200
      ayw2         gk f   w  a                    v    i     v
      ayw          gk f   w  a                    v    t     v
      ayr          ar f   w  v                    v    t     a
      adr          ar f   w  v                    v    t     v
      adr4         ar f   g  v                    v    t     v
      adw          ak y   w  v                    v    t     a
      adw2         ak y   w  v                    v    t     a
      adr2         ar f   w  v                    a    t     v
Consensus  PIPSSWAF-- -LWE-AS-RF SWLSLLVPFV QWF-GLSP-V WLS-IWMMWY

          201                          226
      ayw2       s l    l             a
      ayw        s l    l             v
      ayr        n l    l             v
      adr        n l    l             v
      adr4       n l    l             v
      adw        s v    i             v
      adw2       s v    i             v
      adr2       n l    l             v
Consensus  WGPSLY-I-S PF-PLLPIFF CLW-YI
```

certain amino acids at specific positions in HBsAg with the subtype of the virus (figure 1.4.2, section 1.4.3). Direct evidence comes from serotypic mapping using synthetic peptides (Gerin et al., 1983), and mutational approaches which have implicated certain amino acids in subtype specific epitopes (Peterson et al., 1984; Okamoto et al., 1986; Okamota et al., 1987a, b; Antoni and Peterson, 1988).

One of the aims of this thesis was to assess the contribution of three amino acids occurring at various positions in determining d/y subtype specificity. The residues at position 122 correlate exactly with d/y variation, at 134 correlation is less than exact. Both sites have been implicated in d/y subtype determination by experimental means (Gerin et al., 1983). Position 113 exhibits amino acid variation in correlation with w/r subtype variation, but not with d/y so would represent a negative control for the effect of mutations on d/y subtype. By the use of oligonucleotide site-directed mutagenesis (Zoller and Smith, 1983) the effect of substitutions of amino acids at positions 113, 122, and 134 on the subtype of HBsAg (Pasek et al., 1979) (subtype ayw) was examined. Outlined in figure 1.5.1 are the nucleotide substitutions in the S gene necessary for the desired substitution of y to d specific amino acids. The effect of these mutations on the subtype of the yeast expressed protein (Murray et al., 1984) were assessed using an immunoprecipitation based subtyping assay, in this manner the contribution of the substituted amino acids in the determination of d/y subtype was assessed (Burrell et al., 1978).

Figure 1.5.1 Subtype variable amino acids substituted by Site-Directed Mutagenesis

The mutations made in adw2 HBsAg at sites of amino acid variation are indicated by boxes. The corresponding changes in nucleotide sequence, brought about by oligonucleotide-mediated, site-directed mutagenesis, are also indicated.

Amino Acid	Serotype					Mutation	Amino Acid Change
	ayw2	adw	adr	ayw	ayr		
110	ile	ile	leu	ile	leu		
	* TCA		* ACA			T → A	ser → thr
113	ser	ser	thr	ser	thr		
	* AGA	* AAA				G → A	arg → lys
122	arg	lys	lys	arg	arg		
125	thr	thr	thr	met	thr		
127	pro	pro	pro	thr	pro		
131	ile	asp	thr	thr	thr		
	* TAT	* TIT				A → T	tyr → phe
134	tyr	phe	phe	tyr	phe		

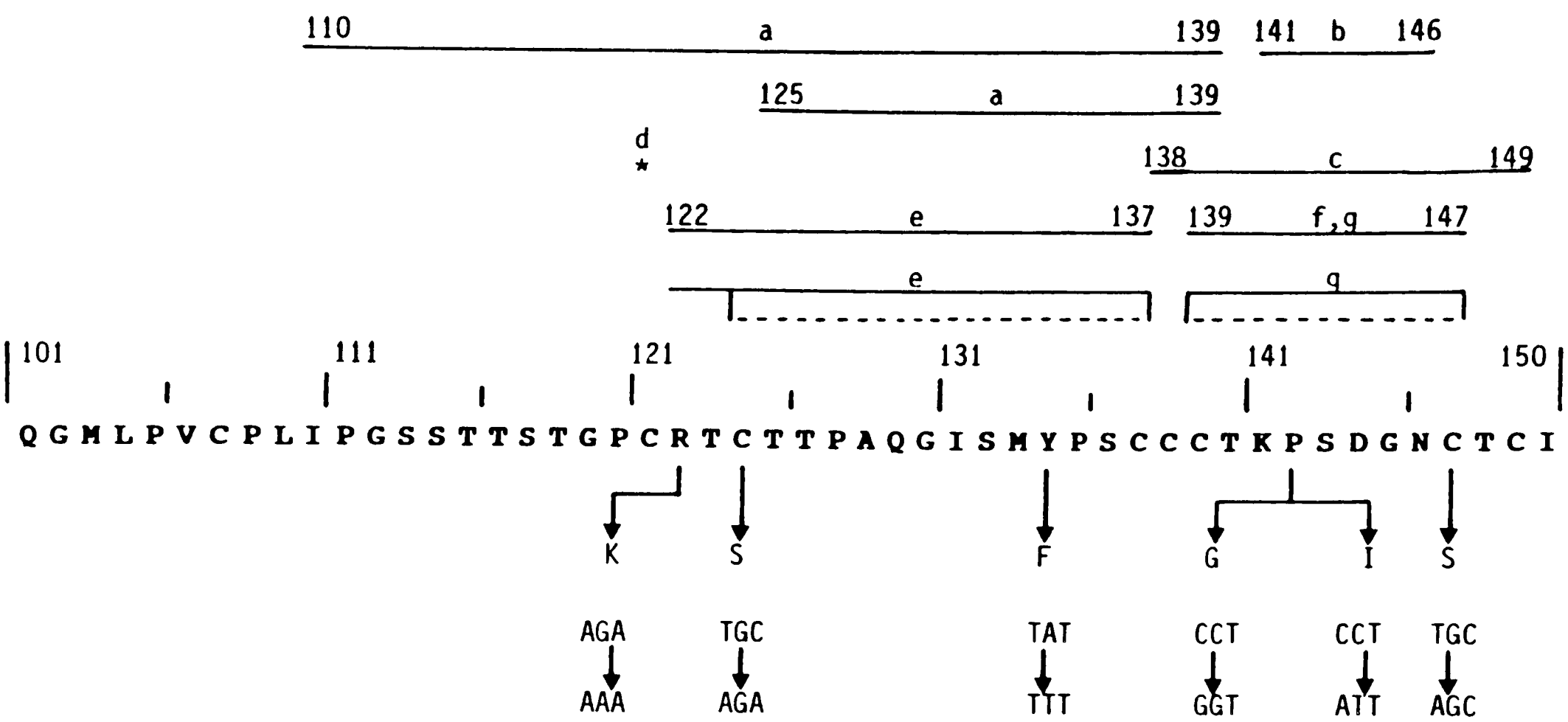
It has been emphasised previously that HBsAg owes much to conformational integrity for its antigenicity. However, in the absence of extensive structural studies (such as X-ray diffraction), the spatial arrangement of the a-antigenic region is largely unknown. Shown in figure 1.5.2 is a summary of the information derived from synthetic peptide based and mutational studies of the immunoreactive region of HBsAg. Using mutants in the a-region it was hoped that amino acids important in determining a-antigenicity could be identified. The work of Brown et al. (1984) and Dreesman et al. (1982) implicated the cysteine residues 139 and 147, 124 and 137 in contributing to antigenicity through the formation of disulphide bridges. To assess the importance of these cross linkages, substitution of the cysteines at 124 and 147 for serine in yeast expressed HBsAg was performed. The contribution of prolines to the spatial arrangement of polypeptides is well known - they introduce turns into the α-carbon backbone of proteins (Garnier et al., 1978). The substitution of proline 142 for glycine and isoleucine was made to determine the contribution of the residue, and by inference the turn at 142, to the antigenicity of the a-region.

The influence of the substitutions $122^{arg \rightarrow lys}$ and $134^{tyr \rightarrow phe}$ on a-antigenicity were also assessed. Both of these residue positions are located in the a-region, but are not important for a-antigenicity since the amino acids present vary between subtypes with no effect on HBsAg antigenicity (McAuliffe et al., 1980). This allowed one to assess the contribution of non-specific amino acid substitutions on a- antigenicity.

Figure 1.5.2 Antigenic region of HBsAg

The peptide sequence of HBsAg, deduced from the gene, cloned by Pasek et al. (1979) from position 101 to 150 is shown. Represented above it is a summary of the results from examinations of HBsAg B-cell antigenicity using synthetic peptides. The region of a-antigenicity has been mapped to between positions 141 and 146 (b) by Hopp + Woods et al. (1981), 138 and 149 (c) by Prince et al. (1982), 139 and 147 (f,g) by Bhatnager et al. (1982) and Brown et al. (1984), 122 and 137 (e) by Dreesman et al. (1982). The cyclisation of peptides e and g about cysteines 124, 137 and 139, 147 respectively, increases immunoreactivity. The region of HBsAg implicated in determining d/y subtype reactivity has been located by Gerin et al. (1983) (a) to between 110 and 139 (125 and 139). The substitution of cysteine with serine at position 121 (d) by Antoni and Peterson (1988) resulted in little change in HBsAg reactivity to an anti-HBsAg serum but altered antigen reactivity with several a-specific monoclonal antibodies.

The positions of the amino acid substitutions implicated are indicated below the peptide sequence, as are the corresponding nucleotide substitutions.


110
a
139
141
b
146
125
a
139
d
*
138
c
149
122
e
137
139
f,g
147
e
g
101
111
121
131
141
150
Q G M L P V C P L I P G S S T T S T G P C R T C T T P A Q G I S M Y P S C C C T K P S D G N C T C I
K
S
F
G
I
S
AGA
TGC
TAT
CCT
CCT
TGC
AAA
AGA
TTT
GGT
ATT
AGC

2. MATERIALS AND METHODS

2.1 Materials

2.1.1 Strains and vectors

E. coli K12 strains

	Relevant Genotype	Use	Reference
HB101	hsdS20 R^- M^-, recA 13, ara14, $proA_2$, lacY1, gal K_2, rpsL20, $(SM)^R$, xyl-5, mtl-1, supE44, λ^-	general host for plasmid cloning	Boyer et al. (1969)
BMH71-18	Δ(lac-pro), SupE, thi, /F', $proA^+B^+$, lacZ ΔM15, lacIq	plating cells for BMH71-18 mutL transfections	Gronenborn et al (1976)
NM522	hsd Δ5 derivative of BMH71-18	host for M13 mp8 phages	Gough and Murray (1983)
BMH71-18 mutL	as BMH71-18 except mutL::Tn10	transfection host for SDM produced M13 mp8 duplexes	Kramer et al. (1984a)

TG1	Δ(lac-pro), SupE, thi, hsdD5/F'tra D36, proA$^+$B$^+$, lacIq, lacZ ΔM15	transfection host for thio-deoxy C SDM reaction products	Amersham International (unpublished)

S. cerevisiae strains

DBY746	α, leu2-3 112, his3 Δ-1, trp1-289, ura3-52, gal-2	host for pHIN/G-2 mediated HBsAg production	Botstein D (unpublished)

Plasmids	Relevant Features	Use	Reference
pHIN/G-2	S. cerevisiae: LEU2 gene, 2μ replication origin, PH05 promoter and terminator E. coli: Tn903 KAN^R, pAT153 replication origin HBV: S gene	E. coli/S. cerevisiae shuttle vector. Allows expression of HBsAg in S. cerevisiae under control of yeast PH05 promoter	Murray et al. (1984)

Bacteriophages

M13 mp8	-	Cloning vector for SDM and sequencing experiments	Messing and Viera (1982)

2.1.2 Microbiological media

Bacteriological L-broth

10g Difco Bacto Tryptone

5g Difco Bacto Yeast extract

10g NaCl

Made to 1 litre, and adjusted to pH 7.2.

L-agar

As L-broth, but containing 15g agar per litre.

BBL top layer agar

10g BBL Trypticase

5g NaCl

10g Agar

Made to 1 litre.

Minimal medium

300ml Water agar

80ml 5 x Spizein salts

4 ml 20% (w/v) Glucose

100μl 2mg/ml Vitamin B1

2 x TY medium

16g Bacto Tryptone

10g Bacto Yeast extract

5 g NaCl

Made to 1 litre

Antibiotics

Ampicillin stock solution at 100mg/ml in H_2O.

Used at 100μg/ml.

Yeast Media

YPDA

10 g Difco Bacto Yeast extract

20g Difco Bacto-peptone

20g Glucose

10ml Adenine Sulphate 4 mg/ml

Made to 1 litre.

YPDA agar

YPDA containing 20g per litre agar.

Yeast minimal medium

6.7g Difco yeast nitrogen base without amino acids

20 g Glucose

20mg Amino acid/supplement as required

20g Agar

Made to 1 litre

Supplements for minimal medium: stock at 4mg/ml in H_2O. Final concentration of 20mg/l. supplements were: leucine, histadine, tryptophan and uracil.

Synthetic medium for expression

In order to control the concentration of culture medium phosphate, a synthetic medium was devised in accordance with Schurr and Yagil (1971)

2g Aspargine

20g Dextrose

30mg Potassium phosphate monobasic

0.5g Magnesium sulphate

0.1g Sodium chloride

0.1g Calcium chloride

Made to 1 litre

to every litre of synthetic minimal medium 2ml "Vitamin Cocktail" and 3 ml "Metal Cocktail" were added.

Vitamin cocktail

2mg	Biotin
0.4g	Calcium pantothenate
2mg	Folic acid
2g	Inositol
0.4g	Niacin
0.2g	p-Aminobenzoic acid
0.4g	Pyridoxine hydrochloride
0.2g	Riboflavin
0.4g	Thiamine hydrochloride

per litre H_2O

Metal cocktail

0.5g	Boric acid
40mg	Copper sulphate
0.1g	Potassium iodide
0.2g	Ferric chloride
0.4g	Manganese sulphate
0.2g	Sodium molybdate
0.4g	Zinc sulphate

per litre H_2O.

The medium was supplemented with the relevant amino acid or purine at a concentration of 20mg/litre.

"high" phosphate medium contained 1g/litre monobasic potassium phosphate.

"low" phosphate medium contained 30mg/litre monobasic potassium phosphate.

2.1.3 General solutions

SSC, 20X

3M NaCl

0.3M Sodium citrate.

Denhardt's solution, 10X

0.2% (w/v) BSA

0.2% (w/v) Polyvinyl-pyrrolidine (PVP)

0.2% (w/v) Ficoll 400 DL

TBE, 10X

0.1M Tris.HCl pH 8.3

89mM Boric acid

2mM EDTA

TE

10mM Tris.HCl pH 8.0

1mM EDTA

LTB

20mM Tris.HCl pH 7.9

20mM NaCl

1mM EDTA

Sequencing TE (SETE)

10mM Tris.HCl pH 8.0

0.1mM EDTA

PBS

4mM KH_2PO_4

6mM K_2HPO_4 pH 7.0

0.15M NaCl

RIP buffer

0.5% (w/v) BSA in PBS

(0.1% (w/v) Sodium azide as preservative)

2.1.4 Enzyme buffers

10X high salt restriction endonuclease buffer: (For use with BamHI)

0.5M Tris.HCl pH 8.0

0.1M $MgCl_2$

0.5M NaCl

10mM DTT

1mg/ml BSA

10X low salt restriction endonuclease buffer: (For use with ClaI)

as above with the omission of 0.5M NaCl

T4 DNA ligase buffer:

10X blunt ligation buffer

0.5M Tris.HCl pH 7.8

0.1M $MgCl_2$

0.2M DTT

10mM ATP

5X sticky ligation buffer

0.2M Tris.HCl pH 7.4

50mM $MgCl_2$

5mM DTT

2mM ATP

10X T4 DNA polynucleotide kinase buffer

0.25M Tris.HCl pH 7.5

0.1M $MgCl_2$

50mM β-mercaptoethanol

2.1.5 Solutions for dideoxy-nucleotide DNA sequencing

Termination mixes: made as following (volumes in μl)

	T°	C°	G°	A°
50mM dTTP		2.5	2.5	2.5
50mM dCTP	2.5		2.5	2.5
50mM dGTP	2.5	2.5		2.5
10mM ddTTP	15			
10mM ddCTP		7.5		
10mM ddGTP			15	
1mM ddATP				7.5
0.5mM dTTP	12.5			
0.5mM dCTP		12.5		
0.5mM dGTP			12.5	
TE	500	500	500	250
H_2O	500	500	500	750

Chase

dTTP dCTP dGTP dATP each at 0.25mM.

TM

0.1M Tris.HCl pH 8.5

50mM $MgCl_2$

Primer mix

1μl 17mer primer (0.5pm for universal primer, New England Biolabs).

1μl TM

9μl template DNA (~1μg)

Klenow mix (per clone)

4μCi α-[^{35}S]-dATP (Amersham, 600Ci/mM)

1.5U *E. coli* DNA polymerase I Klenow fragment (Boehringer)

0.8μl 10mM DTT

6.5μl 10mM Tris.HCl pH 8.5

Single Dideoxy-nucleotide Tracking

Primer mix

0.4μl 17mer primer (0.5pm for universal primer)

0.4μl TM

2.4μl H_2O

per clone.

3μl of the primer mix added to 3μl of template (0.3μg), the composition of termination chase, and Klenow mixes were the same as those used for complete sequencing.

Extended DNA sequencing

Extension mix 5X

7.5μM dCTP

7.5μM dGTP

7.5μM dTTP

Formamide dye

0.1% (w/v) Xylene cyanol

0.1% (w/v) Bromophenol blue

10mM EDTA

in 1ml deionized Formamide.

Polyacrylamide gel separation of dideoxy-nucleotide sequencing reaction products

Gels were made as follows:

17g	Urea
6ml	40% (w/v) Acrylamide (38% (w/v) Acrylamide/2% (w/v) Bisacrylamide)
4ml	10 X TBE
0.24ml	10% (w/v) Ammonium persulphate
40μl	N,N,N',N'-tetramethyethylenediamine (TEMED)

Final volume 40ml with H_2O.

Polyacrylamide gel for separation of DNA oligonucleotides

Gels were made up as follows:

21g	Urea
25ml	40% (w/v) Acrylamide(38% (w/v) Acrylamide/2% (w/v) Bisacrylamide
5ml	10 X TBE
0.35ml	10% (w/v) Ammonium persulphate
40µl	TEMED

Final volume 50ml with H_2O.

Solution for fixing DNA in urea-polyacrylamide gels

10% (v/v) Methanol

10% (v/v) Acetic acid

in H_2O.

2.1.6 Solutions for protein techniques

For a 13% (w/v) SDS-polyacrylamide gel, the solutions were:

Separating gel

23.2ml	30% (w/v) Acrylamide
23.2ml	0.8% (w/v) Bisacrylamide
6.25ml	3M Tris.HCl pH 8.8
0.4ml	10% (w/v) SDS
0.2ml	10% (w/v) Ammonium persulphate
20µl	TEMED

Stacking gel

1.9ml 30% Acrylamide

1.9ml 0.8% Bisacrylamide

1.9ml 1M Tris.HCl pH 6.8

0.15M 10% (w/v) SDS

9.35ml dH_2O

0.15ml 10% (w/v) Ammonium persulphate

10μl TEMED

Electrophoresis buffer 10X

0.2M Glycine

0.25M Tris.HCl pH 8.0

0.1% (w/v) SDS

Loading buffer

50mM Tris.HCl pH 6.8

1% (w/v) SDS

20% (v/v) glycerol

60mM DTT

0.1% (w/v) Bromophenol blue

Ponceau S stain

0.5% (v/v) Ponceau S in 5% (w/v) TCA solution. Destain in dH_2O.

Western Blotting

Transfer buffer

20mM Tris.HCl pH 8.1

0.14M Glycine

in 20% (v/v) Methanol.

10 X TS

0.1M Tris.HCl pH 8.8

1.5M NaCl

2.1.7 Immunological reagents

Sources HBsAg

Purified HBsAg produced from pHIN/G-2 expression in yeast was obtained from either the Green Cross Corporation (Osaka) or Biogen Inc.

The Green Cross material (>99% of total protein 698μg/ml) in 20mM sodium phosphate pH 7.2, 0.1% (w/v) sodium azide. The Biogen material was of similar purity (0.6mg/ml) and stored in PBS with 0.1% (w/v) sodium azide.

Human HBsAg was kindly provided by Dr J Peutherer, Department of Bacteriology, Edinburgh University. HBsAg was purified from carrier serum to 70 - 80% of total protein, using the method of Burrell

(1975). This purified material was labelled with ^{125}I by the iodogen based process (Burrell, 1973).

Preparation of Antisera

Anti-HBsAg mouse monoclonal antibody (clone 1E6) was kindly provided by the Green Cross Corporation. HBsAg of subtype adr had been used to immunise mice and antibodies from clone 1E6 purified to a level of >99% total protein (500μg/ml) and stored in 10mM acetic acid, pH 6.7, 60mM NaCl, 0.002% (w/v) sodium azide, 1% (w/v) mannitol.

The antiserum NRA/B7-4B was kindly provided by Dr P Wingfield (Biogen Inc) and prepared in the following manner: HBsAg (100μg) produced and partially purified from yeast as in Murray et al. (1984) was denatured by incubation in 1% (w/v) SDS, 60mM DTT at 80°C for five minutes. The antigen preparation was then dialysed against PBS to remove SDS and DTT and used to immunise New Zealand white rabbits subcutaneously in Freund's complete adjuvant, the animals were boosted with the same preparation after five weeks, then again after two weeks.

Anti-human HBsAg sera were obtained from Dr J Peutherer. The monospecific anti-y and anti-d antisera were prepared as follows: purified human HBsAg (100 - 150μg) was used to immunise New Zealand white rabbits subcutaneously in Freund's complete adjuvant, then after 6 weeks the animals were re-immunised and, if necessary again two weeks later. Antibodies which were a and subtype specific were

removed by absorbing against antigen of the opposite d or y serotype. This absorbtion was carried out using 100μg HBsAg for 16 hours at 4°C followed by centrifugation at 30,000 x g for 10 minutes at 4°C to remove antibody-antigen complexes. This process was then repeated a further two times.

Donkey anti-rabbit IgG and sheep anti-mouse IgG were obtained from the Scottish Antibody Production Unit.

2.2 Methods

2.2.1 DNA methods

Phenol extraction of nucleic acids

Distilled phenol was equilibrated with 1M Tris.HCl pH 8.0 prior to use. Equal volume of phenol was added to aqueous nucleic acid solution then mixed vigorously by vortexing. Following centrifugation at 17,000 x g for 10 minutes, at room temperature, the upper aqueous layer was removed.

Ethanol precipitation of DNA

Ethanol (98% (v/v)) stored at -20°C was added to aqueous solutions of DNA in a 2.5:1 volumetric ratio; 3M sodium acetate pH 5.2 was also added at a volume 0.1 of that of the DNA solution. The mixture was vortex mixed and incubated at -20°C for 10 minutes. Precipitated

DNA was recovered by centrifugation at 4°C at 17,000 x g for 15 minutes. The pellet was then washed with 70% (v/v) ethanol at -20°C and dried under vacuum.

Restriction endonuclease digestion

Enzymes were purchased from Pharmacia, Boehringer-Mannheim, and New England Biolabs. DNA was digested at 37°C for up to 90 minutes in 1X restriction endonuclease buffer and enzyme added to give a ratio of 10U:1ug DNA. Reactions were terminated by phenol extraction and DNA recovered by ethanol precipitation.

DNA ligation

T4 DNA ligase was purchased from Boehringer. Ligations between DNA bearing cohesive ends were carried out for 16 hours at 15°C by T4 DNA ligase (0.1U/μg DNA) in 1X sticky ligation buffer at a concentration of 5 - 30μg/ml.

Blunt end ligations were carried out at 15°C for 16 hours by T4 DNA ligase (1U/μg) in 1X blunt ligation buffer at a DNA concentration of 200 - 500μg/ml.

Transformation of E. coli

M13 RF DNA: L-broth (100ml) was inoculated with 1ml of a stationery phase culture of NM522, BMH71-18 mutL or TG1 cells and grown to

OD_{600nm} 0.2 by shaking at 37°C. Cells were then pelleted by centrifugation at 4,000 x g for 5 minutes at 4°C, and resuspended in 10ml ice-cold 0.1M $CaCl_2$ and left on ice for 30 minutes. After repelleting and resuspension in 2ml ice-cold 0.1M $CaCl_2$, 0.2ml of cell suspension was added to M13 RF or single-stranded DNA in glass tubes and left on ice for 30 minutes. The cells were heat-shocked at 42°C for 4 minutes and plated in 3.5ml BBL top with 40μl of stationery phase cells on minimal plates. Isopropyl-β-D-thiogalactopyranaside (IPTG) (20μl) and 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (30μl) were added to BBL top to detect β-galactosidase positive M13 clones. IPTG and X-gal were stored at a concentration of 20mg/ml in dimethylformamide at -20°C.

Plasmid DNA: The preparation of competent cells was performed as above except that the cells resuspended in 2ml $CaCl_2$ were left for a further 60 minutes on ice instead of being used immediately. The heat-shocked cells were diluted into 1ml L-broth and incubated at 37°C for 60 minutes. The mixture (0.2ml) was then spread on selecting L-amp plates.

Transformation of S. cerevisiae using sphaeroplast method

The method used was that developed by Beggs (1982). A culture (50ml) of DBY746 was grown in YPDA at 30°C to OD_{600nm} 0.5 then centrifuged at 4,000 x g for 5 minutes at room temperature, and cells resuspended in ice-cold, sterile dH_2O (20ml). After re-pelleting, cells were resuspended in 1M sorbitol, 25mM EDTA pH 8.0,

50mM DTT (20ml) and incubated at 30°C with gentle shaking for 20 minutes. Cells were pelleted, washed with 1.2M sorbitol (ice-cold) and resuspended in 1.2M sorbitol, 10mM EDTA, 0.1M sodium citrate pH 5.8 (20ml) containing 1% (v/v) β-glucuronidase (Sigma H-I) then incubated for 90 minutes at 30°C. Sphaeroplasts were pelleted and washed 3 times with ice-cold 1.2M sorbitol then resuspended in 1.2M sorbitol, 10mM $CaCl_2$ (0.5ml). DNA was added to 50μl of sphaeroplast suspension to give a final concentration of 20μg/ml, and incubated at room temperature for 15 minutes. Then 10mM Tris.HCl pH 7.5, 10mM $CaCl_2$, 10% (w/v) polyethylene glycol 4000 (PEG 4000) solution (0.5ml) was added and sphaeroplasts immediately pelleted by centrifugation at 17,000 x g for 2 minutes at room temperature. Sphaeroplasts were gently resuspended in YPDA (100μl) containing 1.2M sorbitol and incubated at 30°C for 15 - 60 minutes. Dilutions of sphaeroplast suspension were made in 1.2M sorbitol and plated on selecting minimal plates in 3% (w/v) minimal agar. Colonies of transformed cells appeared after incubation for 4 days at 30°C.

Labelling DNA by Random Priming

The method used was that of Feinberg and Vogelstein (1983; 1984). The DNA fragment of interest was purified by electrophoresis through low-gelling temperature agarose (Seakem) and the desired band cut out (~1μg). The agarose was added to dH_2O in a ratio of 3ml dH_2O:1g agarose, then placed in a 100°C water bath for 7 minutes before being incubated at 37°C for 10 to 60 minutes. The labelling reaction was carried out using 35.5μl of agarose slurry added to 10μl oligo-

labelling buffer (OLB), 2μl BSA (10mg/ml), 50μCi α-[^{32}P]-dCTP (Amersham, 3000Ci/mM) and 1μl (5U) E. coli DNA polymerase I Klenow fragment (Boehringer). The reaction was left for between 6 and 20 hours at room temperature.

OLB was made from: 50μl solution A (1.25M Tris.HCl pH 8.0, 0.125M $MgCl_2$, 25mM β-mercaptoethanol, 0.5mM each of dTTP, dGTP, dATP), 125μl solution B (2M HEPES (4-(2-hydroxyethyl)-1-piperazine-ethanesulphoric acid) adjusted to pH 6.6 with NaOH) and 75μl solution C (random hexanucleotides OD_{260nm} 90, in TE).

Removal of Unincorporated Labelled Nucleotides using Sephadex G-50

A 1ml column of sephadex G-50 was equilibrated in STE (10mM Tris.HCl pH 8.0, 0.15M NaCl, 1mM EDTA) by centrifugation at 2,000 x g for 3 minutes at room temperature. The probe was diluted in 100μl STE and added to the column. Unincorporated α-[^{32}P]-dCTP was trapped in the column and labelled DNA collected after centrifugation at 2,000 x g for 5 minutes.

Colony Hybridisation

The method used was essentially that of Grunstein and Hogness (1975). Colonies were blotted on to nitrocellulose filters, which were then placed on blotting paper saturated with 0.5M NaOH, 1.5M NaCl, for 2 minutes. Filters were neutralised by treatment with 0.5M Tris.HCl pH 7.4, 3M NaCl twice, for 5 minutes. After drying at 37°C, the filters

were baked under vacuum at 80°C for 90 minutes, and prehybridised for 60 minutes at 37°C in 1X Denhardt's solution, 5 x SSC, 0.3% (w/v) SDS, 0.2mg/ml sonicated salmon sperm DNA, 50% (v/v) formamide. Hybridisation was performed by incubation under the same conditions for between 16 and 20 hours with a sephadex G-50 purified probe (10^8cpm/μg DNA). Filters were washed for 10 minutes in 2 x SSC, 0.1% (w/v) SDS at room temperature with shaking; then for 10 minutes in 0.5X SSC, 0.1% (w/v) SDS. After drying the nitrocellulose filters were autoradiographed by exposure to X-ray film (pre-flashed), and incubation at -70°C between two intensifying screens.

Agarose Gel Electrophoresis

Agarose gels were made and electrophoresed in 1 x TBE gel buffer. The concentration of agarose varied according to the size of DNA fragments to be separated, but was typically 0.6 - 1.0% (w/v). Generally, agarose gels were cast containing 0.5μg/ml ethidium bromide. Gels were electrophoresed at 7.1V/cm for between 1 and 3 hours or at 1.4V/cm for to 20 hours. DNA was visualised under short-wave ultraviolet radiation.

Isolation of DNA from Agarose Gels

The DNA band of interest was cut out of low-gelling point agarose (Seakem) and added to an equal volume of 1 x TBE, 0.2M NaCl solution. The agarose was melted by incubation at 65°C for 15 minutes; the

agarose solution was then phenol extracted twice and the DNA recovered by ethanol precipitation.

Preparation of Plasmid DNA

Large scale: L-broth (10ml) supplemented by the appropriate selecting antibiotic was inoculated with a single colony of the plasmid-carrying strain. The culture was grown to stationery phase then diluted into 500ml of similarly supplemented L-broth. The cells were grown to OD_{600nm} 1.0 at which point chloramphenicol (75mg) was added and the culture grown for a further 16 to 20 hours. Cells were harvested by centrifugation at 4,000 x g, 4°C for 15 minutes then resuspended in 6ml of sucrose mix (50mM Tris.HCl pH 8.1, 40mM EDTA, 25% (w/v) sucrose). Lysozyme (1ml of 10mg/ml (Sigma)) and 0.5M EDTA (0.5ml) were added, the cell suspension was then mixed gently and left on ice for 5 minutes. At this point 13ml of Triton mix (50mM Tris.HCl pH 8.1, 63mM EDTA, 0.1% (w/v) Triton X-100) was added, the cells mixed by swirling and left on ice for 10 minutes. The lysate was cleared by centrifugation at 27,000 x g, 4°C for 30 minutes. Caesium chloride was added to the cleared lysate to a concentration of 0.95g/ml and ethidium bromide to that of 1mg/ml. Supercoiled plasmid DNA was separated from chromosomal and nicked plasmid DNA by isopycnic ultracentrifugation at 38,000rpm, 20°C in a Sorvall TV865 rotor for 16 hours. DNA bands were visualised under short-wave UV illumination and the lower supercoiled plasmid removed. Ethidium bromide was removed by 4 rounds of extraction with dH_2O saturated

butan-1-ol, and the CsCl by dialysing against TE. DNA was recovered by ethanol precipitation.

Small Scale (Mini-Plasmid Preparation)

The method used was essentially that of Birnboim and Doly (1979). Cells from 1.5ml of a stationery phase culture, grown in the presence of a selecting antibiotic, were pelleted by centrifugation at 17,000 x g, room temperature for 5 minutes and resuspended in 0.1ml 34mM Tris.HCl pH 8.0, 10mM EDTA, 1% (w/v) glucose solution, then incubated at room temperature for 5 minutes. After this, 0.2ml of 0.2M NaOH, 1% (w/v) SDS (prepared fresh) was added, mixed gently by inversion and placed on ice for 5 minutes. Upon the addition of 0.15ml ice-cold 5M potassium acetate pH 4.8 and incubation on ice for 5 minutes, protein, SDS, and chromosomal DNA were precipitated. The precipitate was clarified by centrifugation at 17,000 x g, room temperature for 10 minutes, and plasmid DNA recovered from the salt solution by ethanol precipitation. When DNA derived from mini-plasmid preparations was digested with restriction endonucleases, RNase H (IU) (Boehringer) was included in the reaction to remove RNA prior to agarose gel electrophoresis and visualization of DNA.

Preparation of M13 RF DNA

A single fresh plaque of M13 was picked into 2ml of NM522 or TG1 cells at OD_{600nm} 0.2 and shaken for 6 hours at 37°C; then 0.5ml of these infected cells was added to 500ml of L-broth along with 5ml of

cells at OD_{600nm} 0.2 and shaken at 37°C for 7 hours. Cells were harvested by centrifugation at 14,000 x g at 4°C. From this point onwards the procedure was that of a standard large scale plasmid preparation.

2.2.2 Dideoxy-nucleotide DNA sequencing methods

Preparation of single-stranded M13 template

Small scale

A fresh single plaque of M13 was picked into 1.5ml of 2 x TY medium along with 15μl of a stationery phase culture of TG1 or NM522 cells. The culture was shaken vigorously for 5.5 hours at 37°C, then the cells were pelleted by centrifugation at 17,000 x g, room temperature, for 5 minutes. Phage were precipitated from the supernatant by the addition of 0.2ml 2.5M NaCl, 20% (w/v) PEG 6000 solution. The mixture was incubated at room temperature for 30 minutes then centrifuged at 17,000 x g, room temperature, for 10 minutes, and supernatant removed by vacuum aspiration. The PEG pellet was then re-centrifuged and residual supernatant removed. Precipitated phage was resuspended in 0.1ml SETE then mixed by vortexing for 10 seconds with 50μl phenol. The mixture was incubated at room temperature for 10 minutes then centrifuged at 17,000 x g, room temperature for 2 minutes. Phage DNA was then recovered from the aqueous layer by ethanol precipitation.

Large scale

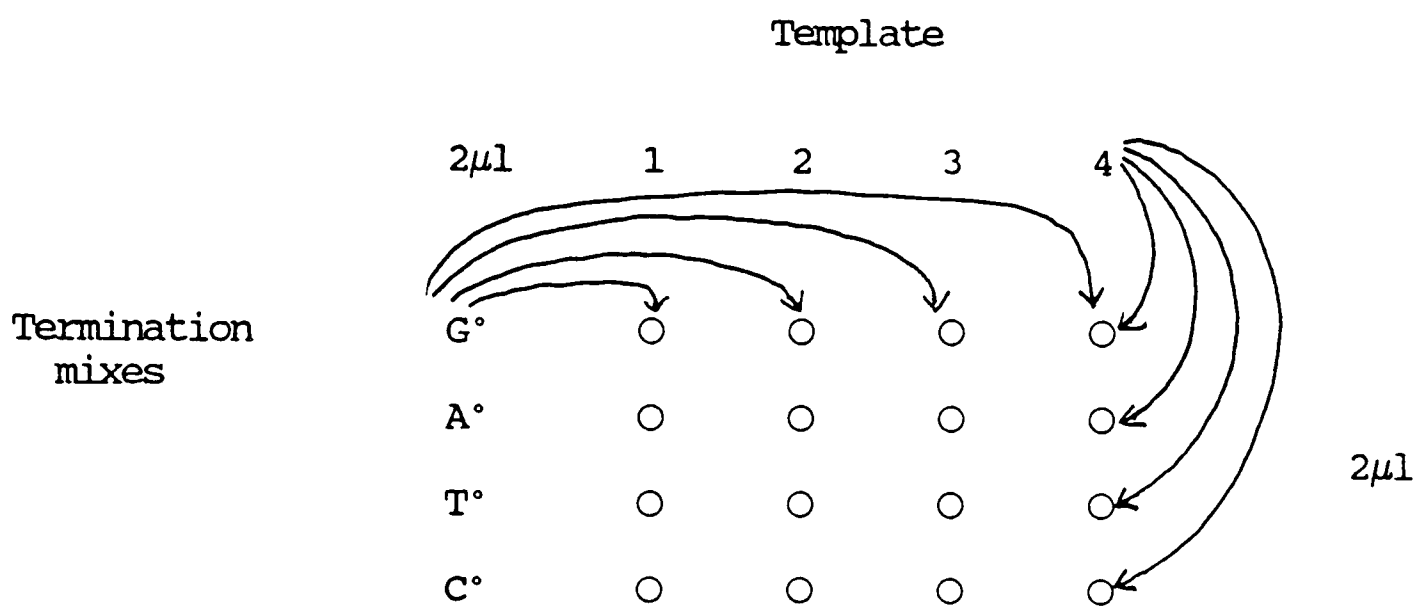
TG1 cells (50μl) from a stationery phase culture grown in 2 x TY were added to 20ml of 2 x TY and grown for 3 hours at 37°C. This culture (100μl) was then used to inoculate 1ml 2 x TY medium along with a fresh M13 plaque. Incubation at 37°C for 4 hours with shaking followed, after which cells were pelleted by centrifugation at 17,000 x g, room temperature for 5 minutes and removed. Phage supernatant (1ml) was then used to inoculate a 100ml 2 x TY culture of TG1 cells (OD_{550nm} 0.3). The culture was grown for 4 hours with shaking, at 37°C then centrifuged at 5,000 x g for 30 minutes at 4°C. The supernatant was carefully removed then mixed thoroughly with 0.2 volume of 2.5M NaCl, 20% (w/v) PEG 6000 and incubated at 4°C for 1 hour. Phage precipitate was removed by centrifugation at 5,000 x g for 20 minutes at 4°C, the supernatant removed and any residue eliminated by re-centrifugation of the PEG pellet. The viral pellet was resuspended in 0.5ml TE and centrifuged at 17,000 x g for 5 minutes at room temperature to remove any remaining cells. PEG/NaCl solution (0.2ml) was then added to the supernatant, mixed well, then incubated at room temperature for 15 minutes or overnight at 4°C. The viral pellet was then recovered by centrifugation at 17,000 x g as before and any residual supernatant was removed by centrifugation. Phenol (0.2ml) was added to the viral pellet suspension (0.5ml) and mixed by vortexing for 10 seconds then incubated at room temperature for 15 minutes. Following centrifugation for 3 minutes at 17,000 x g, room temperature, the aqueous layer was removed and subjected to phenol extraction, performed in the same way as before. Single-

stranded DNA was recovered by ethanol precipitation and resuspended in TE to a concentration of 1μg/ml.

Dideoxy-nucleotide DNA sequencing reactions

The method used was essentially that of Sanger et al. (1977). Template DNA was primed by incubating with a 17mer oligonucleotide of complementary sequence to that of the cloned insert fragment or lacZ gene of M13 mp8, at 80°C for 5 minutes as in section 2.1.5. The primer mix was then cooled over a period of 20 minutes to room temperature.

Siliconised microtitre plates were used to perform the sequencing reactions, which were arranged in the following manner:



Termination mixes (2μl) were placed in the siliconised wells along with primer/template mix (2μl). After the two solutions were mixed by tapping the plate, Klenow mix (2μl) was added to every well, mixed

and incubated at room temperature for 25 minutes. Chase mix (2μl) was added and mixed, and the solution incubated for a further 25 minutes. For single track-sequencing only one termination mix was added to primer/template mix, and the rest of the reactions performed as normal. For extended DNA sequencing, the template-primer reaction was performed as above then 2μl 1 x Extension mix, was added as well as Klenow mix. The extension reaction was incubated for 5 - 15 minutes at room temperature at which point the termination mixed (2μl) were added and incubated as normal. The chase reaction was not performed.

Urea-polyacrylamide electrophoresis of sequencing reaction products

Reactions were stopped by the addition of 2μl of formamide dye then incubated at 80 or 100°C for 5 minutes. Sequencing gel mix, prepared as in section 2.1.5, was poured into a glass gel former sealed by adhesive tape (40 x 20 x 0.035cm). A suitable comb was then inserted; either a Flat former comb enabling the subsequent use of a "shark's tooth" comb (BRL) or a square-toothed sequencing gel comb. The polymerised gel was assembled into an electrophoresis kit and the upper and lower tanks filled with 1 x TBE. The reaction mixes were loaded using drawn out glass capillaries in a G-A-T-C order.

The samples were electrophoresed at 50V/cm for at least 2 hours, and gels fixed in 10% (v/v) acetic acid, 10% (v/v) methanol for 10 minutes, then dried on 3MM paper using a vacuum gel drier (Zabona).

The gel was then autoradiographed for between 16 and 20 hours at room temperature using X-ray film.

2.2.3 Site-directed mutagenesis methods

Phosphorylation of 5' ends of DNA with T4 DNA polynucleotide kinase

Polynucleotide kinase (PK) was used to add 5' phosphate groups to unphosphorylated 17mer oligonucleotides which were then used in SDM reactions. Approximately 100pm of oligonucleotide was incubated in 1 x PK buffer (section 2.1.4) with 1U of T4 DNA polynucleotide kinase (Boehringer) at 37°C for 30 minutes. The substrate providing phosphate was ATP at 20mM. The reaction was stopped by incubation at 70°C for 10 minutes and rapid cooling on ice.

Labelling 5' ends of oligonucleotides with γ-[^{32}P]-ATP using T4 DNA poly-nucleotide kinase

Adenosine 5'-γ-[^{32}P] Triphosphate (Amersham, 3000 Ci/mMol, 30μCi) was used to label 15pm of oligonucleotide in the presence of 1 x PK buffer with 1U of PK (Boehringer) at 37°C for 30 minutes.

Screening M13 plaques by hybridisation with mutagenic oligonucleotide

The method used was essentially that of Zoller and Smith (1983). Plaques arranged in asymmetric grids of a hundred were blotted on to nitrocellulose filters. The filters were then transferred to a sheet

of blotting paper saturated with 0.5M NaOH and left for 3 minutes at room temperature. They were neutralised by transferring to 1M Tris.HCl pH 7.4 soaked paper (twice for 1 minute each), then to 0.5M Tris.HCl pH 7.5, 1.5M NaCl soaked paper (5 minutes). The blotted plaques were dried at room temperature for 20 minutes then baked for 60 minutes at 80°C in a vacuum oven. Filters were pre-wetted in 6 x SSC and prehybridised at 67°C for 5 minutes in 10 x Denhardt's, 6 x SSC, 0.2% (w/v) SDS solution. After rinsing in 6 x SSC each filter was dropped in probe solution (^{32}P labelled oligonucleotide 7.0 x 10^9cpm/μg, diluted in 4ml 6 x SSC), and left at room temperature for 60 minutes.

Filters were initially washed at room temperature in 6 x SSC (3 x 30ml, one minute washes per filter), partially dried then autoradiographed using pre-flashed X-ray film at -70°C for 1 to 3 hours. The filters were then re-washed in 6 x SSC at a temperature 5°C below that of the calculated Tm (Wallace et al., 1979) for the mutant oligonucleotide-template duplex, then autoradiographed. The washing and autoradiographing process was repeated in 6 x SSC at a temperature corresponding to the Tm.

Gapped-Duplex Mediated Site-Directed Mutagenesis

The method used was that of Kramer et al. (1984a). The Gapped-Duplex was formed by adding 0.1pm (0.5μg) M13 mp8 RF DNA cut with BamHI to 0.5pm (1.25μg) template DNA in 30μl 12.5mM Tris.HCl pH 7.5, 0.2M KCl, then heating to 65°C for 1 minute and cooling to room temperature

over 20 minutes. The template-RF mixture (8μl) was then added to 7.5pm (35ng) of 5' phosphorylated mutant oligonucleotide in a final volume of 10μl, heated to 65°C for 3 minutes then cooled at room temperature for 20 minutes. The primed-Gapped-Duplex was "filled-in" by incubation at room temperature for 45 minutes with T4 DNA ligase (2U, Boehringer), E. coli DNA polymerase I Klenow fragment (1U, Boehringer) in the presence of 20mM Tris.HCl pH 7.5, 15mM $MgCl_2$, 12.5mM KCl, 2mM DTT, 124μM ATP, 125μM dNTPs. Competent BMH71-18 mutL cells were transformed with the reaction mix and plated on minimal media with BMH71-18 cells.

Thio-deoxy C mediated site-directed mutagenesis

The method used was based on the procedure developed by Taylor et al. (1985 a, b). Mutant oligonucleotide (5' phosphorylated; 8pm, 40ng) was annealed to single-stranded template DNA (3.8pm, 10μg) by incubation in 0.15M Tris.HCl pH8.0, 0.15M NaCl (34μl) at 70°C for 3 minutes followed by slow cooling at 37°C for 30 minutes then cooling on ice. Extension and ligation of mutant strand DNA was performed by the incubation of template-primer duplex at 16°C for 15 hours with E. coli DNA polymerase I Klenow fragment (12U, Amersham), T4 DNA ligase (12U, Amersham), 10mM $MgCl_2$, 1mM ATP, 0.5mM dATP, 0.5mM dGTP, 0.5mM dTTP and 0.5mM dCTP-α-S (Sp isomer). The reaction mixture was diluted to 0.27ml with NaCl to a concentration of 0.6M and passed through two 13mm Sartorius nitrocellulose filters (SM11336) in a Swinnex Millipore filtration device (SX001300). The dsDNA contained in the filtrate was recovered by ethanol precipitation and

resuspended in 10mM Tris.HCl pH 8.0, 50mM NaCl, 20μM EDTA. A fifth of the DNA recovered (about 3μg) was digested with NciI restriction endonuclease (5U, Amersham) in 15mM Tris.HCl pH 8.0, 20mM NaCl, 10mM $MgCl_2$, 30μM EDTA, 15mM DTT. To this digest, E. coli Exonuclease III (50U, Amersham) was added and incubated at 37°C for 30 minutes in 55mM Tris.HCl pH 8.0, 67mM NaCl, 10mM $MgCl_2$, 0.1mM EDTA, 15mM DTT. Both enzymes were then inactivated by incubation at 70°C for 15 minutes, and the partial duplex repolymerised and ligated using E. coli DNA polymerase I (3U, Amersham) and T4 DNA ligase (2U, Amersham) in 50mM Tris.HCl pH 8.0, 60mM NaCl, 10mM $MgCl_2$, 60μM ATP and 40μM dNTPs at 16°C for 3 hours. Competent TG1 cells were transformed with the reaction mix and plated on minimal medium with TG1 cells.

2.2.4 Protein and immunological methods

Protein extraction from S. cerevisiae

The method used was developed from that of Murray et al. (1984). Cultures of DBY746 LEU^+ cells (1l.) were grown to OD_{600nm} 2.0 in low phosphate synthetic medium, then cells harvested by centrifugation at 5,000 x g for 15 minutes at 4°C. Cells were washed with PBS (100ml) then resuspended in PBS containing 1% (w/v) Triton X-100, 0.1mM phenylmethyl sulphonyl fluoride (PMSF) (10ml). Acid-washed (Ballitoni No. 1) glass beads (14g) were added to the suspension in glass tubes, then agitated by vortex mixing 10 times for 30 seconds, with one minute incubations on ice between each mixing. The lysate

was clarified by centrifugation at 12,000 x g for 10 minutes at 4°C and the supernatant stored at -70°C.

Batch Fermentation

Cultures grown in high phosphate synthetic medium (1.85l.) were diluted into 50l. of low phosphate medium and incubated for approximately 20 hours at 30°C in a Fermatic 50 LB (Biotec) batch fermenter until OD_{600nm} 2.0 was achieved. A cell pellet was recovered (~100g) by continuous-flow centrifugation, washed with PBS, then resuspended in 1/250 volume with PBS containing 1% (w/v) Triton X-100 and 0.1mM PMSF. The cells were broken by 3 x 15 second passages through a French press (700kg/cm^3) and extract clarified by centrifugation at 17,000 x g for 10 minutes at 4°C. Ammonium sulphate precipitation were performed by adding a pre-weighed amount over a period of 30 minutes to crude extract at 4°C with constant mixing. The precipitate was incubated for a further 30 minutes then centrifuged at 17,000 x g for 10 minutes at 4°C and the pellet resuspended in PBS. Dialysis for a period of between 16 to 20 hours against PBS removed any residual ammonium sulphate associated with the precipitated protein.

SDS-polyacrylamide gel electrophoresis (Laemmli, 1970)

Separating gel solution (13% w/v) was prepared as in section 2.1.6, and the gel cast between 25 x 20 x 0.1cm glass plates, then overlaid with dH_2O, after polymerisation stacking gel was poured and the comb

inserted. The gel was assembled into a vertical gel kit and samples boiled for 5 minutes in 1X loading buffer before loading. Electrophoresis in 1X buffer at 4.5V/cm was carried out for 6 hours at room temperature.

Immunoblotting

The method used was largely based on that of Towbin et al. (1979). The nitrocellulose and SDS-polyacrylamide gel were sandwiched together between 6 pieces of blotting paper (of similar size) which were pre-wetted in transfer buffer. The sandwich was placed in the "Trans-blot" apparatus (Biorad), filled with transfer buffer, and protein transferred at 2.7V/cm for 16 hours or 8V/cm for 4 hours at 4°C. The filter was stained in 0.5% (w/v) Ponceau S solution for 30 seconds and destained in dH_2O, to check the efficiency of transfer of protein, then incubated at room temperature in 5% (w/v) milk powder, 0.5% (v/v) NP40, 0.1% (w/v) sodium azide, 1 x TS solution for a period of 4 to 10 hours. After blocking, antibody was added in the same solution and the filter incubated, with shaking, for 16 to 20 hours at room temperature. The filter was then washed 6 times over a period of 30 minutes in 1 x TS then incubated with a 1:7500 dilution of goat anti-rabbit IgG (F_c) alkaline phosphatase conjugate (Promega Biotec) in milk powder buffer, for 3 hours at room temperature. The filter was then washed with 1 x TS as before, and the presence of bound alkaline phosphate detected by incubation in 0.1M Tris.HCl pH 9.5, 0.1M NaCl, 50mM $MgCl_2$, 0.33mg/ml Nitro-blue-tetrazolium (NBT),

0.25mg/ml 5-bromo-4-chloro-3-indolyl phosphate (BCIP) for 2 to 15 minutes at room temperature.

Lowry assay (Lowry et al., 1951)

Protein sample (0.4ml) was added to 2% (w/v) Na_2CO_3, 0.01% (w/v) $CuSO_4$, 0.02% (w/v) sodium citrate, 0.1M NaOH solution (2ml) mixed well and incubated at room temperature for 10 minutes. Folin-Cocateu reagent (50% (v/v) in dH_2O, 0.2ml) was then added, mixed well and left to stand at room temperature for 15 minutes. Absorbances at OD_{550nm} were then recorded; the assay was calibrated using BSA at concentrations of 1 to 100μg/ml.

Radioimmunoassay for HBsAg

The procedure used was based on the Abbott Ausria II kit for the detection of HBsAg using antisera raised against human HBsAg. Samples diluted in RIP buffer (0.2ml) were placed in the wells of microtitre plates along with polystyrene beads coated with guinea pig anti-human HBsAg and incubated at room temperature for between 16 and 20 hours. Samples were then removed by vacuum pump aspiration and the beads washed in the wells 10 times with 1ml of dH_2O. After this [^{125}I]-anti-human HBsAg (0.2ml; 0.74μCi/ml) was added and the beads incubated at 45°C for 1 hour. Label was then removed by aspiration and the beads washed as before. The degree of gamma-radioactivity of each bead was determined using a LKB 1275 mini-gamma counter.

Double antibody radioimmuno-precipitation assay (DARIP)

The assay was first used for the detection of HBsAg or anti-HBsAg antibodies by Burrell et al. (1978). It was used also as the basis of a competition assay determinig the subtypic activity of HBsAg and anti-HBsAg sera. The basic assay was as follows:

Antisera was diluted in RIP buffer, then added to 50μl [^{125}I]-HBsAg (3000cpm in 0.5ng HBsAg). A further 0.1ml of RIP buffer was added, mixed well, then the mixture incubated for between 20 and 48 hours at 4°C. The second antibody, diluted 1:1 in RIP buffer, was then added (0.1ml), and mixed well. The specificity of the second antiserum was directed against the first, so therefore enabled the precipitation of complexes between antibodies from the first serum and antigen. The mixture was incubated for 16 to 20 hours at 4°C then diluted in 0.1% (w/v) starch in RIP buffer (1ml) and centrifuged at 2,500 x g for 10 minutes at room temperature. The supernatant was decanted and the amount of gamma-radioactivity in the pellet determined. The relative avidity of the HBsAg antiserum for HBsAg was expressed as the percentage of total [^{125}I]-HBsAg precipitated.

The relative avidity of HBsAg for anti-d and anti-y monospecific antisera was tested using an inhibition DARIP assay. Antigen preparation (0.1ml, diluted in RIP buffer) was incubated at 37°C for 2 hours with monospecific antiserum (50μl). Then labelled HBsAg was added (50μl), mixed well, and the mixture incubated for between 20 and 48 hours at 37°C. The second, precipitating, antibody was then

added (0.15ml) the solution mixed well and incubated at 4°C for 16 to 20 hours. Labelled antigen-antibody complexes were recovered as described before and the percentage of total [^{125}I]-HBsAg precipitated determined. In this procedure the avidity of the unlabelled test antigen for monospecific anti-HBsAg serum was indicated by the degree of inhibition of reaction between the labelled HBsAg (of appropriate subtype) and antiserum.

Preparation of anti-HBsAg sera

The inoculation of mice with yeast produced HBsAg was kindly performed by G Leadbetter, Department of Bacteriology, Edinburgh University Medical School. Mice (Balb/c) were injected intraperitoneally (0.8ml) and intramuscularly (0.1ml, with 0.1ml Freund's complete adjuvant) with ammonium sulphate fractions (20-35% saturation) of yeast extracts containing about 3μg HBsAg. After 6 weeks the animals were boosted with a similar amount of antigen applied in the same manner. The third inoculation was performed as before 14 days after the second, then after 10 days the animals were bled (0.1ml removed) and the serum tested for the presence of anti-HBsAg antibodies using the DARIP assay.

RESULTS

3. SITE-DIRECTED MUTAGENESIS OF THE S GENE

3. Site-Directed Mutagenesis of the S Gene

In order to investigate the role of specific amino acids in the determination of HBsAg immunochemistry, the site-directed mutagenesis of single-stranded S gene DNA was performed.

3.1 Method of Site-Directed Mutagenesis

The basis of the two methods of SDM used in the project was the hybridisation of a single-stranded DNA oligonucleotide (17 bases long) to a single-stranded cloned copy of the S gene. The oligonucleotide was complementary to the wild type gene sequence except for one or two bases, under the hybridisation conditions used this mismatch did not disrupt the oligonucleotide:template duplex. The oligonucleotide then served as a primer for extension of the template sequence by the copying of the opposite DNA strand by *E. coli* DNA polymerase I Klenow fragment and T4 DNA ligase resulting in a heteroduplex of mutant and wild type DNA. Depending on the method used, this basic process can be optimised to varying degrees so as to increase the production of mutant single-stranded M13 DNA.

The two methods of SDM used in the project were the Gapped-Duplex method (Kramer *et al.*, 1984a; figure 3.1.1A) and the Thio-deoxy C method (Taylor *et al.*, 1985a, b; figure 3.1.1B). The former method optimised the formation of mutant-template duplexes by allowing Gapped-Duplexes to form between single-stranded template and denatured RF M13 molecules. The latter method involved DNA synthesis

template is removed by filtration and the remaining double-stranded DNA subjected to NciI digestion. The enzyme will not cut DNA at its recognition site of CC-GG if dCTP-α-S has been incorporated in the DNA. The result of this is that NciI nicks only the template strand of the heteroduplex, which is removed by subsequent E. coli Exonuclease III digestion. A homoduplex of mutant DNA is produced upon repolymerisation with E. coli DNA polymerase I, T4 DNA ligase and dNTPs, then used to transform TG1 cells.

Figure 3.1.1

A. Gapped-Duplex mutagenesis

As described in section 2.2.3 (Kramer et al., 1984a), linearised RF M13 DNA is heated with an excess of single stranded M13 DNA (containing the target sequence) and mutant oligonucleotide. The mixture is cooled slowly and the Gapped-Duplex formed with hybridised oligonucleotide in place. Extension takes place from the 3'-OH of the RF strand and the 3'-OH of the mutant oligonucleotide by the action of E. coli DNA polymerase I Klenow fragment and T4 DNA ligase until the synthesised mutant strands are ligated to the 5' end of the RF-derived strand and the 5' end of the mutant oligonucleotide. These duplexes are then used to transform mismatch repair deficient BMH71-18 mutL cells.

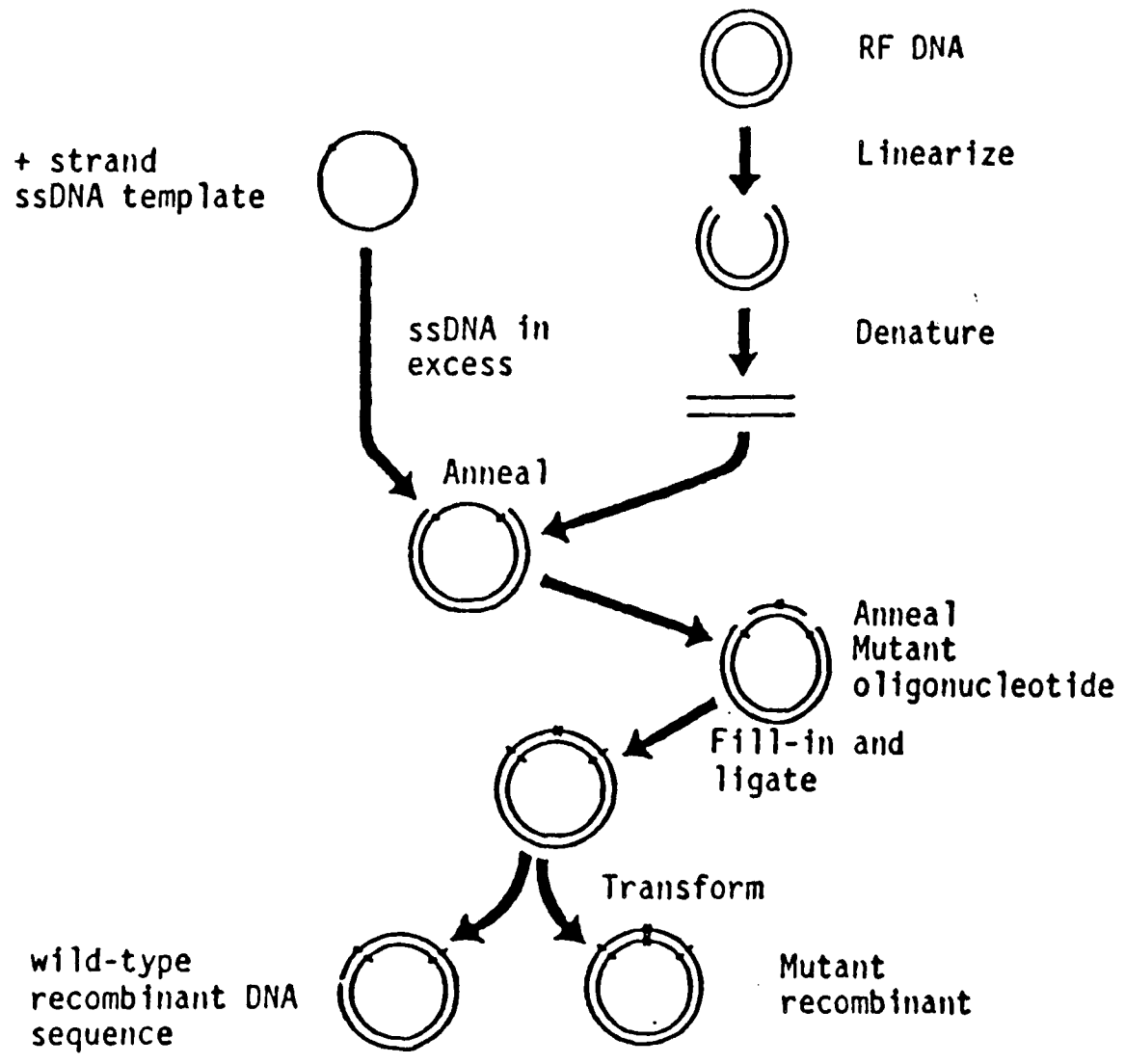
B. Thio-deoxy C method of mutagenesis

As described in section 2.2.3 (Taylor et al., 1985 a and b) oligonucleotide (in excess) is annealed to M13 single-stranded template containing the sequence to be mutated, by heating and slow cooling. The mutant strand of the duplex is extended from the primer with E. coli DNA polymerase I Klenow fragment and T4 DNA ligase where phosphorothioate-modified dCTP is incorporated as well as the standard dATP, dTTP and dGTP. Extension is continued around the template and the synthesised strand ligated to the 5'-end of the mutant oligonucleotide. Unwanted, unprimed

A

Kramer

(Gapped-Duplex)

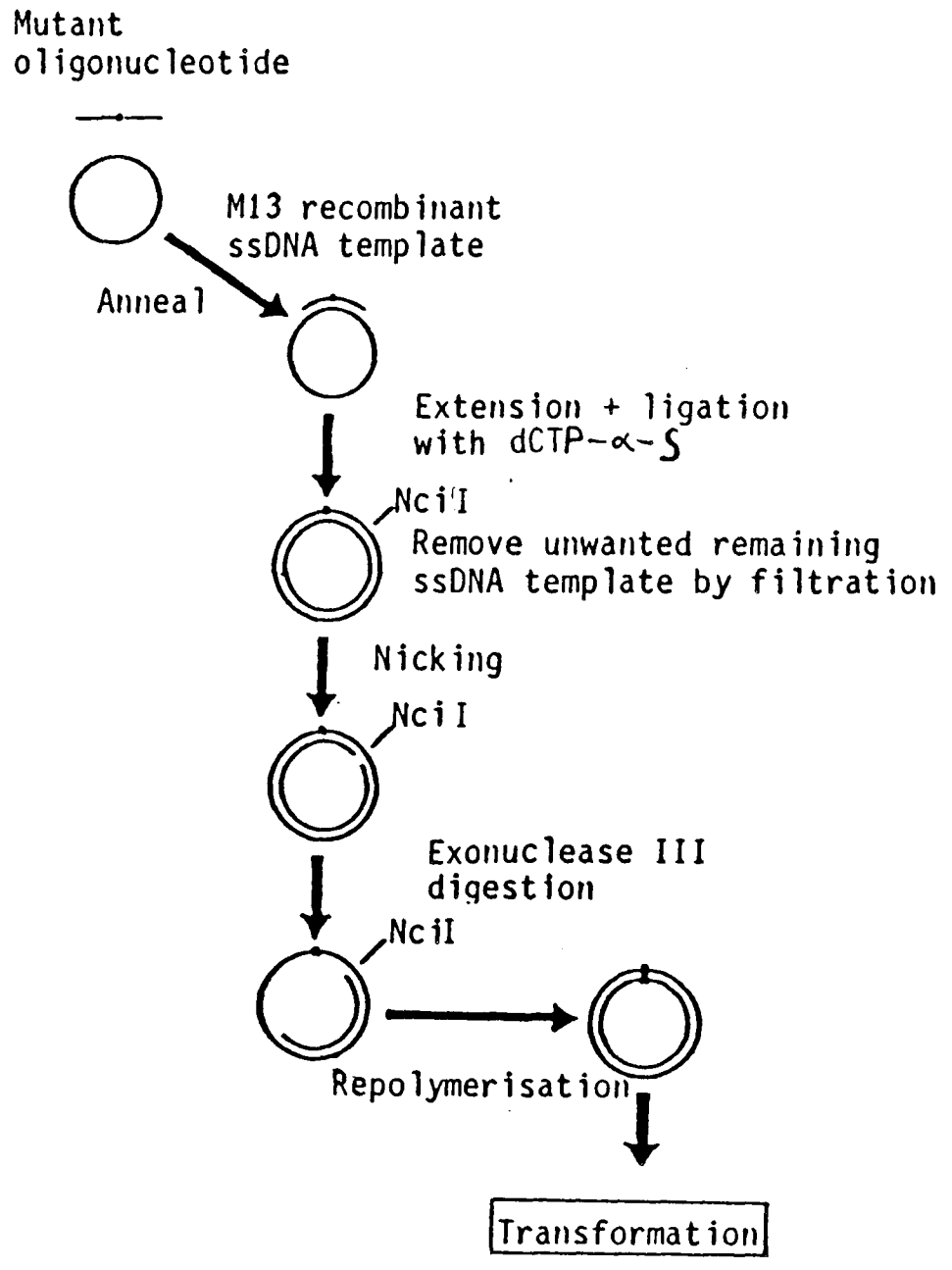


Efficiency ~6%

B

Eckstein

(thio-deoxy C)



Efficiency ~80%

from the hybridised mutant oligonucleotide around a naked template and the removal of contaminating un-primed template by filtration, a series of biochemical steps were then performed to remove the template strand of the heteroduplex and leave a homoduplex of mutant DNA.

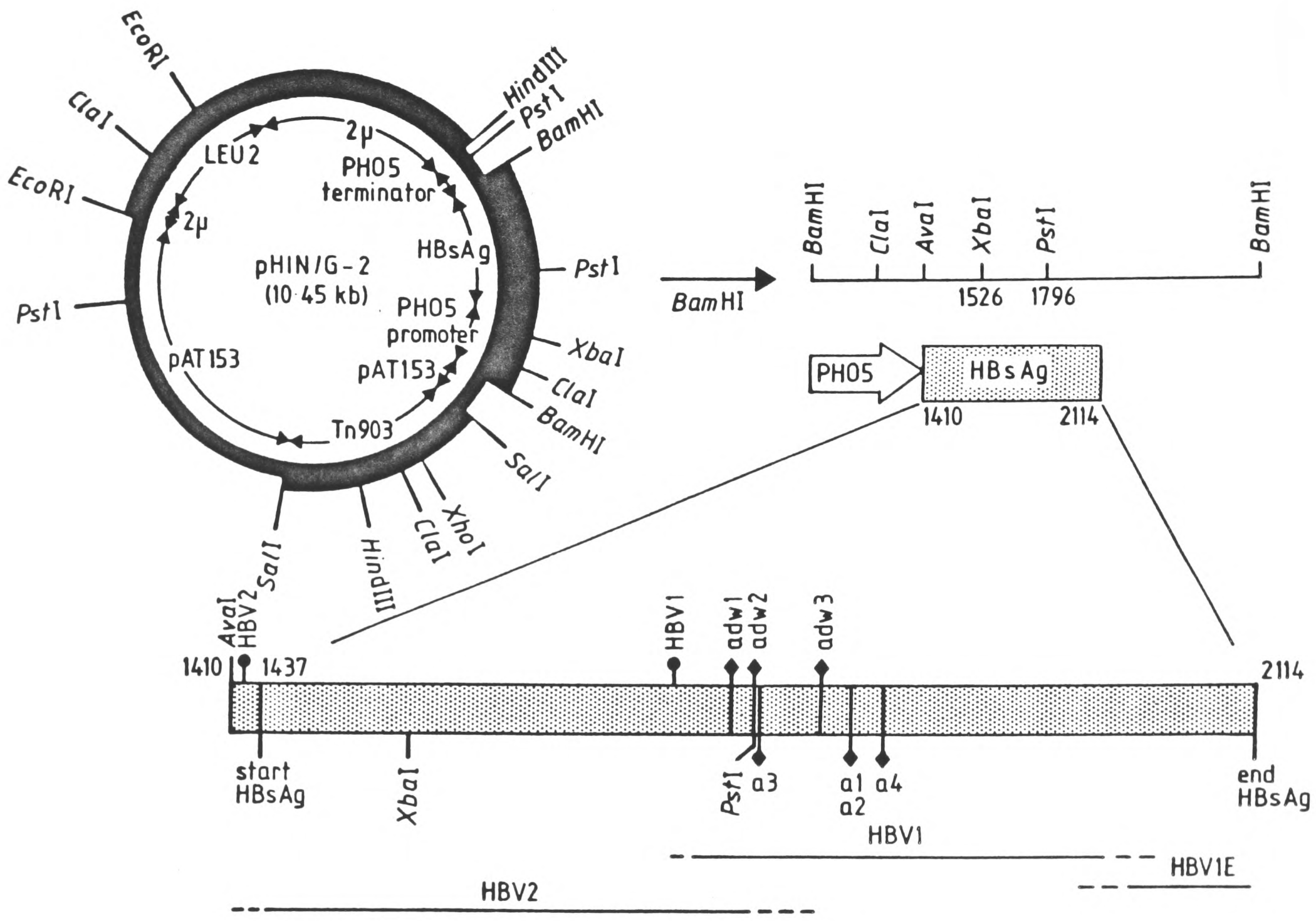
3.2 Sub-Cloning of the S Gene contained in the BamHI Fragment of pHIN/ G-2

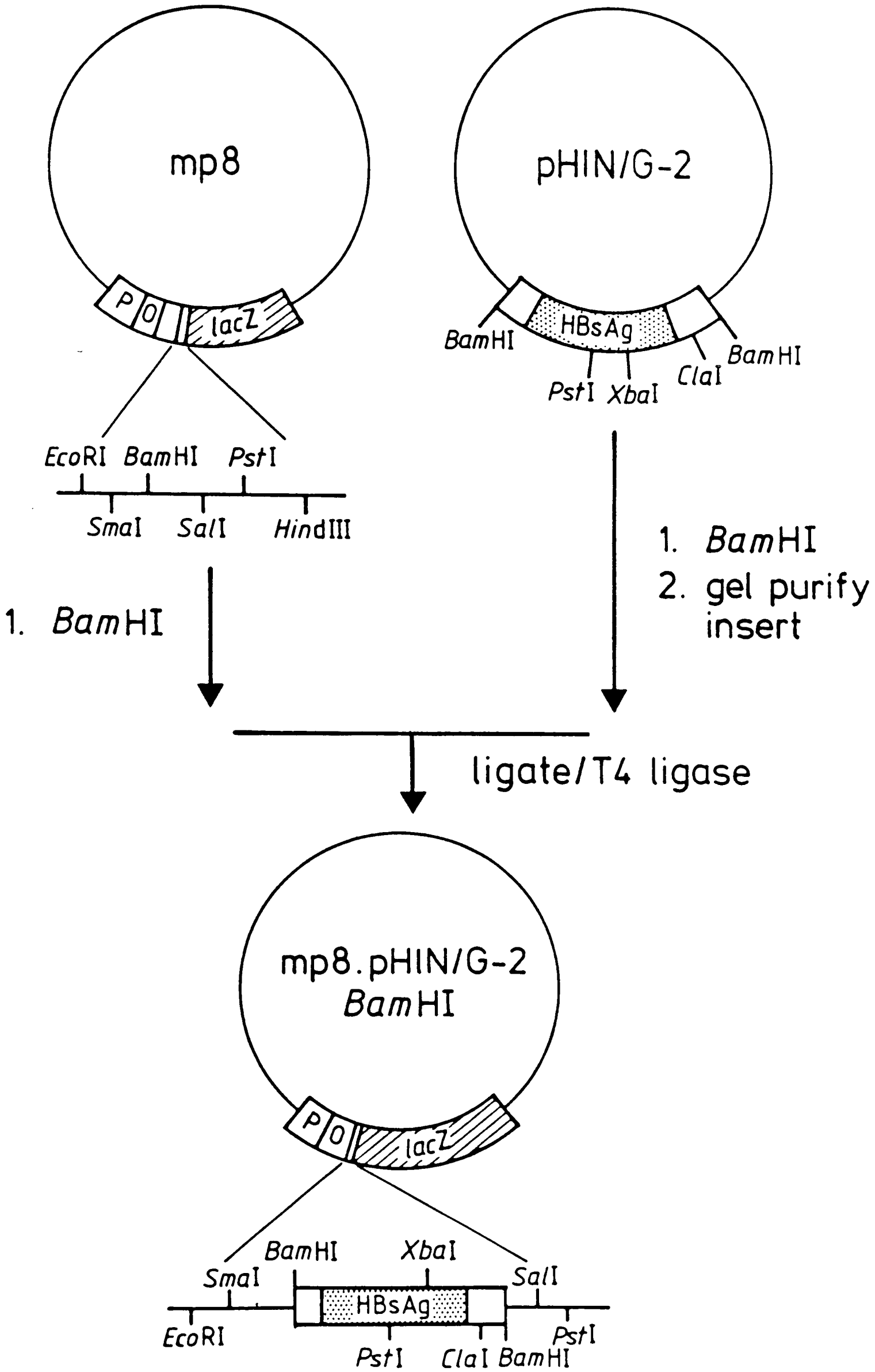
The plasmid pHIN/G-2 (5μg) was digested with BamHI (10U) and the 1.78kb fragment which contained the HBV S gene and S. cerevisiae PHO5 gene promoter (figure 3.2.1) separated from the 8.67kb fragment by agarose gel electrophoresis. The 1.78kb fragment was recovered from the low melting point agarose and resuspended in TE to a final concentration of approximately 0.1μg. The RF form of M13 mp8 (1μg) was similarly digested with BamHI, then phenol extracted and ethanol precipitated and resuspended in TE to a final concentration of 0.1μg. The BamHI digested M13 mp8 DNA (0.1μg) was ligated with the pHIN/G-2 1.78kb BamHI fragment (0.1μg) using T4 DNA ligase (1U) (figure 3.2.2). The ligation mix was then used to transform NM522 cells along with the following control samples; digested unligated M13 mp8 DNA (0.1μg), self ligated digested M13 mp8 DNA (0.1μg) and supercoiled M13 mp8 DNA (1ng). Typically, using the $CaCl_2$ method of transformation about 10^3 plaques per ng of supercoiled M13 DNA were observed. Plaques derived from recombinant clones of M13 mp8 DNA were distinguished from other transformants by Xgal/IPTG colour selection with the white plaques containing Lac^- M13 mp8 recombinant

Figure 3.2.1 Structure of pHIN/G-2

The S. cerevisiae/E. coli shuttle vector pHIN/G-2 (Murray et al., 1984) contains the 2μ origin of replication from S. cerevisiae and the pAT153 derived Col E1 origin of replication. Selection in S. cerevisiae is afforded either by the LEU2 gene or by the Tn903 encoded G418^R gene. The pAT153 region also contains the β-lactamase gene which confers the AMP^R phenotype on transformed E. coli cells. Expression of HBsAg is achieved by the in-frame fusion of the yeast acid phosphatase gene(PH05) promotor to an AvaI site 27bp upstream of the start of the S gene. Transcription from the promotor is repressed by phosphate (Schurr and Yagil, 1971) and S gene transcription terminated by the PH05 terminator. The PH05 promotor, the whole of the S gene, and approximately 0.5kb of HBV DNA is contained in the 1.78kb BamHI fragment.

Also indicated are the sites mutated in the S gene (denoted by stalked diamonds) and the positions of sequencing primers (denoted by stalked circles). The approximate areas sequenced from each primer (HBV1 and HBV2) are indicated by the horizontal lines below the gene.


EcoRI
ClaI
EcoRI
PstI
HindIII
PstI
BamHI
2μ
LEU2
2μ
PHO5
terminator
HBsAg
pHIN/G-2
(10·45 kb)
PHO5
promoter
pAT153
pAT153
Tn903
PstI
XbaI
ClaI
BamHI
SalI
XhoI
ClaI
HindIII
AvaI
HBV2
SalI
BamHI
BamHI
ClaI
AvaI
XbaI
PstI
BamHI
1526
1796
PHO5
HBsAg
1410
2114
1410
1437
HBV1
adw1
adw2
adw3
2114
start
HBsAg
XbaI
PstI
a3
a1
a2
a4
end
HBsAg
HBV1
HBV1E
HBV2


mp8
P
O
lacZ
EcoRI
BamHI
PstI
SmaI
SalI
HindIII
1. BamHI
pHIN/G-2
HBsAg
BamHI
PstI
XbaI
ClaI
BamHI
1. BamHI
2. gel purify insert
ligate/T4 ligase
mp8.pHIN/G-2
BamHI
P
O
lacZ
BamHI
XbaI
SmaI
SalI
HBsAg
EcoRI
PstI
ClaI
BamHI
PstI

clones and the blue Lac$^+$ non-recombinant clones. Recombinant clones were then used to make single-stranded template with NM522, which was then sequenced to determine which clones contained the BamHI fragment in the orientation indicated in figure 3.2.2.

3.3 Gapped-Duplex Mutagenesis

The Gapped-Duplex method of SDM (figure 3.1.1A) was used to make the d/y set of mutants of the pHIN/G-2 S gene. The oligonucleotides used were all 17mers and kindly synthesised by Biogen Inc. The designation of name, oligonucleotide and nucleotide position associated with each mutation is shown in figure 3.3.1.

The adw1 (nucleotide position 1775), adw2 (nucleotide position 1799), and adw3 (nucleotide position 1835) mutants were the first mutants to be made using the SDM protocol of section 2.2.3, using oligonucleotides phosphorylated by T4 DNA polynucleotide kinase. The SDM reactions were performed on template DNA (0.1μg) prepared from 1.5ml cultures of M13 mp8.pHIN/G-2 BamHI phase infected NM522 cells. Several control SDM annealings/reactions were carried out alongside the reaction proper. They were: I template + oligonucleotide, II template + RF, III template, IV RF. These control mixes and the reaction proper mix were used to transform competent BMH 71-18 mutL cells. The "mutL" cells were used for this experiment since they apparently increased the efficiency of SDM by being deficient in the ability to correct aberrant DNA duplexes such as the ones produced by the Gapped-Duplex method (Kramer et al., 1984b). Typical numbers of

Figure 3.3.1 Oligonucleotides used for site-directed mutagenesis

All oligomers (17mers) were synthesised by a solid-phase phosphodiester procedure by Biogen Inc (adw 17mers) or Department of Chemistry, University of Edinburgh (a-region 17mers and internal sequencing primers). The designation of name to oligonucleotide was as follows: adw1 - $1775^{T \rightarrow A}$ (ser → thr), adw2 - $1799^{G \rightarrow A}$ (arg → lys), adw3 - $1835^{A \rightarrow T}$ (tyr → phe), a1 - $1860/1861^{CC \rightarrow GG}$ (pro → gly), a2 - $1860/1861^{CC \rightarrow AT}$ (pro → ile), a3 - $1806^{T \rightarrow A}$ (cys → ser), a4 - $1878^{T \rightarrow A}$ (cys → ser).

T

adw1 5' T C C A G G A|A C A|T C A A C C A 3'

Thr 113

G

adw2 5' A C C C T G C|A A A|A C C T G C A 3'

Lys 122

A

adw3 5' C T C T A T G|T T T|C C C T C C T 3'

Phe 134

C C

a1 5' T A C A A A A|G G T|T C G G A T G 3'

Gly 142

C C

a2 5' T A C A A A A|A T T|T C G G A T G 3'

Ile 142

T

a3 5' C A G A A C C|A G C|A C G A C T C 3'

Ser 124

T

a4 5' T G G A A A C|A G C|A C C T G T A 3'

Ser 147

transformants from successful SDM reactions of this type (i.e. 5% of transformants were mutants) were as follows: I template + oligonucleotide-656, II template + RF-176, III template-2, IV RF-42, V (Reaction Proper) template - oligonucleotide + RF-816 (figures from adw2 SDM experiment). The controls allowed the monitoring of components of the SDM system; if the background number of transformants from template (III) or RF (IV) rose significantly, then screening was aborted and the reactions re-performed with new DNA. The level of transformants from II (RF + template) indicated the efficiency of the "filling in" reaction, since duplexes made were primed from the 3'-OH of the cut RF strand and continued to the 5'-P of the RF strand of the Gapped-Duplex. The efficiency of oligonucleotide hybridisation and priming was indicated by the level of template (I) transformants. The plating cells for the transformation of BMH71-18 mutL cells by SDM reaction mixes were BMH71-18, which were used for all further experiments with the mutant phage.

Initial screening for adw1, adw2 and adw3 mutants was performed by differential hybridisation with the original mutagenic oligonucleotide (Zoller and Smith, 1983; section 2.2.3). The Tm value for each duplex between oligomer and target sequence was determined using the Wallace rule (Wallace et al., 1979). The Tm values were determined as: adw1 52°C, adw2 50°C, adw3 50C°. By washing at 22°C, Tm - 5°C and Tm, in 6 x SSC, the necessary discrimination between unmutated and mutated sequence was achieved (figure 3.3.2). The observed frequency of occurrence of mutant

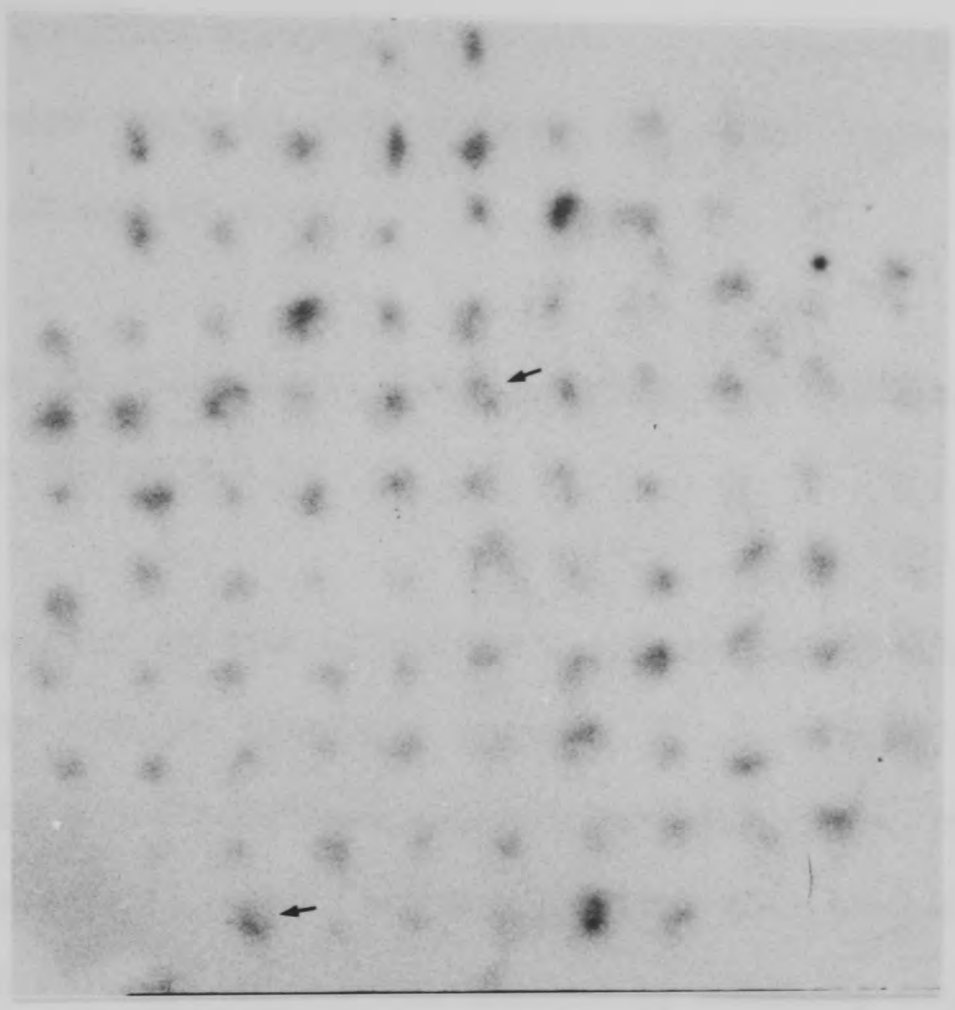
Figure 3.3.2 Use of differential hybridisation to identify a mutant sequence

Oligonucleotide (adw2) was labelled with γ-[^{32}P]-ATP using T4 DNA polynucleotide kinase and hybridised to DNA from M13 phage immobilised on nitrocellulose according to the procedure described in section 2.2.3. The filter was washed at 45°C (Tm - 5°C) in 6 x SSC (A). Upon increasing the temperature of the wash to 50°C (Tm) the necessary discrimination between unmutated and adw2 mutated sequence was achieved (B). The plaques indicated by arrows contained adw2 mutated sequence, as judged by the relative stability of oligomer-DNA duplexes at the Tm.

The Wallace rule states (Wallace et al., 1979): The melting temperature (Tm), defined as the temperature when half the duplex molecules have disassociated into their constituent strands, can be determined by the formula:

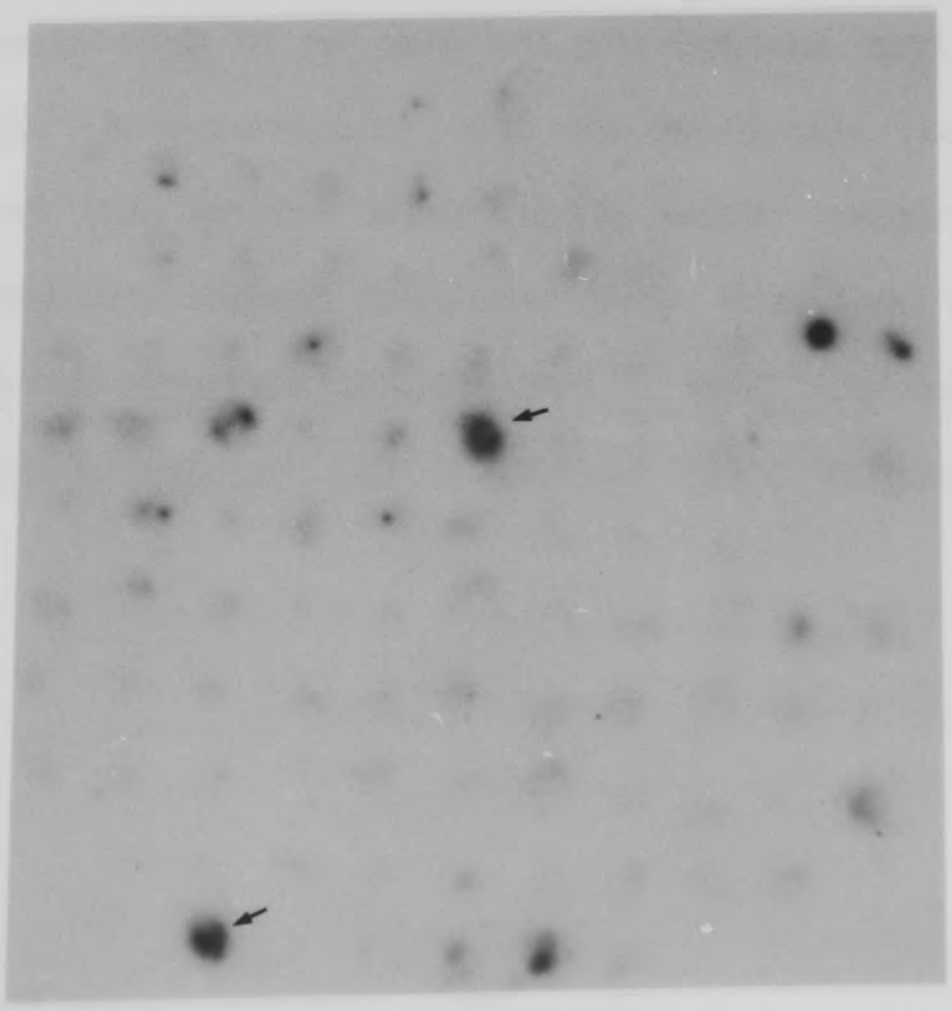
$$Tm\ (°C) = 4\ (G + C) + 2\ (A + T)$$

where G, C, A and T indicate the number of the corresponding nucleotides in the oligomer. The formula applies for DNA oligomers (16 - 20 nucleotides long) hybridised in 6 x SSC.



A

Tm -5°C



B

Tm

clones was 2 - 6% for all transformants. Candidate mutant phage were then plaque purified from the LTB stock (one plaque diluted in 50μl), single-stranded template DNA prepared and then sequenced; mutant clones were stored at -20°C.

The sequences of the mutated sites in adw1, adw2 and adw3 mutants are presented in figure 3.3.3.

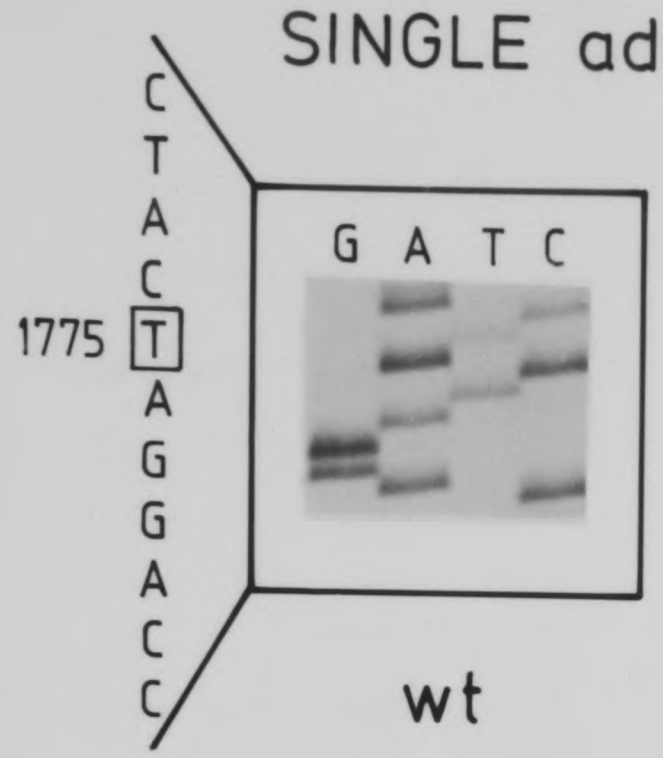
The investigation of the effect of combinations of adw mutants was initiated by taking template mutated at a single position and mutating with a different adw oligonucleotide. Thus, by mutating an adw2 template with adw1 oligonucleotide the adw2.1 mutant was produced. Similarly, adw3.1 was created by mutating an adw3 template with adw1 oligonucleotide and by using the adw3 oligonucleotide on the adw2.1 template, adw2.1.3 was produced. The screening procedure used in these mutagenesis experiments involved the single-dideoxy nucleotide sequencing of 60 template preparations from original SDM transformants. A sequence for just one nucleotide (usually A) was produced, and the desired substitution easily recognised (figure 3.3.4). Candidate mutant templates were then used to transform BMH71-18 and the resulting plaques used to make template DNA. Fully sequenced mutant template was stored at -20°C (figures 3.3.5 and 3.3.6).

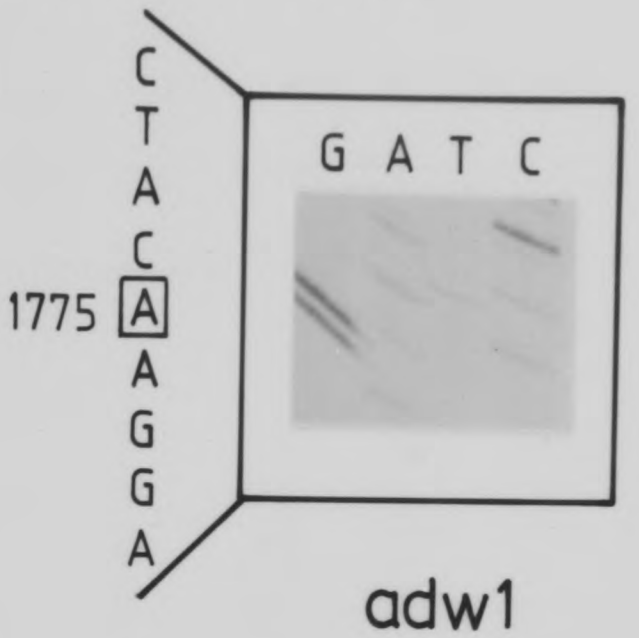
Figure 3.3.3 Nucleotide sequence of S genes mutated singly at the adw1, adw2, and adw3 sites

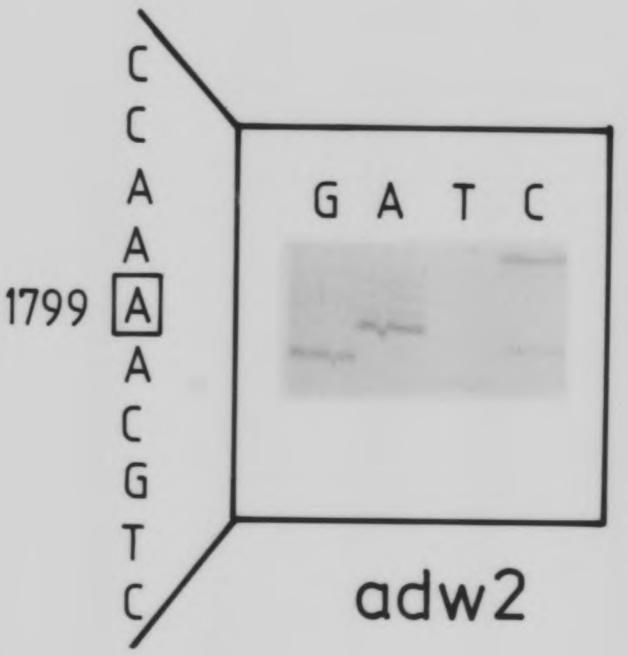
The sequence of the mutated DNA was determined by performing dideoxy chain terminator reactions (Sanger et al., 1977) which were resolved on 6% (w/v) urea-polyacrylamide gels (section 2.2.2).

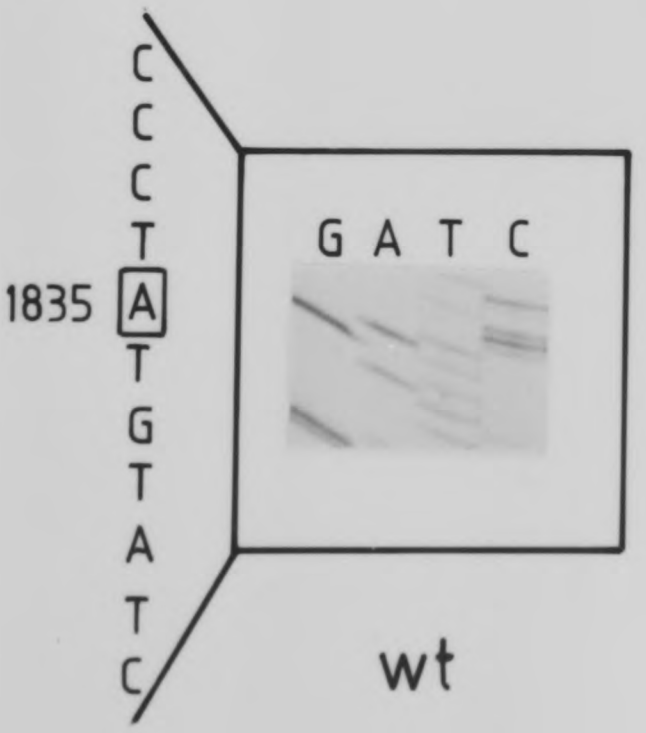
The nucleotide position where wild type (wt) nucleotide was substituted for mutant nucleotide is indicated by a box.

SINGLE adw MUTATIONS









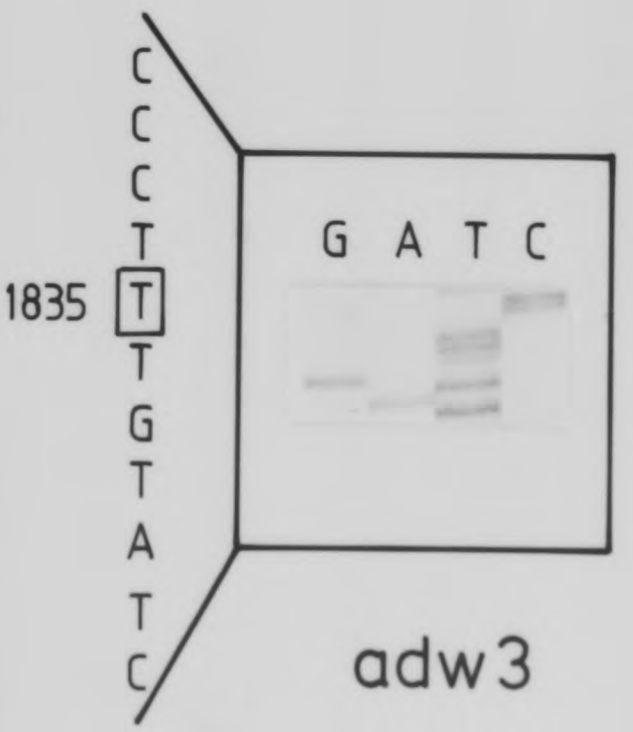
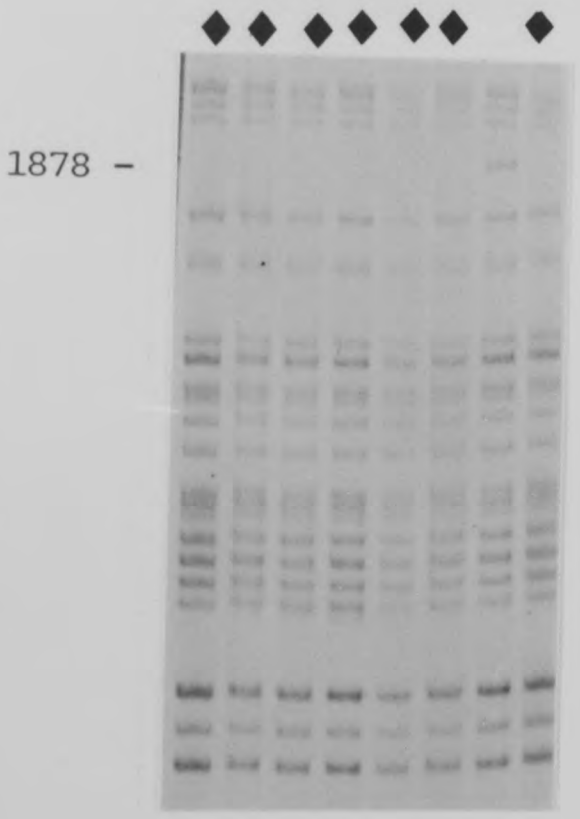


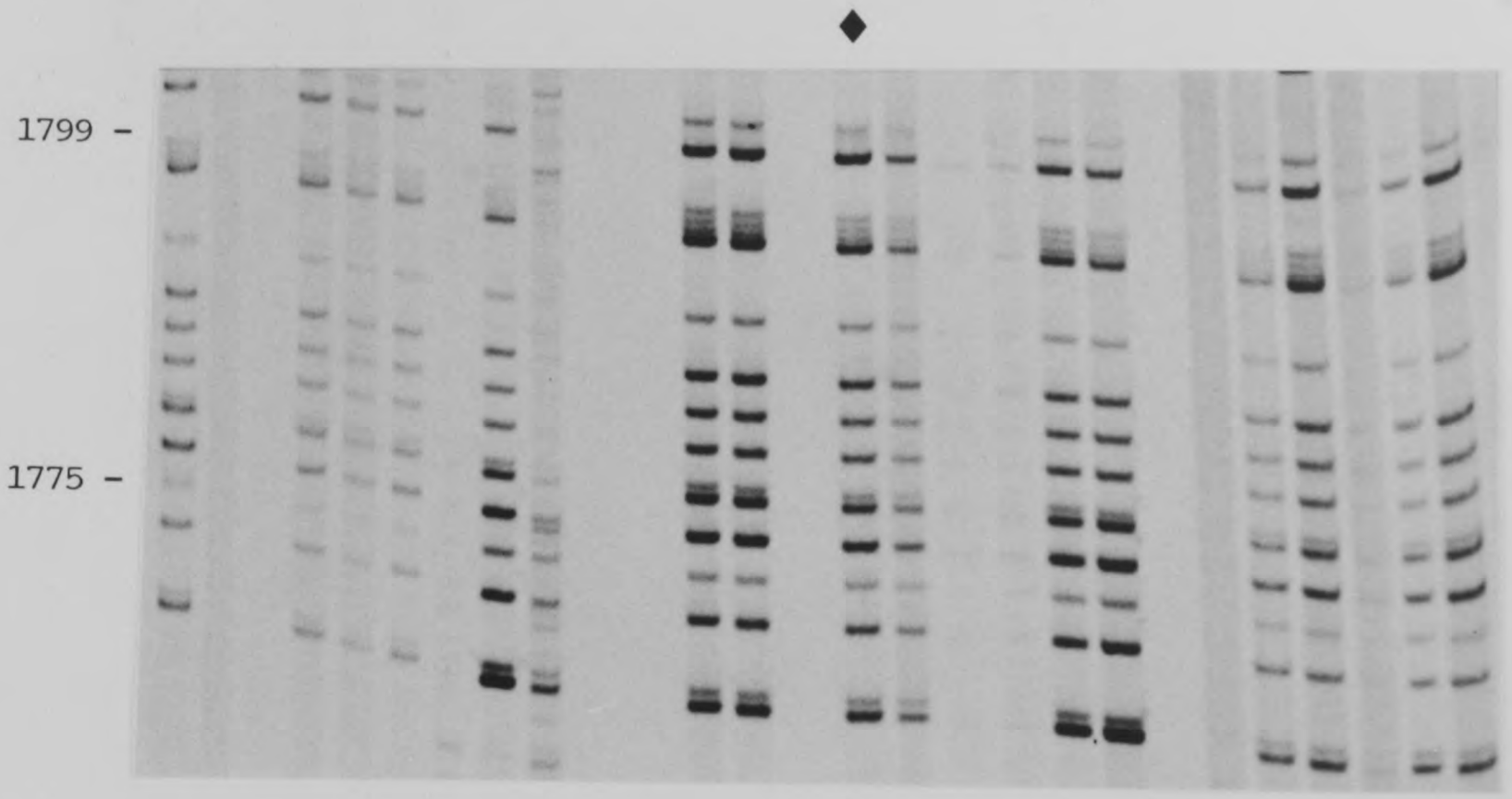
Figure 3.3.4 Single dideoxy nucleotide screening for mutants

A. T° screen of 8 single-stranded DNA preparations from phage produced after thio-deoxy C mutagenesis of S gene by a4 oligomer (section 3.4). The desired nucleotide substitution $1878^{T \rightarrow A}$ was present in 7/8 templates sequenced (filled diamonds) and was inferred by the loss of the T at position 1878.

B. A° screen of 16 single-stranded DNA preparations from phage produced after Gapped-Duplex mutagenesis of adw2 S gene by adw1 oligonucleotide (section 3.3). The desired nucleotide substitution $1775^{T \rightarrow A}$ was present in 1/16 templates sequenced (filled diamonds) and was inferred by the appearance of an extra A band at 1775. The extra A nucleotide present at 1799 was consistent with starting template being an adw2 mutant.



A



B

Figure 3.3.5 Nucleotide sequence of S genes harbouring double adw mutations

The sequencing of mutated DNA was as described in figure 3.3.3.

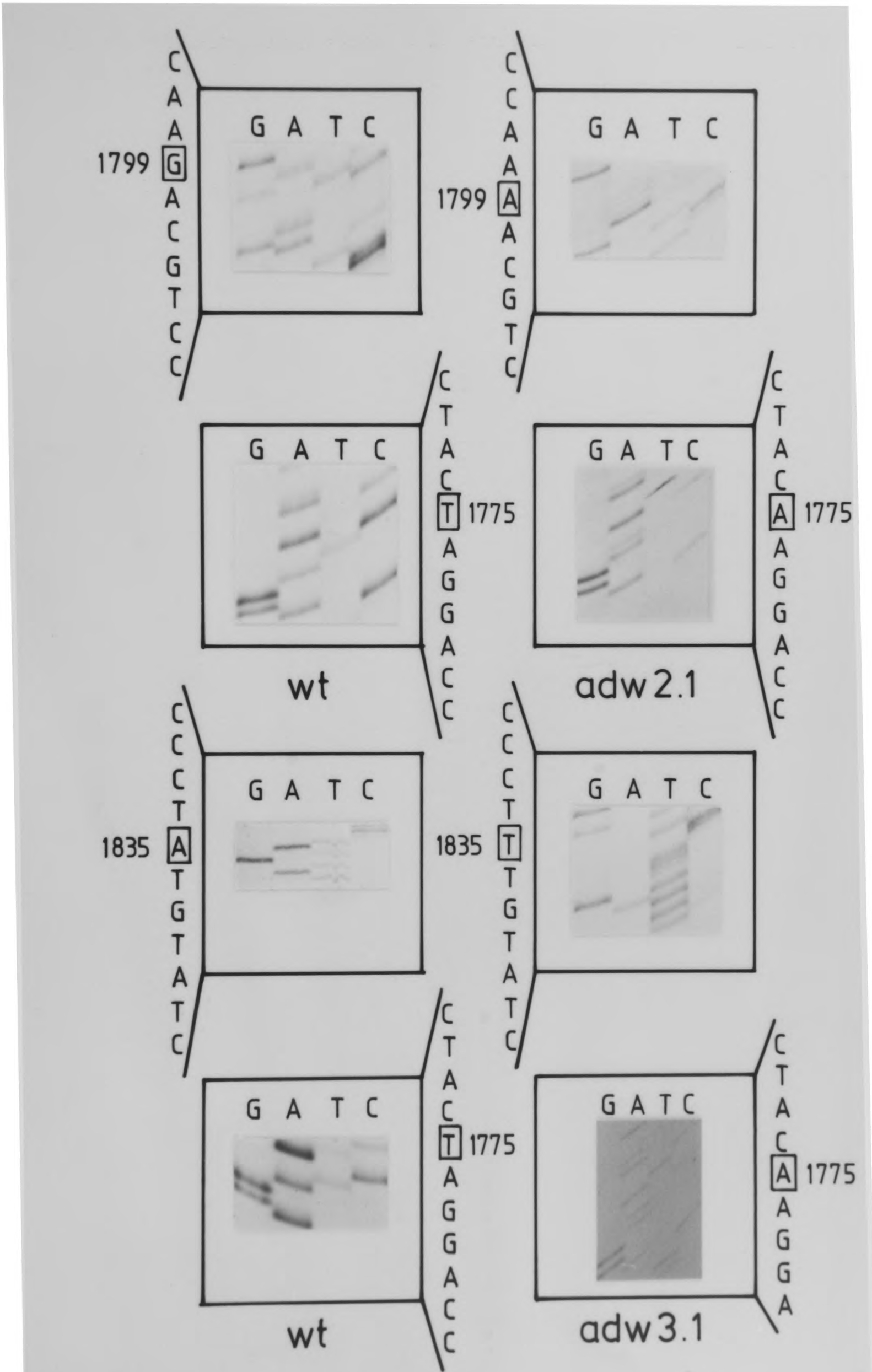
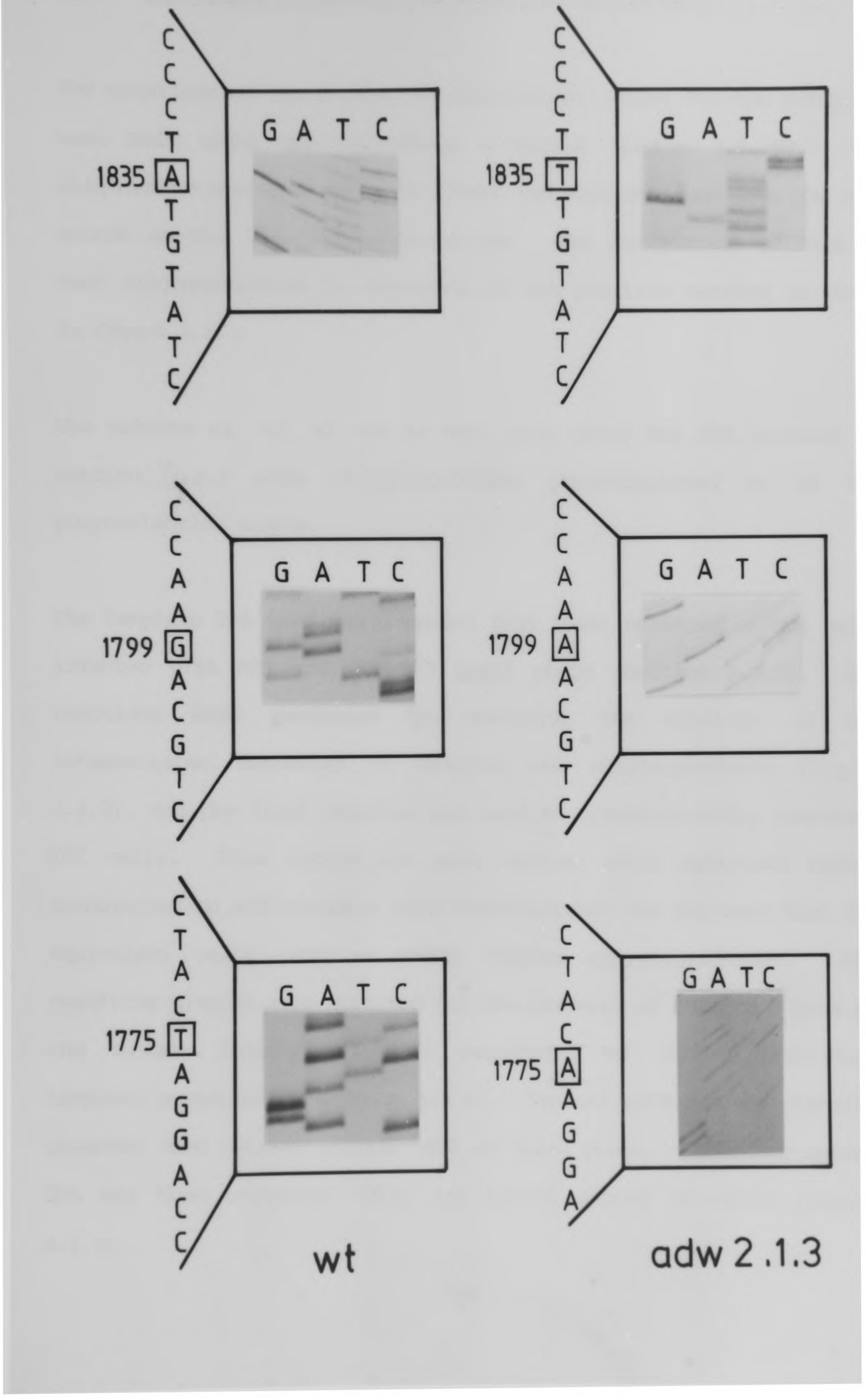

1799 G
CAAGACGTCC
G A T C
1799 A
CCAAAACGTC
G A T C
wt
CTACTAGGACC
T 1775
G A T C
adw2.1
CTACAAGGACC
A 1775
G A T C
1835 A
CCTTATGTATC
G A T C
1835 T
CCTTTGTATC
G A T C
wt
CTACTAGGACC
T 1775
G A T C
adw3.1
CTACAAGGA
A 1775
G A T C

Figure 3.3.6 Nucleotide sequence of S gene harbouring the triple adw mutation

The sequencing of mutated DNA was as described in figure 3.3.3.


C C C T
1835 A
T G T A T C
G A T C
C C C T
1835 T
T G T A T C
G A T C
C C A A
1799 G
A C G T
G A T C
C C A A
1799 A
A C G T
G A T C
C T A C
1775 T
A G G A C C
G A T C
wt
C T A C
1775 A
A G G A
G A T C
adw 2 .1.3

3.4 Thio-deoxy C mediated Site-Directed Mutagenesis

The mutations of the S gene, at nucleotides coding for the a-region, were made using the thio-deoxy C method (figure 3.1.1B). The oligonucleotides used were all 17mers and were obtained from the same source as the 'adw' oligonucleotides. The designation of name to each oligonucleotide in reference to the position mutated is shown in figure 3.3.1.

The mutants a1, a2, a3 and a4 were made using the SDM protocol of section 2.2.3 with oligonucleotides phorphorylated by T4 DNA polynucleotide kinase.

The template DNA used was prepared from 100ml cultures of TG1 cells infected with M13 mp8.pHIN/G-2 BamHI phage (section 2.2.3). The reactions were monitored by observing the mobility of DNA intermediates subjected to agarose gel electrophoresis (figure 3.4.2), and the final reaction mix used to transform $CaCl_2$ competent TG1 cells. This strain was used because cells exhibited higher transformation efficiencies with hemithiolated DNA duplexes than did equivalent cells such as NM522 (Taylor et al., 1985a). The resulting plaques were screened for the presence of mutated S gene by the single dideoxy-nucleotide sequencing of twenty individual template preparations (figure 3.3.4). Typical mutation efficiencies observed were between 65 and 90% of total phage. Candidate mutant DNA was then sequenced fully and mutant stored at -20°C (figure 3.4.3).

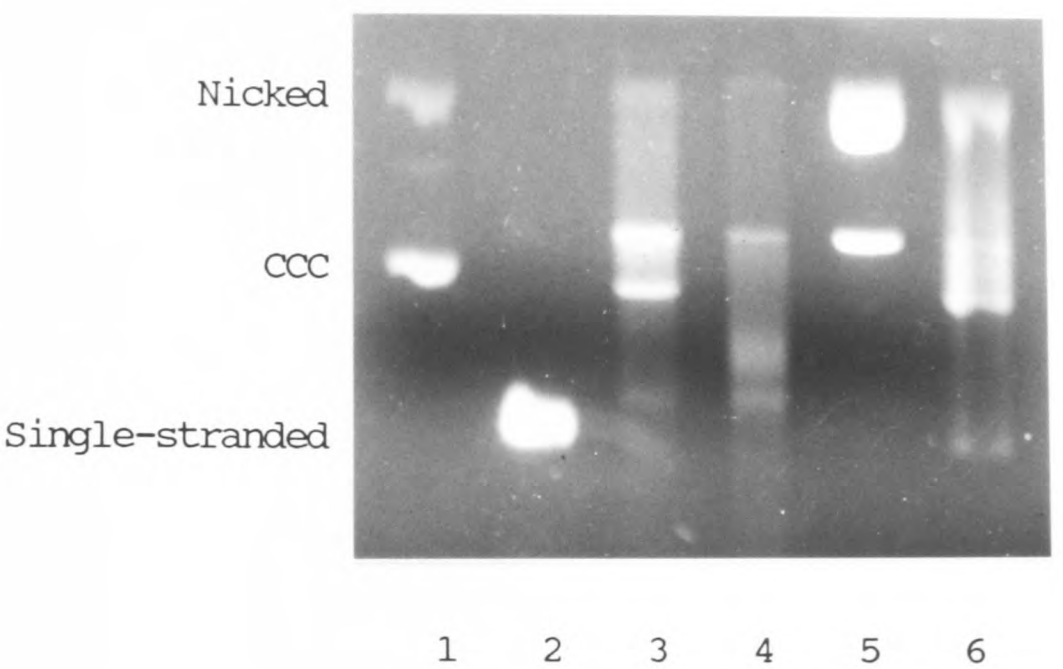
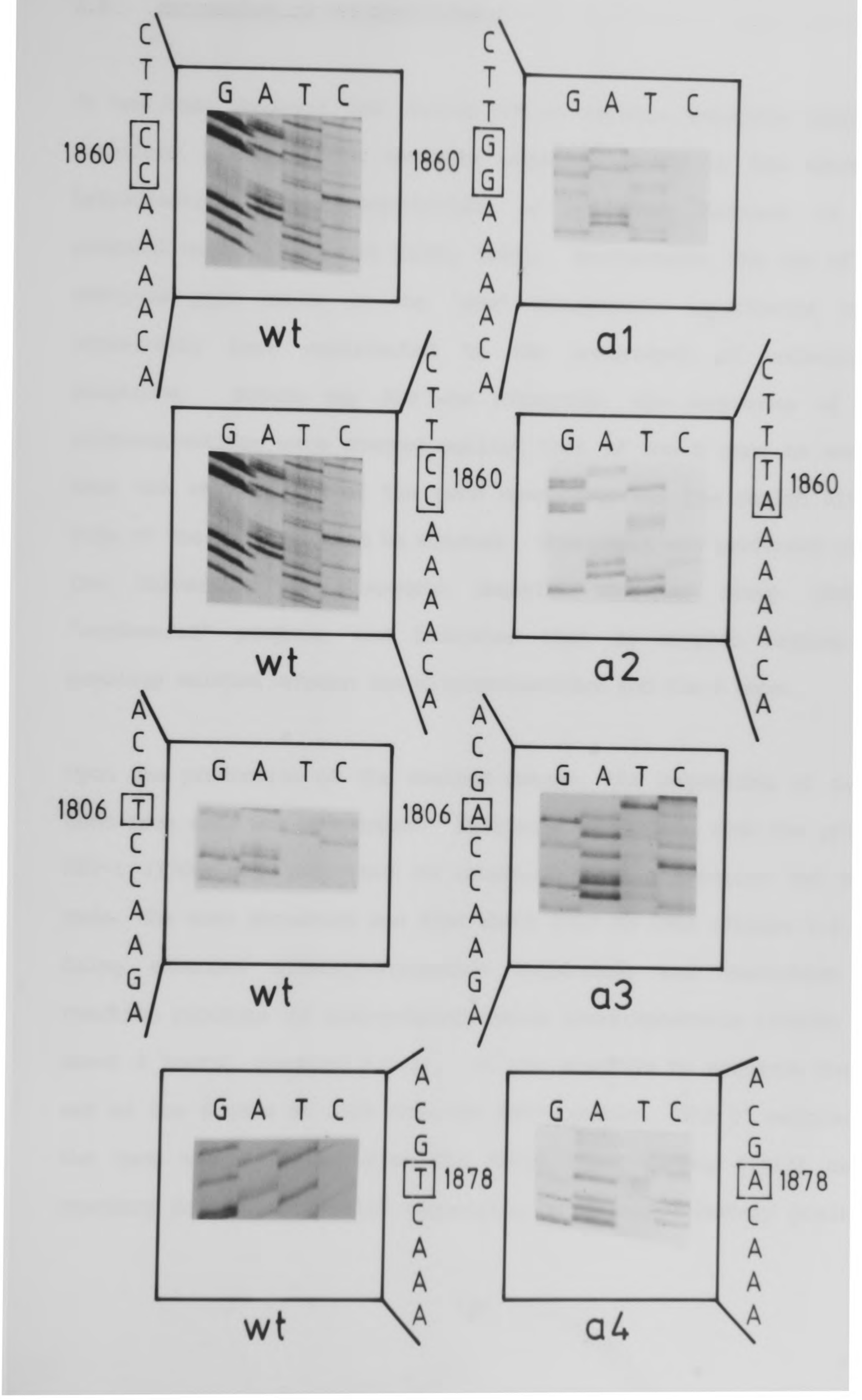


Figure 3.4.2

Photograph of 1% (w/v) agarose gel showing products of thio-deoxy C SDM. Sample 1 - 200ng double-stranded RF DNA; Sample 2 - 200ng single-stranded DNA template; Sample 3 - DNA after repolymerisation; Sample 4 - DNA after exonuclease III digestion (note presence of partially single-stranded DNA); Sample 5 - DNA after filtration and nicking (note absence of single-stranded DNA and large amount of nicked DNA); Sample 6 - DNA after extension and ligation.

Figure 3.4.3 Nucleotide sequences of S genes mutated at the a1, a2, a3 and a4 sites

The sequencing of mutated DNA was as described in figure 3.3.3.


G A T C
1860
C T T C C A A A A C A A
wt
1860
C T T G G A A A A C A A
a1
1860
C T T C C A A A A C A A
wt
1860
C T T T A A A A A C A A
a2
1806
A C G T C C A A G A A
wt
1806
A C G A C C A A G A A
a3
1878
A C G T C A A A A
wt
1878
A C G A C A A A A
a4

3.5 Sequencing of Mutated S Genes

It has been observed that during SDM of certain sequences spurious mutations arise in the template sequence caused by the aberrant hybridisation of oligonucleotides or by other features of the protocol used (Zoller and Smith, 1983). Furthermore, the use of the BMH71-18 mutL cells in the 'adw' mutagenesis experiments could conceivably have contributed to the occurrence of undesirable mutations. Before any SDM was attempted, the sequences of the oligonucleotides were checked against that of the S gene to ensure that the only region of the gene homologous was the region either side of the nucleotide to be mutated. This check was performed using the University of Wisconsin Genetics Computer Group (UWGCG) "wordsearch" program, and indicated that no cryptic regions of homology existed between the oligonucleotides and the S gene.

Upon the production of the desired mutant, the sequencing of the S gene from each was undertaken. Initially sequencing from the primer HBV-1 (1704) was performed to establish that a mutation had been made, the area sequenced was from about 1710 to 1950 (figure 3.2.1). Using extended dideoxy-nucleotide sequencing and resolution of reaction products by urea-polyacrylamide electrophoresis (50V/cm for about 6 hours) (section 2.2.2), it was possible to sequence the 3' end of the S gene at 2114 from the HBV-1 primer. The 5' section of the gene was sequenced from the HBV-2 17mer primer (1417) using standard dideoxy-nucleotide sequencing up to approximately position

1730. The full sequence of the S genes from all 10 mutants revealed no undesirable mutations.

4. EXPRESSION OF HBsAg IN S. CEREVISIAE

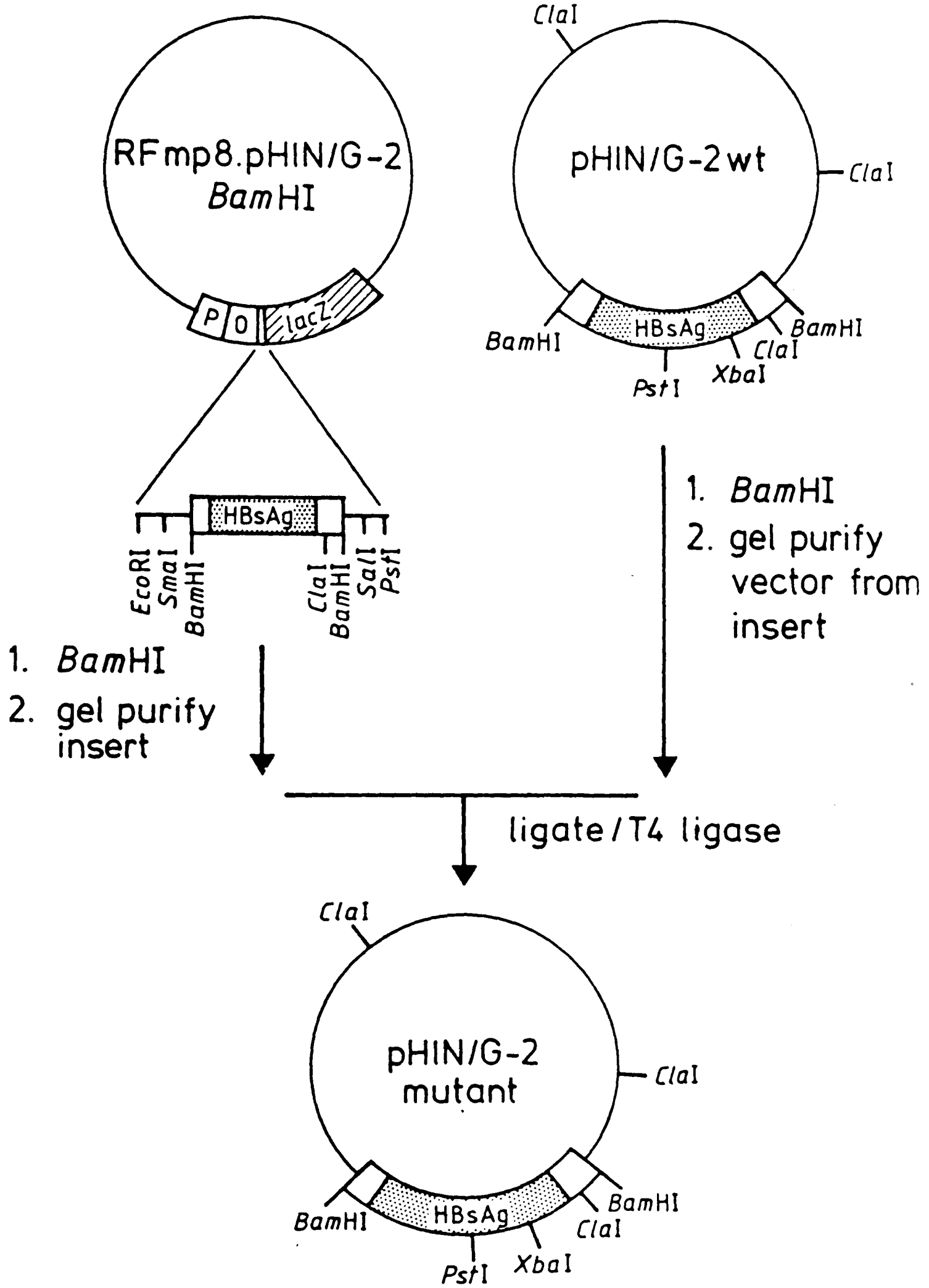
4. Expression of HBsAg in S. cerevisiae

In order to assess the phenotypic implications of the 'adw' and a group mutations made in section 3, it was necessary to clone and express the mutant S genes in yeast. This was done using the E. coli/ S. cerevisiae shuttle vector pHIN/G-2 (figure 3.2.1; Murray et al., 1984) to express the mutant genes in cells from the appropriate S. cerevisiae strain.

4.1 Sub-Cloning Mutant S Gene into pHIN/G-2

Sequenced mutant template (10ng) was used to transform TG1 or NM522 cells and the resulting plaques used to prepare RF DNA (section 2.2.1) to a final concentration of 0.1μg. Resulting RF M13 DNA (5μg) was digested with BamHI (10U) as was pHIN/G-2 DNA (5μg) which contained the original 1.78kb BamHI insert (figure 4.1.1). Both digests were subjected to agarose gel electrophoresis; the 1.78kb fragment from the M13 clone and the 8.67kb fragment from pHIN/G-2 were then purified from low melting point agarose and resuspended at a concentration of 0.1μg in TE. Using T4 DNA ligase (1U) the two BamHI fragments were ligated together (0.25μg 1.78kb fragment, 0.1μg 8.67kb fragment) and used to transform $CaCl_2$ competent HB101 cells. Selection for transformants was afforded by plating on L-amp plates (100μg/ml) and recombinants detected by colony hybridisation (section 2.2.1) of nitrocellulose-immobilised colonies with pHIN/G-2 1.78kb BamHI probe (10^8cpm/μg) prepared by random-primed, polymerase reactions (section 2.2.1). In order to establish that contaminating

Figure 4.1.1 Sub-cloning of the mutated S gene containing BamHI fragment into pHIN/G-2


RFmp8.pHIN/G-2
BamHI
P
O
lacZ
HBsAg
EcoRI
SmaI
BamHI
ClaI
BamHI
SalI
PstI
pHIN/G-2wt
ClaI
ClaI
BamHI
HBsAg
BamHI
ClaI
PstI
XbaI
1. BamHI
2. gel purify insert
1. BamHI
2. gel purify vector from insert
ligate / T4 ligase
pHIN/G-2
mutant
ClaI
ClaI
BamHI
HBsAg
BamHI
ClaI
PstI
XbaI

1.78kb fragment, from original pHIN/G-2, did not co-purify with the extracted 8.67kb fragment used for cloning the mutant fragments, colonies from self-ligated 8.67kb DNA were screened alongside those derived from every mutant fragment ligation. None of these colonies gave a positive signal in hybridisation experiments; whereas the typical frequency of occurrence of positive colonies from ligations involving isolated mutant fragments was between 10% and 20% of total colonies. Hybridisation-positive colonies were analysed for the orientation of the BamHI insert by digestion of mini-preparation, plasmid DNA with ClaI (figure 4.1.1) followed by agarose gel electrophoresis (figure 4.1.2). The correct pHIN/G-2 recombinant plasmids were then produced from 500ml L-broth cultures (containing ampicillin at 100μg/ml), and stored at a final concentration of 1μg/ml.

4.2 Expression of HBsAg in Yeast

Spheroplasts of the yeast cells DBY746 (LEU^-) were prepared, then transformed with pHIN/G-2 plasmid DNA (LEU^+) and plated in 3% (w/v) agar on leucine deficient minimal plates. Transformation efficiencies were about 1 x $10^2/\mu g$. Yeast cells were grown in minimal synthetic medium (section 2.1.2; Schurr and Yagil, 1971) containing 1g/l. KH_2PO_4, to an OD_{600nm} of about 2.0, then diluted into low phosphate medium (30mg KH_2PO_4/l.) and grown again to OD_{600nm} 2.0 inducing HBsAg production by derepressing transcription from the PHO5 promoter. Preparation of extracts from cells of 1l. cultures was performed as in section 2.2.4 (Murray et al., 1984). Total

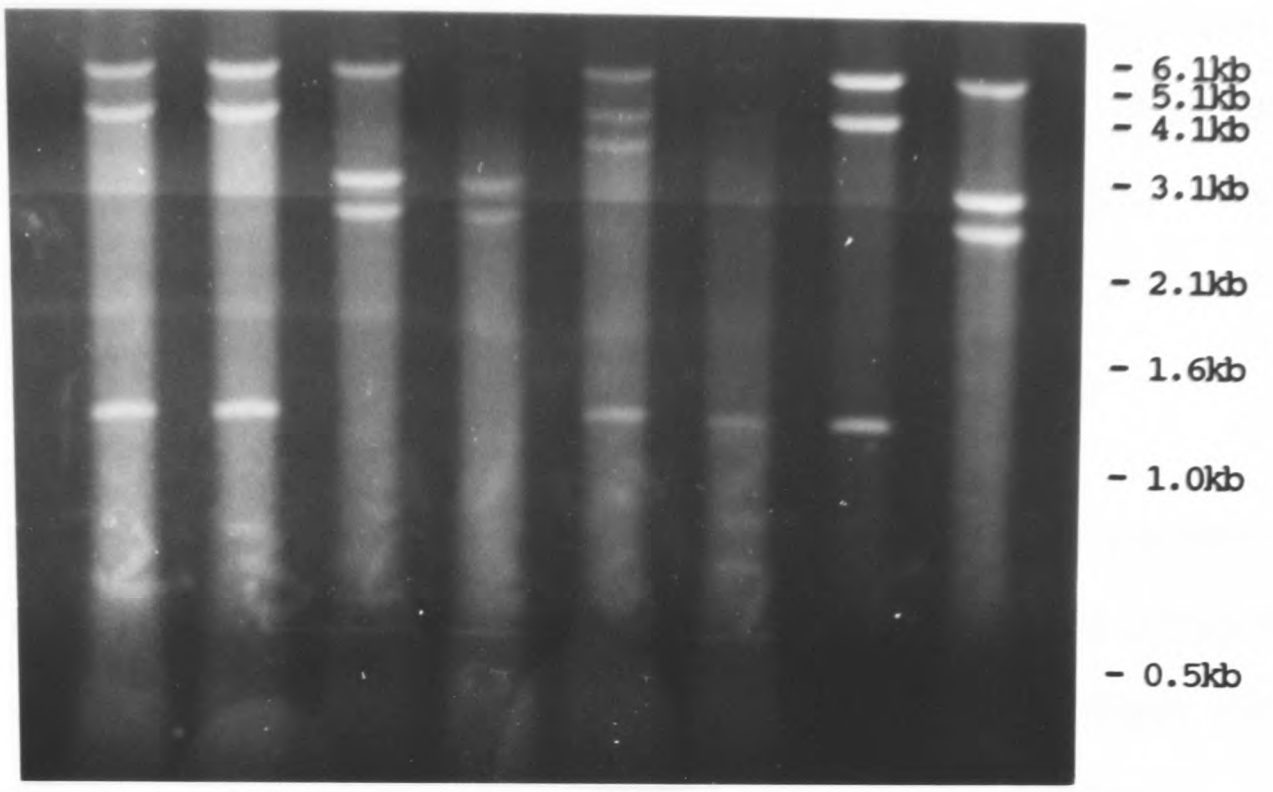


Figure 4.1.2

Separation of ClaI digest products from mini-plasmid preparations of pHIN/G-2 BamHI recombinants, by electrophoresis through 1% (w/v) agarose. Depending on the orientation of the 1.78kb BamHI fragment, the sizes of ClaI generated fragments were: 1.2kb, 4.1kb, 5.0kb or 2.5kb, 2.7kb, 5.1kb. The former set of ClaI fragments indicated the desired orientation of BamHI insert in pHIN/G-2.

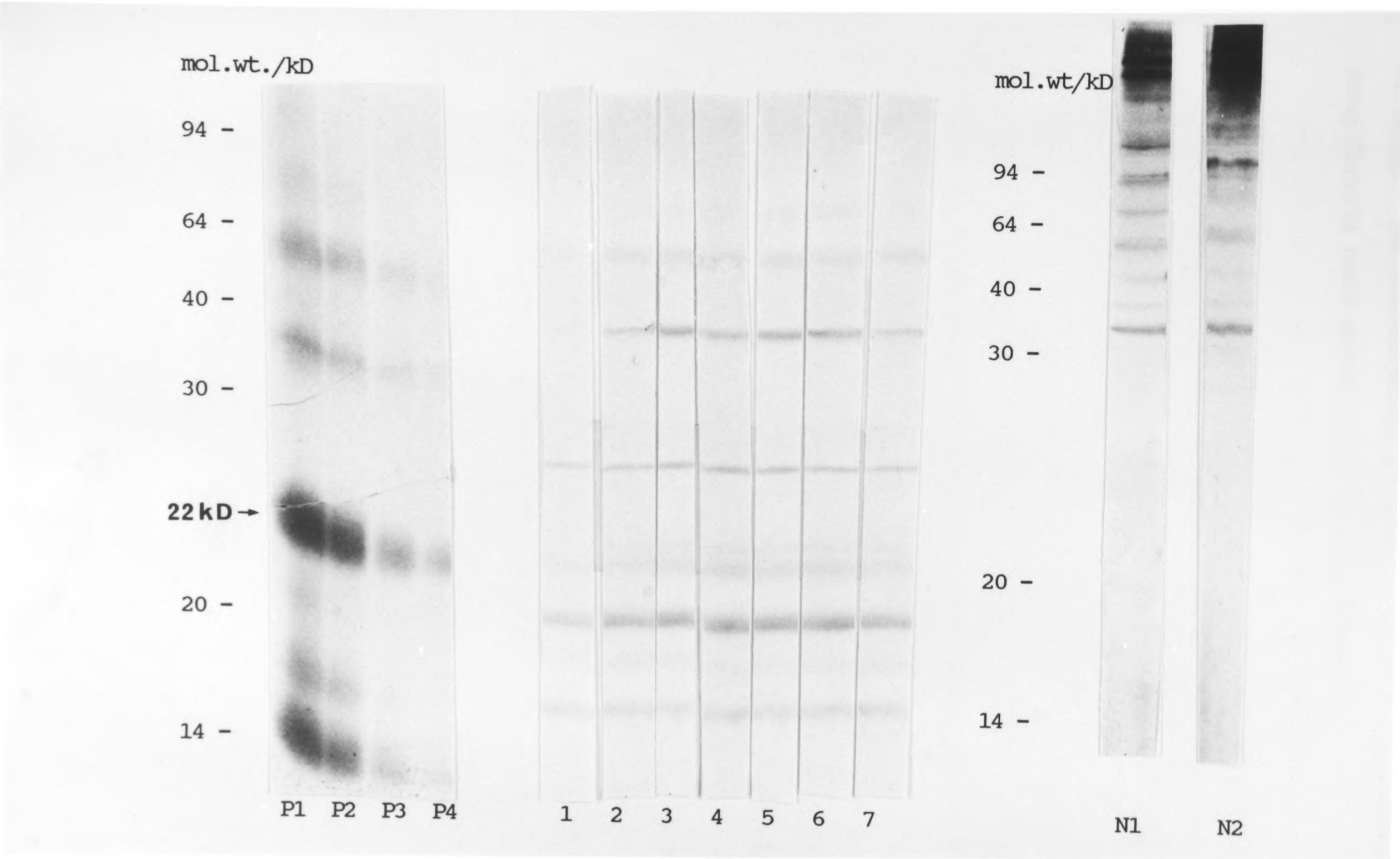
protein extracts were prepared from 1l. cultures of all the pHIN/G-2 'adw' mutants as well as from the unmutated strain. The level of antigen was estimated using the Abbott Ausria II radioimmunoassay system for detection of human HBsAg. The system was calibrated for use with yeast derived HBsAg, by testing with 99% pure pHIN/G-2 derived HBsAg (a gift from the Green Cross Corporation). The level of HBsAg present in the extracts was about 50ng/mg protein (about 0.05% total cell protein).

Estimation of expression levels for pHIN/G-2 clones of mutants in the a-region could not be performed by the Ausria II system since antigen produced may have had impaired reactivity in the HBsAg detection system. To estimate the amount of HBsAg produced in 50l. fermentations of cells harbouring unmutated, 'adw' and a-region mutated S genes (section 2.2.4), a Western blot experiment was performed. An IgG preparation of a rabbit antiserum (NRA/BC-4B) was used in the experiment (section 2.1.7). The antiserum was raised against SDS treated, reduced HBsAg produced from pHIN/G-2 transformed yeast. The antiserum was considered as having anti-denatured HBsAg specificity, therefore in the experiment detection of HBsAg was not affected by any antigenic property of the native protein (figure 4.2.1).

Using this approach the level of denatured HBsAg present in ammonium sulphate fractions (20 - 35% saturation) of protein from cells containing S genes mutated in the a-region (a1 - a4), at positions $122^{arg \rightarrow lys}$ (adw2), $134^{tyr \rightarrow phe}$ (adw3) and wild type were

Figure 4.2.1 Detection of denatured HBsAg using immunoblotting

Protein samples were electrophoresed at 4.5V/cm through a 13% (w/V) SDS-polyacrylamide gel (Laemmli, 1970) then immunoblotted on to nitrocellulose (Towbin et al., 1979). Denatured HBsAg was detected by incubation with NRA/BC-4B anti-denatured yeast HBsAg serum, then goat anti-rabbit IgG Fc-alkaline phosphatase conjugate. Using transmission densitometry, the degree of staining was related to the amount of the 22kD form of HBsAg present using the positive control samples (P) (99% pure, pHIN/G-2 derived HBsAg of concentration 0.6mg/ml), of which 40% was in the 22kD form, as standards. Lanes: P1 - 12μg 22kD HBsAg, P2 - 6μg 22kD HBsAg, P3 - 2.4μg 22kD HBsAg, P4 - 1.8μg 22kD HBsAg; 1 - 7 were ammonium sulphate fractions (20 - 35% saturation; 3.5mg protein) of DBY746 cell extracts, 1 - unmutated, 2 - $124^{cys \rightarrow ser}$, 3 - $142^{pro \rightarrow gly}$, 4 - $142^{pro \rightarrow ile}$, 5 - $147^{cys \rightarrow ser}$, 6 - $122^{arg \rightarrow lys}$, 7 - $134^{tyr \rightarrow phe}$. N1 and N2 were ammonium sulphate fractions (20 - 35% saturation; 3.5mg protein) of extracts of DBY746 cells not transformed by pHIN/G-2.


mol.wt./kD
94 -
64 -
40 -
30 -
22 kD →
20 -
14 -
P1
P2
P3
P4
1
2
3
4
5
6
7
mol.wt/kD
94 -
64 -
40 -
30 -
20 -
14 -
N1
N2

estimated. The concentrations observed ranged from about 160ng to 300ng HBsAg/mg total protein (0.56 to 1.5μg HBsAg/ml).

5. EFFECT OF MUTATIONS ON HBsAg IMMUNOREACTIVITY

5. Effect of Mutations on HBsAg Immunoreactivity

5.1 Mutations and d/y subtype

5.1.1 Assay for HBsAg subtype

The effect of mutations at $113^{ser \rightarrow thr}$ (adw1), $122^{arg \rightarrow lys}$ (adw2), and $134^{tyr \rightarrow phe}$ (adw3) on the d/y subtype of HBsAg was determined using an assay which examined the reactivity of the protein with d and y monospecific antisera. A modification of a double antibody-radioimmunoprecipitation (DARIP) was used for the subtyping assay (Burrell et al., 1978; section 2.2.4). The extent of inhibition of an anti-y or anti-d:[^{125}I]-HBsAg (adw or ayw) reaction by test antigen was taken as an indication of the degree of avidity the antigen had for the monospecific antiserum (figure 5.1.1). The specificity of the assay was demonstrated using a series of standard HBsAg preparations of known subtype (figure 5.1.1B). The reactivity of HBsAg with a subset of anti-HBsAg antibodies was the original feature used in the definition of subtype (Le Bouvier et al., 1971).

5.1.2 Effect of mutations on d/y subtype

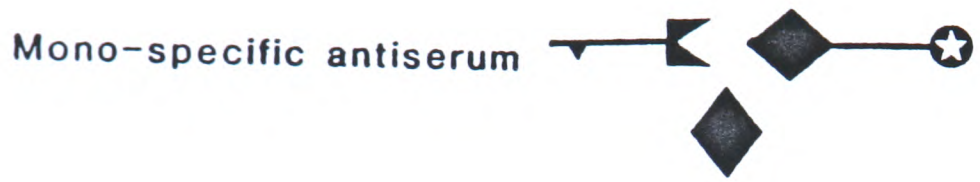
The subtype of unmutated HBsAg produced from the S gene of pHIN/G-2 (wt) was defined as y^+d^- by competition DARIP (figure 5.1.2). The three single amino acid substitutions in HBsAg (adw1, adw2, adw3) had no detectable influence on the d/y subtype.

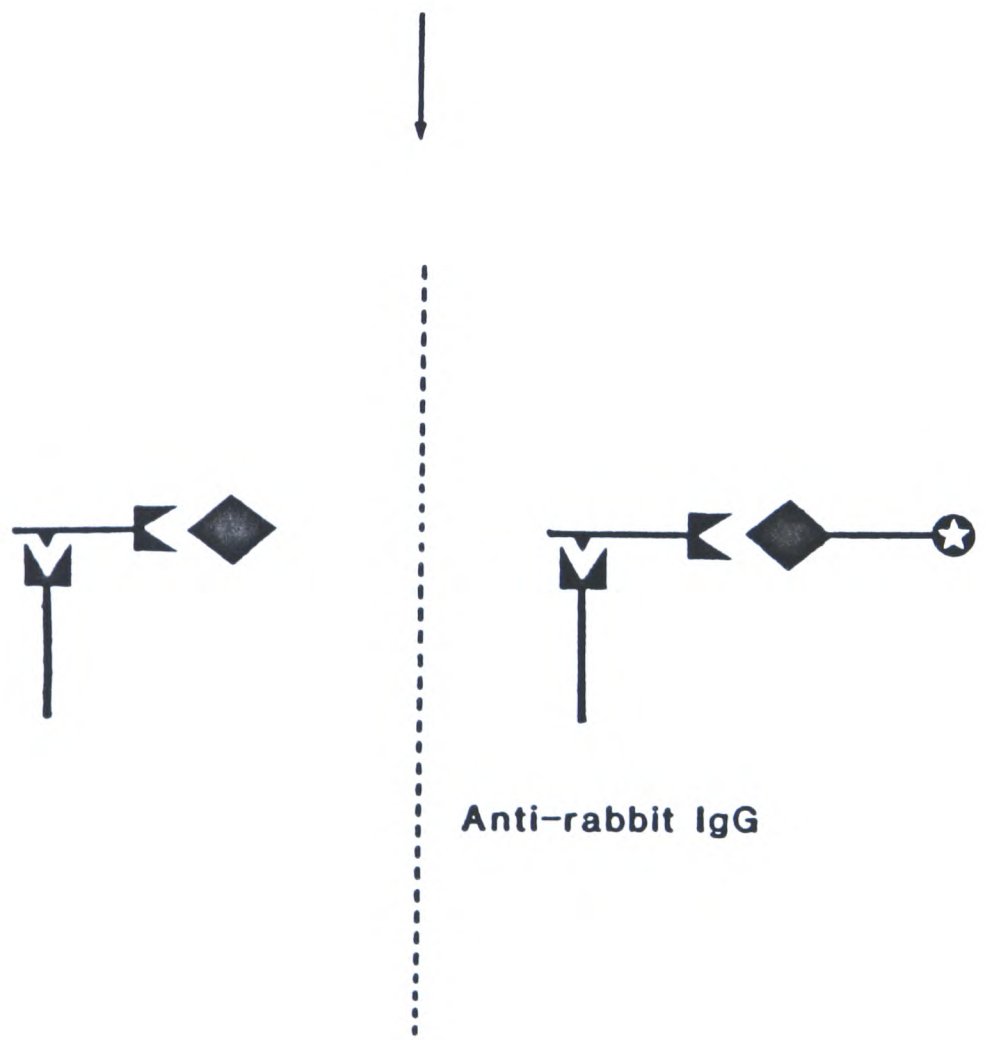
Figure 5.1.1 Use of monospecific antisera to define the d/y subtype of HBsAg

A. The relative avidity of test antigen for d or y monospecific antisera was indicated by the extent of inhibition of the reaction between monospecific antiserum and labelled HBsAg of subtype adw or ayw. The complexes formed between antibody and antigen were precipitated by donkey anti-rabbit IgG serum; with the extent of reduction in label precipitated indicating the relative avidity of test antigen for monospecific antiserum.

B. Standard HBsAg preparations, of known subtype, were used to establish the specificity of the competition double antibody radioimmunoprecipitation assay (Burrell et al., 1978). The reaction between monospecific anti-d or anti-y sera and labelled HBsAg of subtype adw or ayw was inhibited in a subtype specific manner by the panel of HBsAg proteins (obtained from Dr J Peutherer).

A





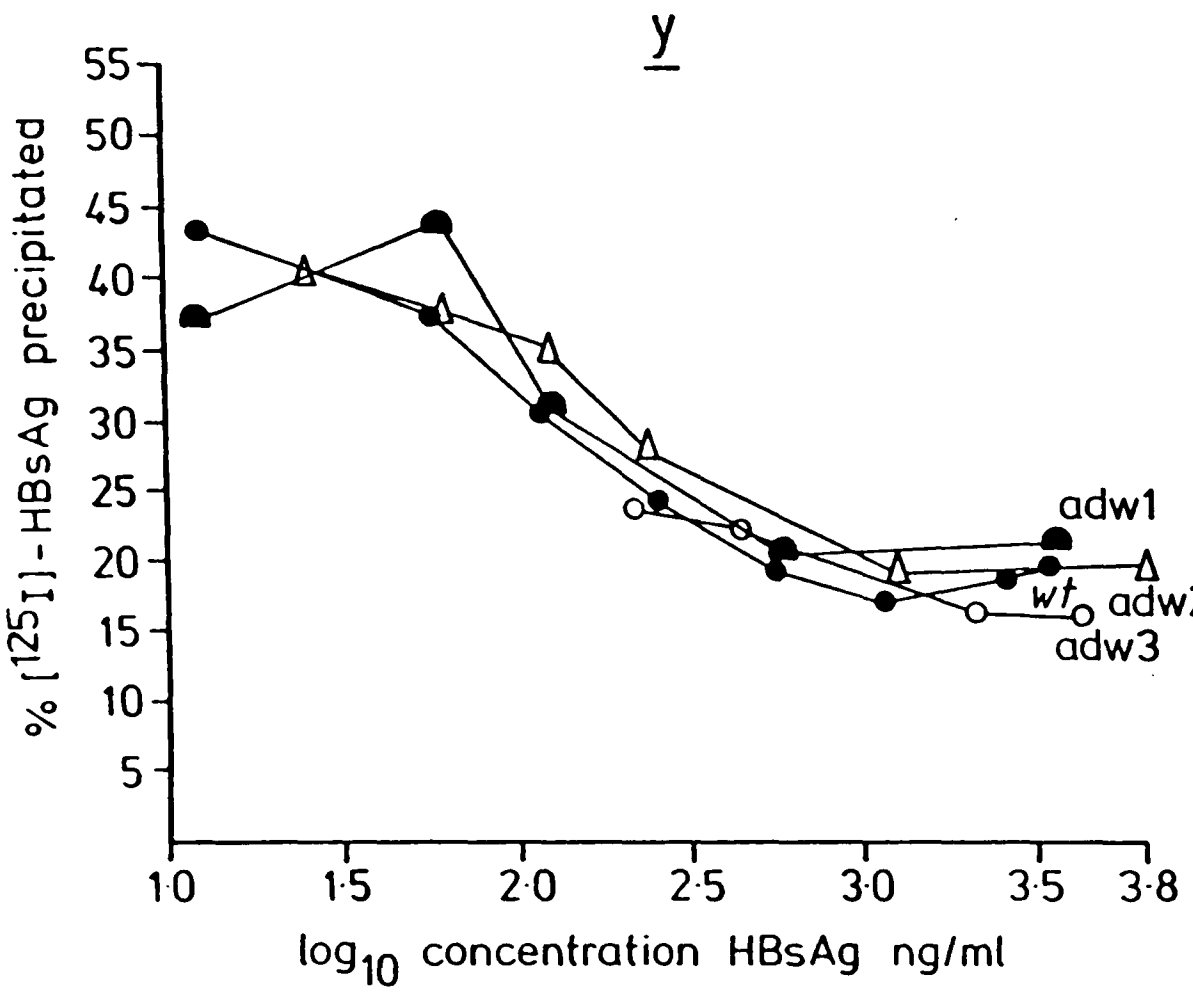
B

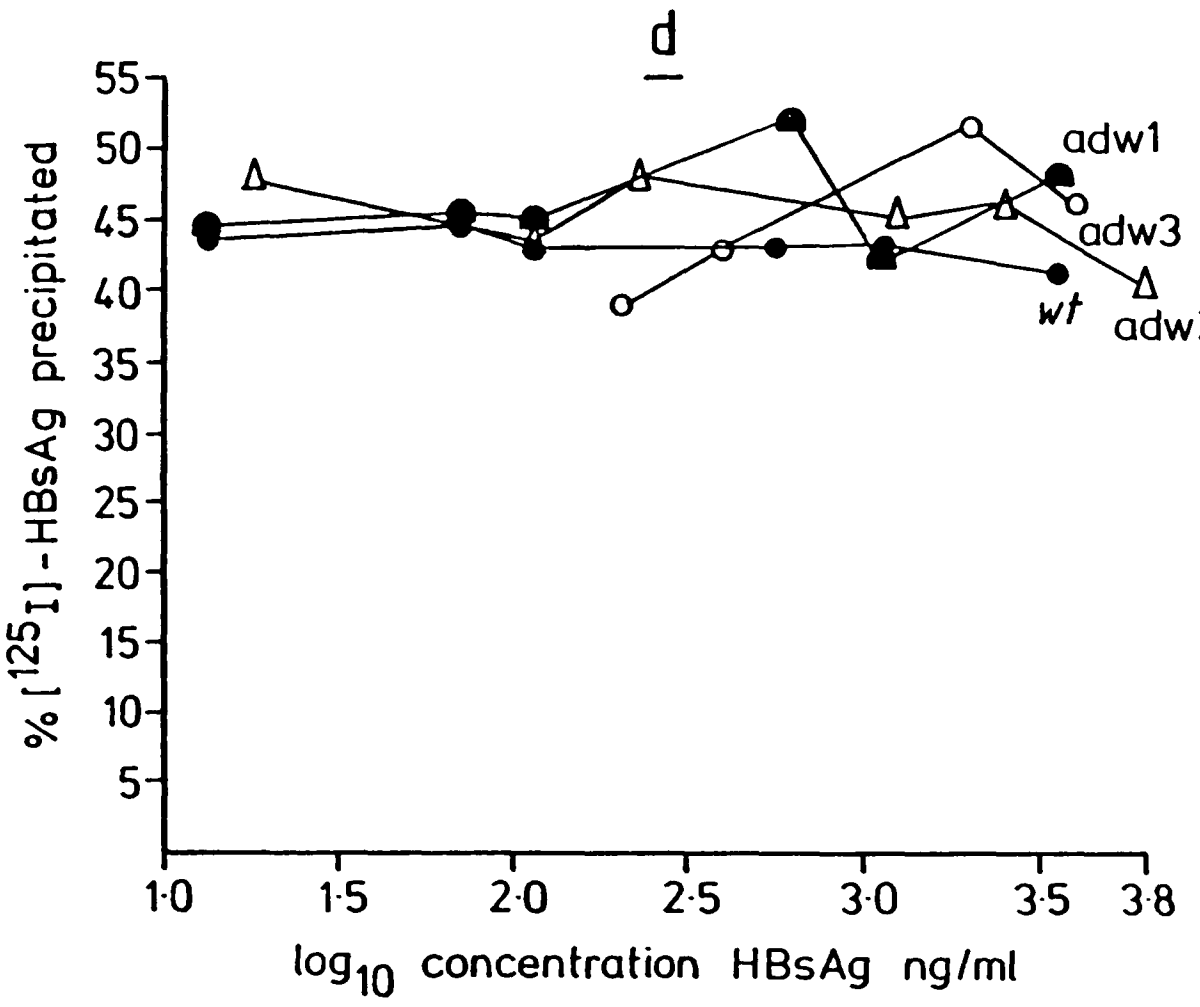
Controls on Competition RIP Assay		
% [^{125}I]-HBsAg precipitated		
	y	d
Assay buffer (RIP)	47	44
ayw1	23	44
ayw2	16	44
ayw3	20	43
ayw4	29	46
ayr	19	42
adw2	49	8
adw4	50	9
adr	44	7

controls at concentration 50ng/ml

Figure 5.1.2 Effect of single amino acid substitutions on the d/y subtype of HBsAg

The subtype of S. cerevisiae expressed HBsAg was defined by the degree of inhibition of anti-d:HBsAg adw or anti-y:HBsAg ayw reaction. The concentration of antigen was determined by the Abbott Ausria II radioimmunoassay for human HBsAg, calibrated with a standard preparation of pHIN/G-2 derived HBsAg.


y
% [125I]-HBsAg precipitated
55
50
45
40
35
30
25
20
15
10
5
1·0
1·5
2·0
2·5
3·0
3·5
3·8
log10 concentration HBsAg ng/ml
adw1
wt
adw2
adw3


d
% [125I]-HBsAg precipitated
55
50
45
40
35
30
25
20
15
10
5
1·0
1·5
2·0
2·5
3·0
3·5
3·8
log10 concentration HBsAg ng/ml
adw1
adw3
wt
adw2

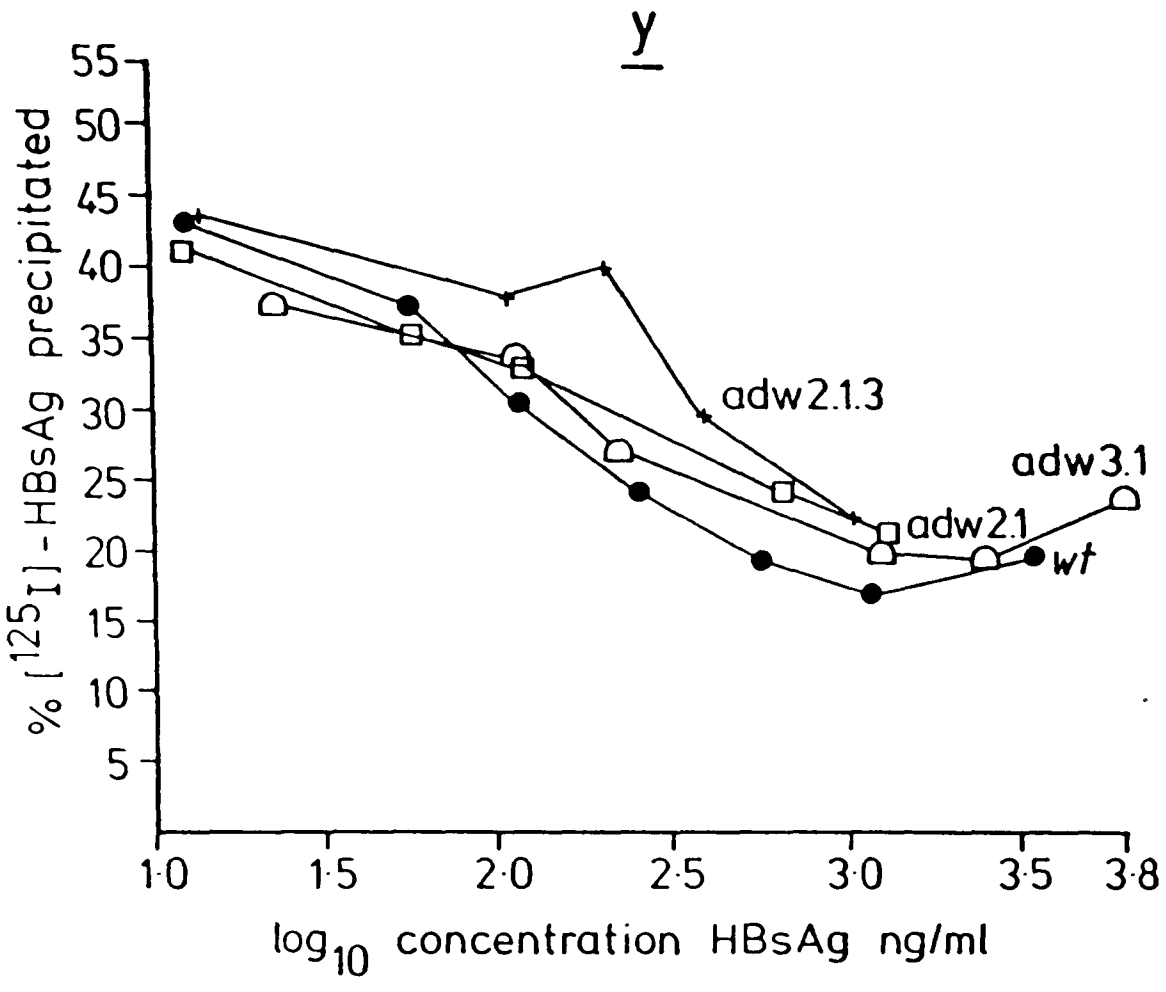
A combination of either mutation $122^{arg \rightarrow lys}$ or $134^{tyr \rightarrow phe}$ with $113^{ser \rightarrow thr}$ resulted in a switch from a $\underline{y}^+\underline{d}^-$ to a $\underline{y}^+\underline{d}^+$ subtype (figure 5.1.3). The increase in $\underline{d}$-reactivity was slightly more pronounced in adw2.1 HBsAg than adw3.1. The degree of $\underline{y}$-antigenicity was detectably diminished by both double mutations. The effect of combining all three mutations was to slightly increase $\underline{d}$-antigenicity and to further decrease $\underline{y}$-antigenicity. These results indicate that the appearance of the $\underline{d}$-subtype depends on the nature of the amino acids present at two positions. The complete switch from $\underline{y}$ to $\underline{d}$ can almost be achieved by the substitution of amino acids at three positions. However, the nature of the residue at one particular position, even if the residue at that position correlates exactly with a subtype (such as 122, figure 1.4.2), had no detectable bearing on the $\underline{d}/\underline{y}$ subtype of the antigen.

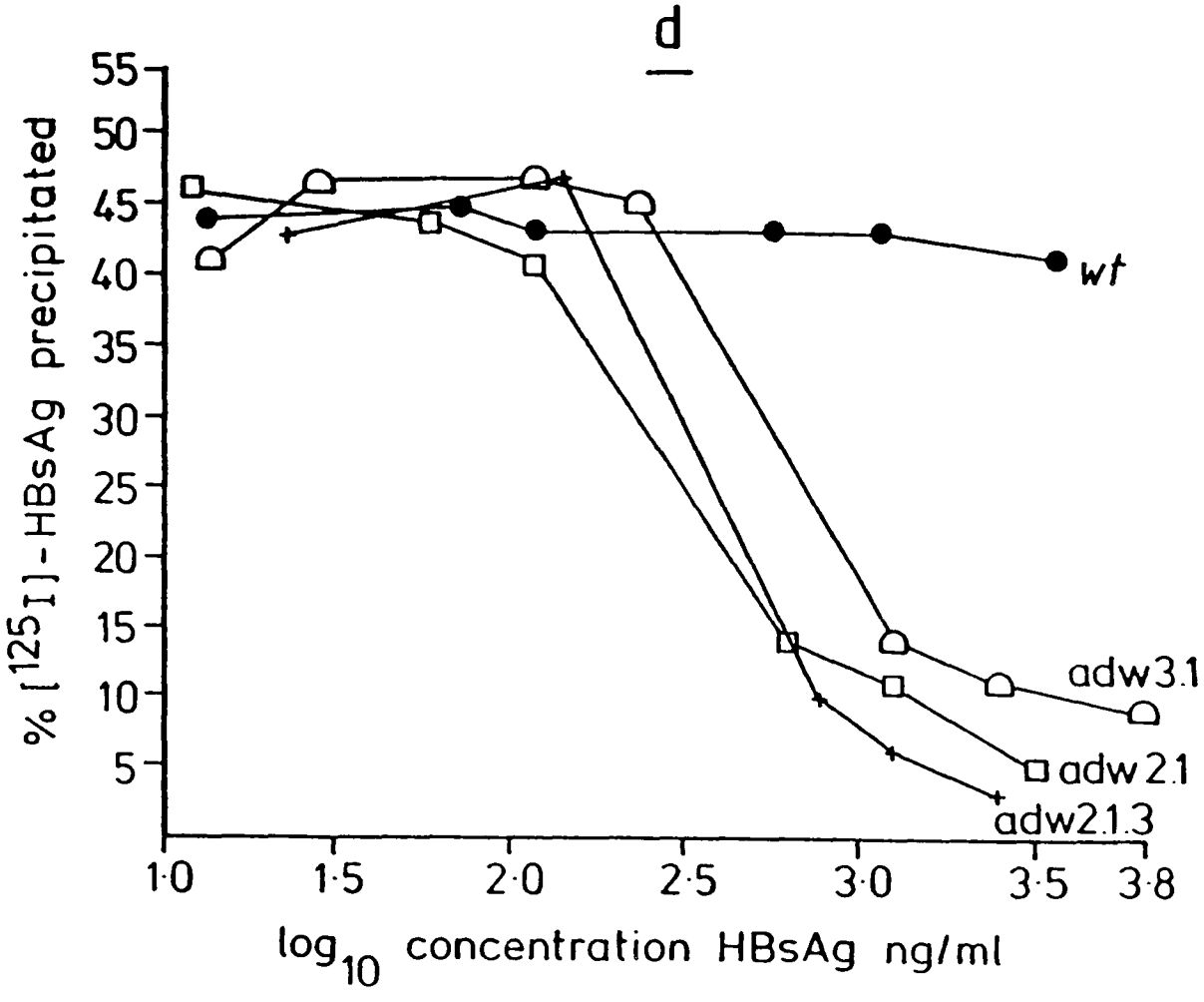
5.1.3 Effect of mutations on the binding of a $\underline{d}$-specific monoclonal antibody

The specificity of the mouse monoclonal antibody 1E6 (a gift from The Green Cross Corporation, Osaka) was demonstrated using HBsAg of subtype $\underline{adw}$ and $\underline{ayw}$ in a radioimmunoassay. The monoclonal antibody was immobilised on polyvinyl beads (60μg/ml antibody in 0:2M $NaHCO_3$ pH 9.2 for 16 hours at room temperature), and used with [^{125}I]-anti-human HBsAg second antibody according to the Abbott Ausria II protocol.

Figure 5.1.3 Effect of multiple amino acid substitutions on the d/y subtype of HBsAg

The assay, controls and determination of concentration of HBsAg were as described in figure 5.1.2.


y
55
50
45
40
35
30
25
20
15
10
5
% [125I]-HBsAg precipitated
adw2.1.3
adw3.1
adw2.1
wt
1·0
1·5
2·0
2·5
3·0
3·5
3·8
log10 concentration HBsAg ng/ml


d
55
50
45
40
35
30
25
20
15
10
5
% [125I]-HBsAg precipitated
wt
adw3.1
adw2.1
adw2.1.3
1·0
1·5
2·0
2·5
3·0
3·5
3·8
log10 concentration HBsAg ng/ml

The monoclonal antibody exhibited subtype specificity in the assay, reacting with the adw, but not the ayw, antigen (figure 5.1.4A). The unmutated antigen (wt) reacted with the d-specific monoclonal antibody to the same extent as a bona fide adw antigen. The three singly mutated HBsAg proteins, adw1, adw2, and adw3, also exhibited similar reactivity with the antibody. The multiply-mutated antigens (adw2.1, adw3.1, adw2.1.3) showed no significant deviation from the behaviour of the unmutated antigen to the d-specific monoclonal antibody.

These results indicate that HBsAg subtyped as a $\underline{d}^-$ shows the same relative affinity to a d-specific monoclonal antibody as that of a $\underline{d}^+$ antigen. Mutations which bring about a $\underline{d}^+$ subtype do not alter the reactivity of the antigen to the monoclonal antibody compared with that of $\underline{d}^-$ antigen.

5.2 Subtype-Specificity of Antisera Raised Against Mutant HBsAg

Mutant and wild type (ayw) HBsAg was used to immunise Balb/c mice (section 2.2.4). The relative anti-HBsAg titres of the antisera produced were determined by observing relative avidity with [^{125}I]-HBsAg in a DARIP assay (Burrell et al., 1978). Titres were expressed as the dilution capable of precipitating 50% of the maximum precipitable HBsAg (figure 5.2.1). The relative avidity of antisera for d and y determinants was indicated by the ability of the mouse sera to inhibit anti-d or anti-y:[^{125}I]-HBsAg (adw or ayw) reactions.

Figure 5.1.4 Effect of amino acid substitutions on the reactivity of HBsAg to a d-specific monoclonal antibody

The reactivity of test antigen with the monoclonal antibody was assessed by radioimmunoassay. The relative amount of HBsAg bound to monoclonal antibody coated beads (in a standard volume) was expressed as the ratio: amount of [^{125}I]-anti-HBsAg antibody bound to immobilised antigen (P) to the amount of [^{125}I]-anti-HBsAg antibody associated with the negative control (RIP buffer) (N). The standard concentration of HBsAg used was determined as in figure 5.1.2.

A. Effect of single adw mutations on reactivity.

B. Effect of multiple adw mutations on reactivity.

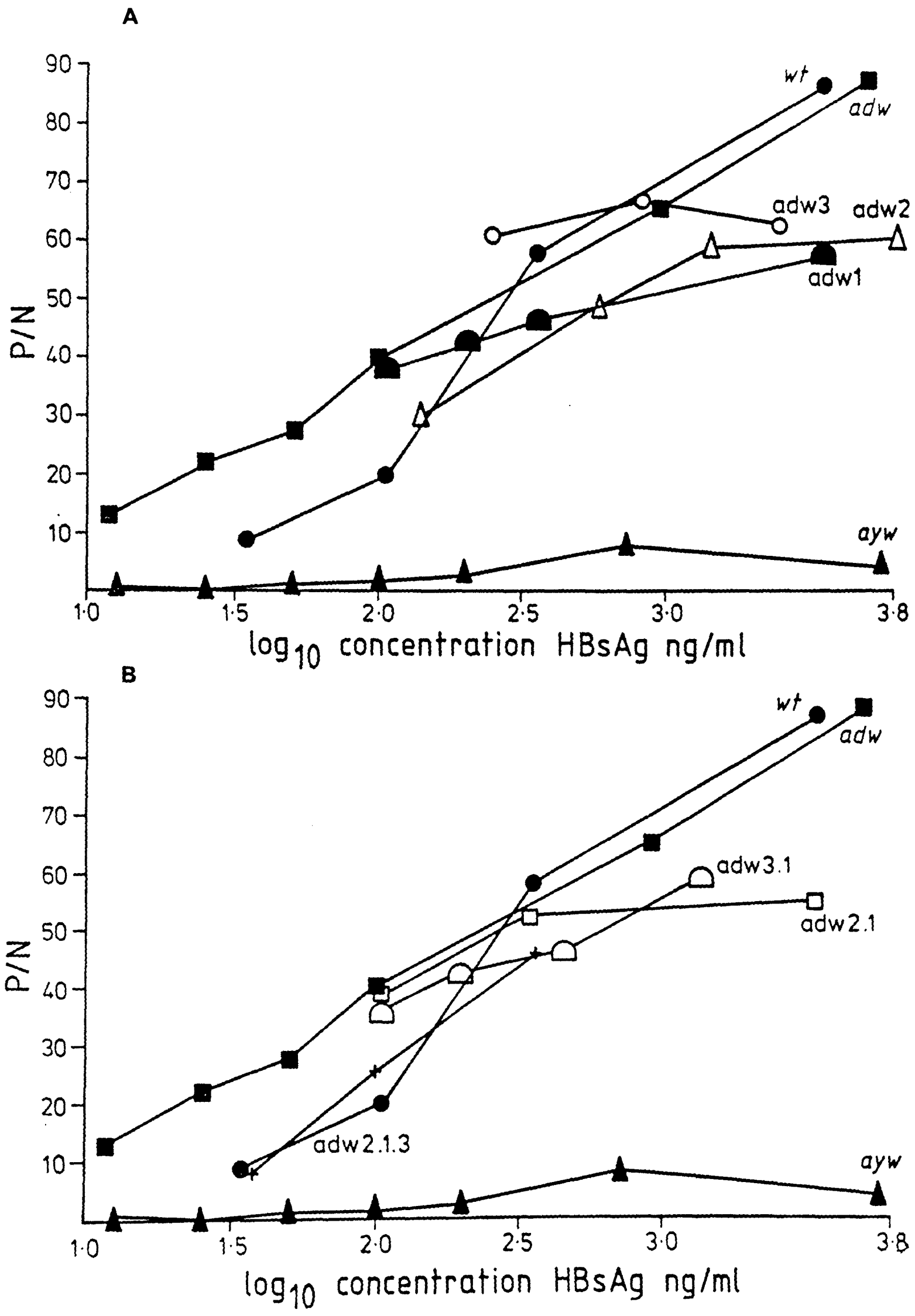
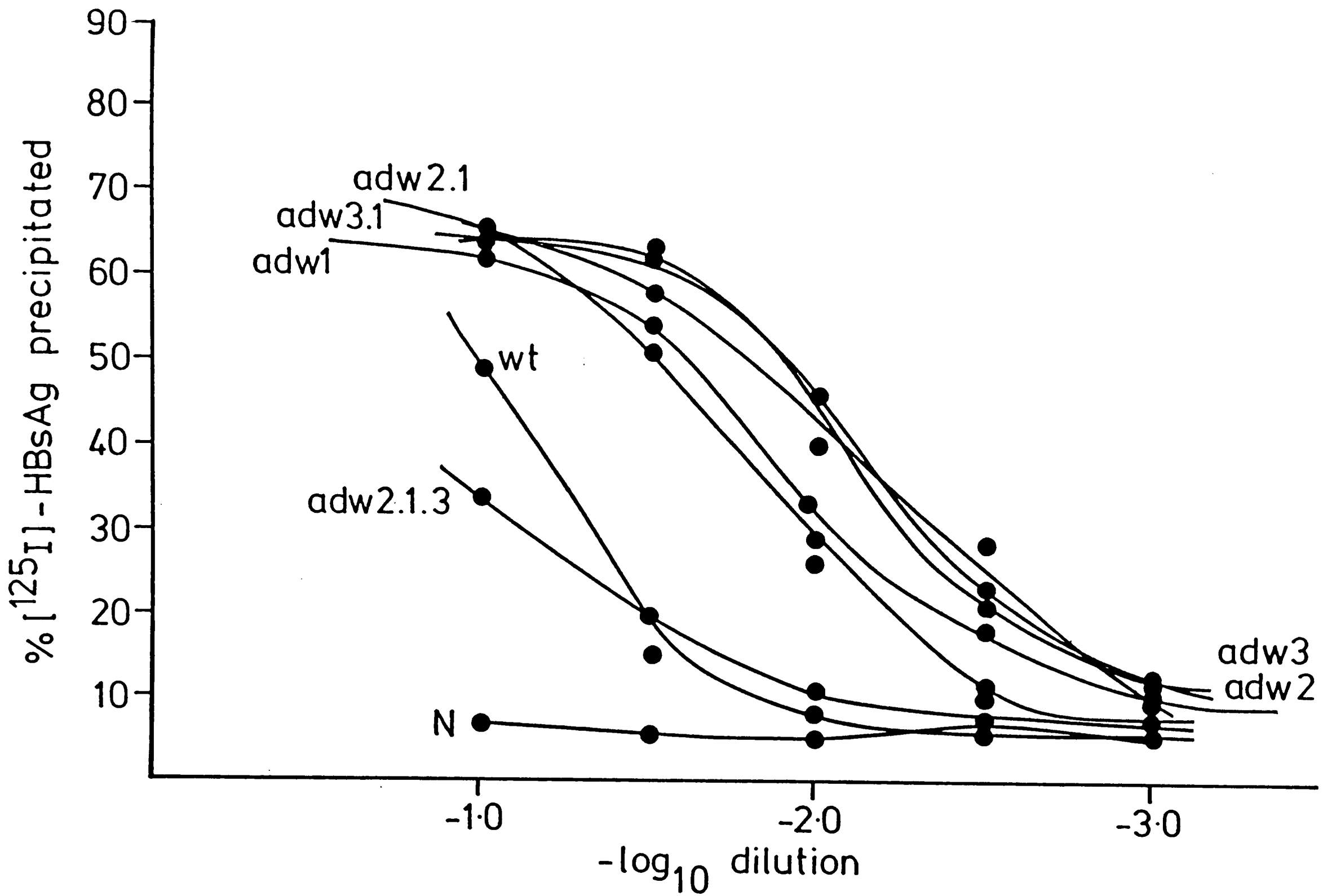

A
90
80
70
60
50
40
30
20
10
P/N
wt
adw
adw3
adw2
adw1
ayw
1·0
1·5
2·0
2·5
3·0
3·8
log10 concentration HBsAg ng/ml
B
90
80
70
60
50
40
30
20
10
P/N
wt
adw
adw3.1
adw2.1
adw2.1.3
ayw
1·0
1·5
2·0
2·5
3·0
3·8
log10 concentration HBsAg ng/ml

Figure 5.2.1 Titre of anti-yeast HBsAg sera from Balb/c mice

Animals were inoculated once and boosted twice with pHIN/G-2 derived HBsAg from mutant DBY746 strains (adw1, adw2, adw3, adw2.1, adw3.1, adw2.1.3) and unmutated cells (wt). Cells not transformed with pHIN/G-2 were also used to make extracts which were used to produce mouse antiserum (N). The titre of anti-HBsAg antisera was indicated by the degree of precipitation of [^{125}I]-HBsAg in a double antibody radioimmunoprecipitation assay, using sheep anti-mouse second antibody.

The titre of each antiserum was indicated by noting the dilution necessary to give 50% of the maximum level of precipitated antigen, which was 65% of total label. This precipitation level (32%) was used and the titres were: wt - 1/19, adw1 - 1/94, adw2 - 1/178, adw3 - 1/154, adw2.1 - 1/188, adw3.1 - 1/82, adw2.1.3 - 1/11. These titres were used to then normalise all future measurements taken on the antisera for the relative avidity of each anti-HBsAg serum.


90
80
70
60
50
40
30
20
10
% [125I]-HBsAg precipitated
adw2.1
adw3.1
adw1
wt
adw2.1.3
N
adw3
adw2
-1·0
-2·0
-3·0
-log10 dilution

Test antiserum (mouse) complexes with antigen were not precipitated by the donkey anti-rabbit IgG second antiserum (figure 5.2.2).

The antiserum raised against the unmutated HBsAg (wt) displayed the specificity anti-$\underline{y}^{+}$anti-$\underline{d}^{-}$ (figure 5.2.3). The effect of the mutations $113^{ser \rightarrow tyr}$, $122^{arg \rightarrow lys}$, $134^{tyr \rightarrow phe}$ on the subtype specificity of the antisera was to induce the production of anti-$\underline{d}$ antibodies and reduce the avidity of anti-$\underline{y}$ antibodies. Mutation at residue 122 (arg → lys) had greater effect than the mutation at 113 (ser → thr) and 134 (tyr → phe) on introducing $\underline{d}$-avidity and decreasing $\underline{y}$-avidity.

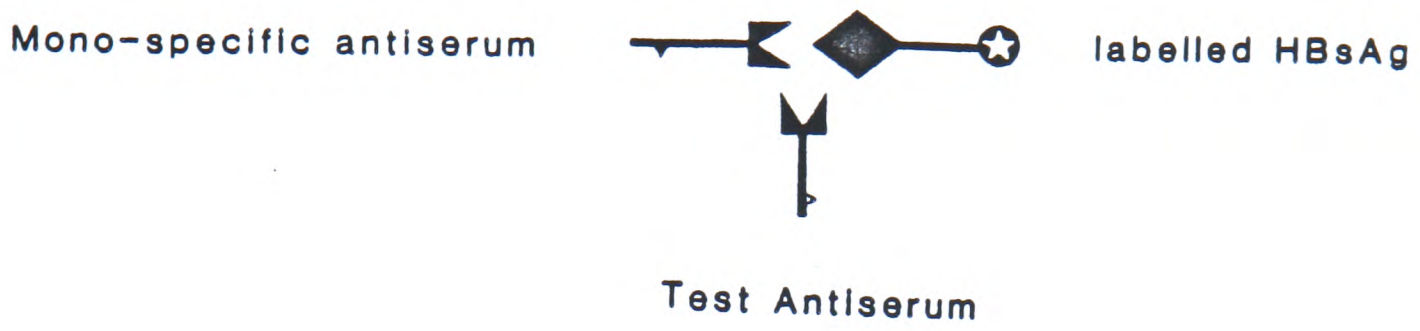
The antisera derived from the double (adw2.1, adw3.1) and triple mutants (adw2.1.3) revealed a greater relative avidity for $\underline{d}$-antigen than the antisera derived from unmutated or singly mutated HBsAg (figure 5.2.4). All three antisera exhibited a marked decrease in anti-$\underline{y}$ avidity, with anti-adw 2.1.3 serum possessing no detectable anti-$\underline{y}$ avidity. The substitutions at 113, 122, and 134, if present in the same protein, are capable of redirecting the production of subtypic antisera from $\underline{y}$ to $\underline{d}$. The amino acid position which had the greatest influence on $\underline{d}/\underline{y}$ subtype immunogenicity was 122 ($\underline{y}$ - arg, $\underline{d}$ - lys).

5.3 Mutations in the $\underline{a}$-Region and HBsAg Antigenicity

The mutations $124^{cys \rightarrow ser}$, $142^{pro \rightarrow ile}$, $142^{pro \rightarrow gly}$, $147^{cys \rightarrow ser}$ in HBsAg were made to test the contribution of each of the residues

Figure 5.2.2 Determination of the subtype specificity of anti-HBsAg sera

The relative avidity of mouse test-antiserum for the d or y-epitopes present on labelled HBsAg was indicated by the degree of inhibition the serum exerted on the monospecific antiserum labelled HBsAg reaction. The d/y-specific sera were produced from rabbit anti-HBsAg sera so only complexes between these antibodies and HBsAg will be precipitated by the second antibody (donkey anti-rabbit IgG). Labelled HBsAg:mouse test-antibody complexes will not be precipitated and so will diminish the degree of label recovered in the precipitate. The relative level of the decrease in precipitation indicated the avidity of the test antiserum for d or y features of HBsAg.


Mono-specific antiserum
labelled HBsAg
Test Antiserum

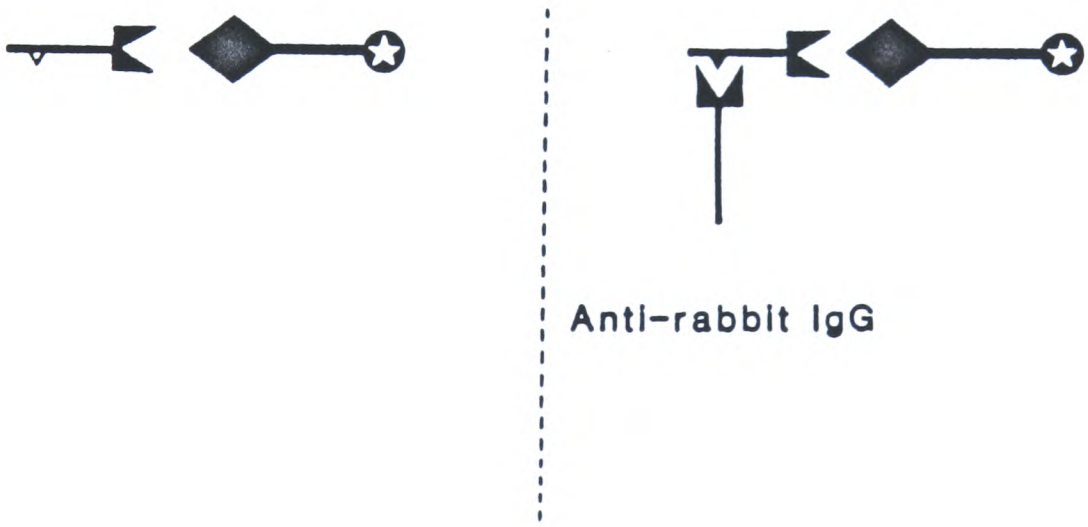
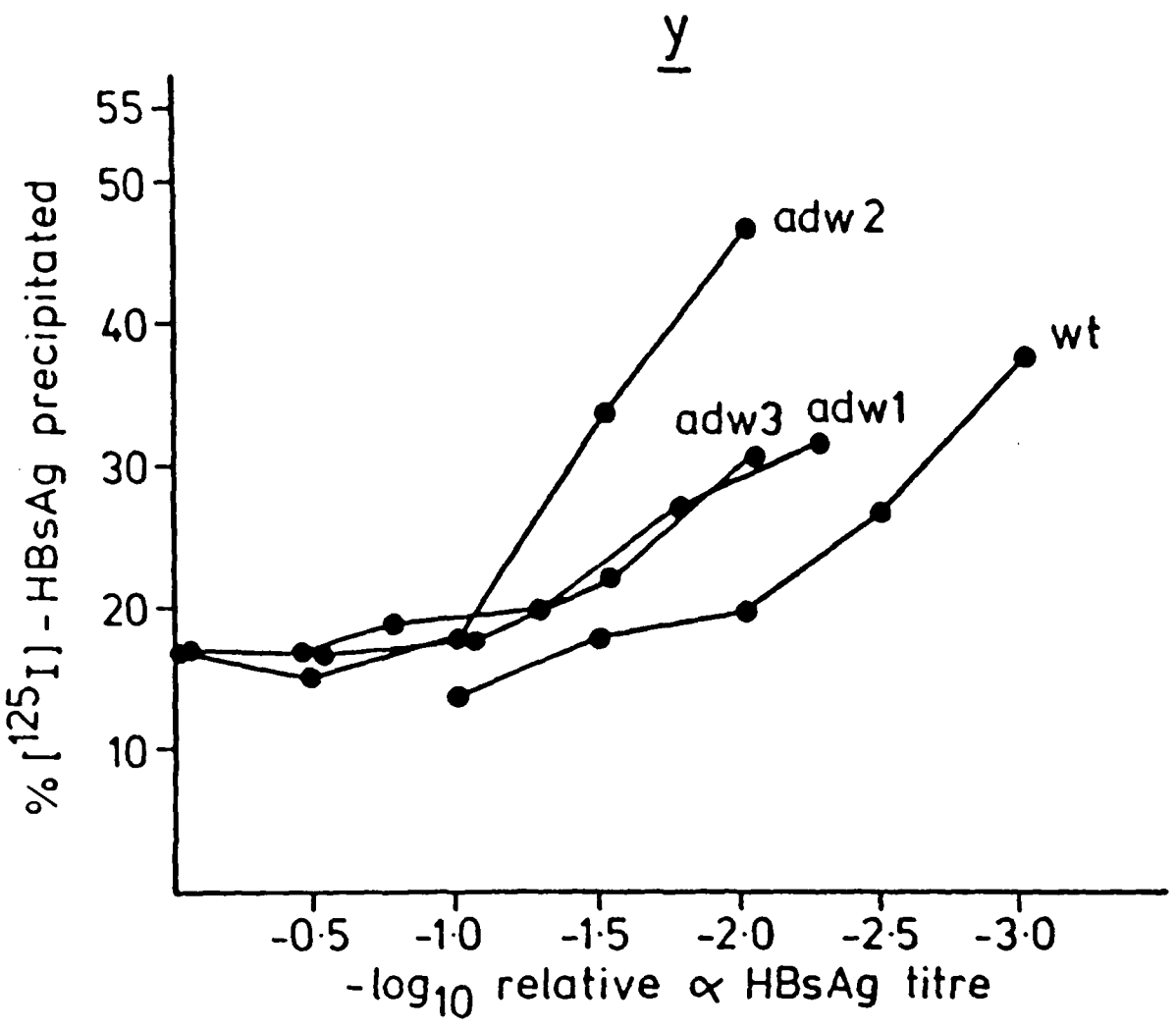

Anti-rabbit IgG

Figure 5.2.3 Effect of single amino acid substitutions in HBsAg on subtype specificity of anti-HBsAg antibodies

The assay for relative avidity of anti-HBsAg serum for d or y determinants was carried out using the assay described in figure 5.2.2. The negative control antiserum (N - raised against extracts from untransformed DBY746 cells) when tested in the assay gave a level of about 45% precipitated [^{125}I]-HBsAg over the relevant range of dilution. The reagents used in the assay (standard monospecific antisera, labelled antigen) were from batches previously tested for subtype specificity (figure 5.1.1B). The relative anti-HBsAg titre of antisera raised against mutated HBSAg was determined as described in figure 5.2.1.


y
55
50
40
30
20
10
% [125I]-HBsAg precipitated
adw2
adw3
adw1
wt
-0·5
-1·0
-1·5
-2·0
-2·5
-3·0
-log10 relative ∝ HBsAg titre

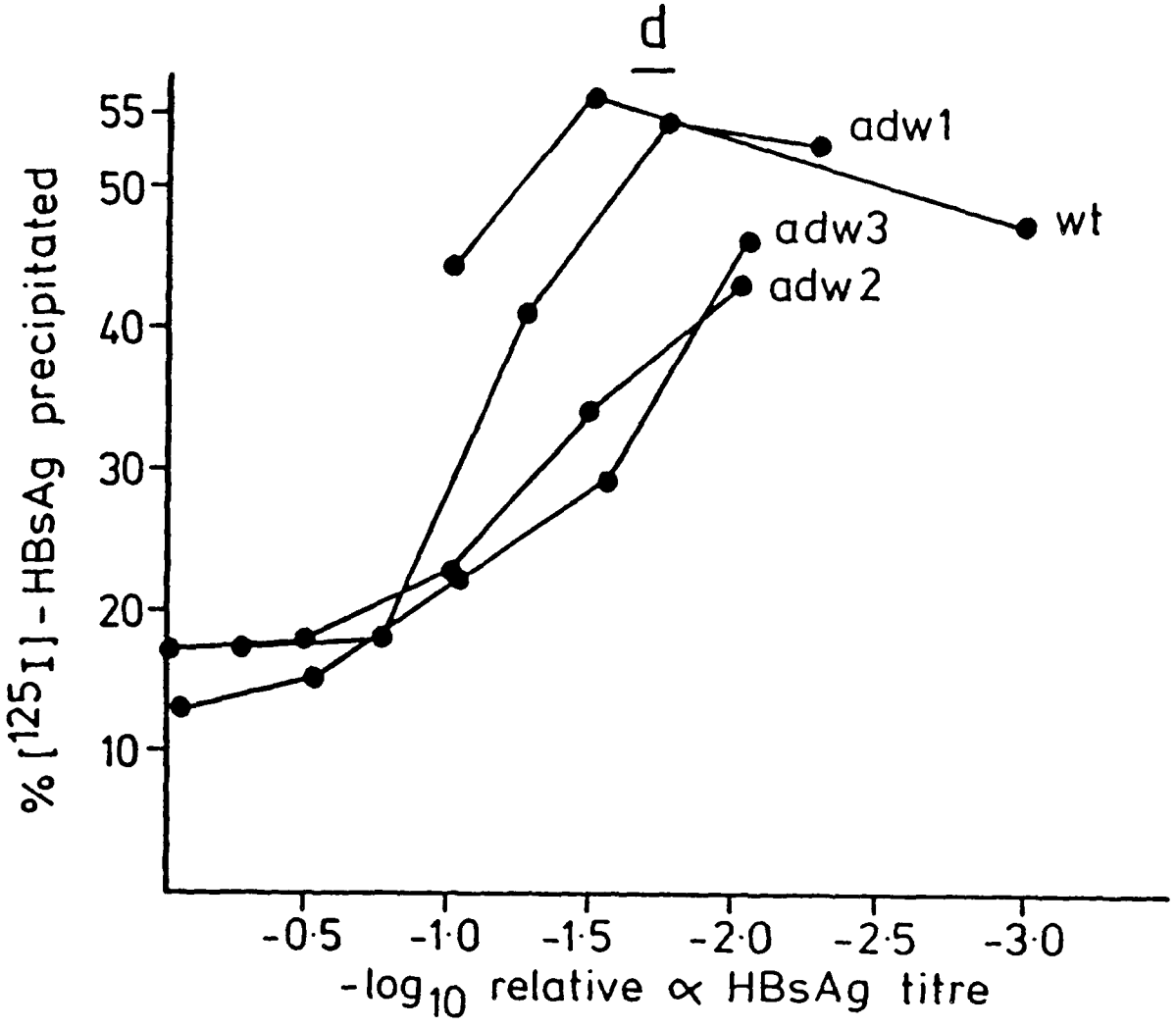
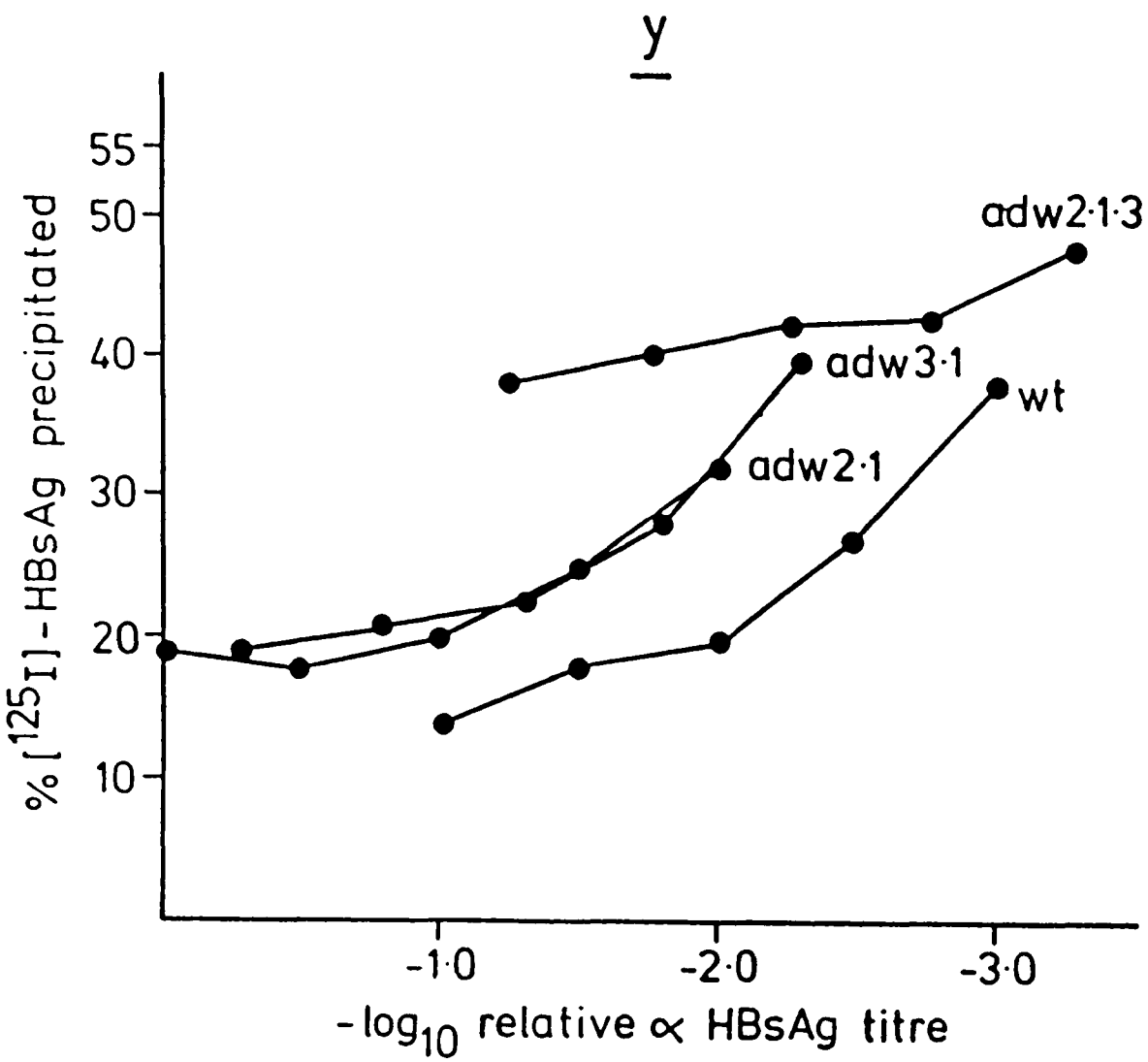
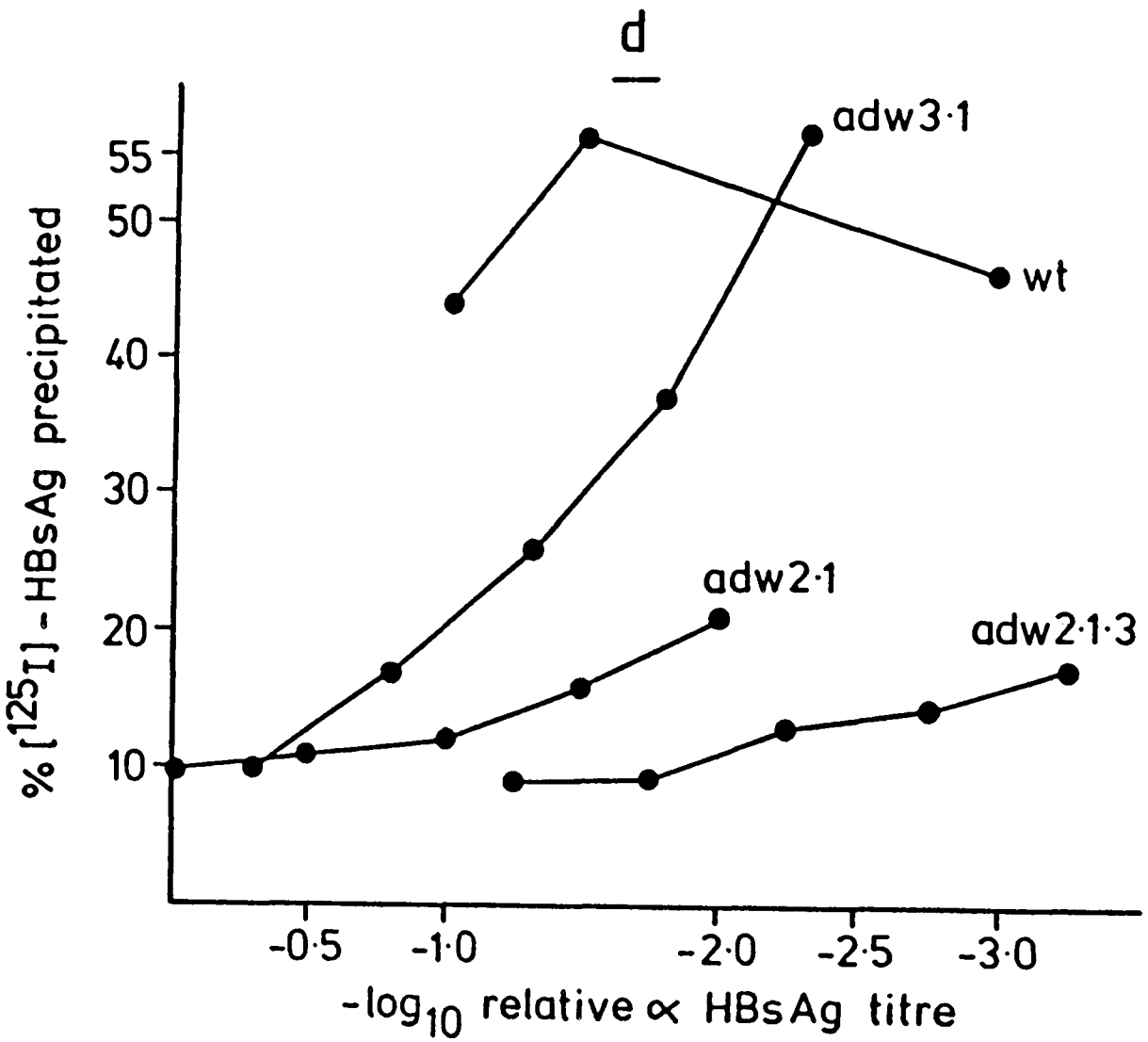

d
55
50
40
30
20
10
% [125I]-HBsAg precipitated
adw1
wt
adw3
adw2
-0·5
-1·0
-1·5
-2·0
-2·5
-3·0
-log10 relative ∝ HBsAg titre

Figure 5.2.4 Effect of multiple amino acid substitutions in HBsAg on subtype specificity of anti-HBsAg antibodies

The assay, controls and relative anti-HBsAg titre of antisera were the same as described in figure 5.2.3.


y
55
50
40
30
20
10
% [125I] - HBsAg precipitated
adw2·1·3
adw3·1
wt
adw2·1
-1·0
-2·0
-3·0
-log10 relative ∝ HBsAg titre


d
55
50
40
30
20
10
% [125I] - HBsAg
adw3·1
wt
adw2·1
adw2·1·3
-0·5
-1·0
-2·0
-2·5
-3·0
-log10 relative ∝ HBsAg titre

in determining HBsAg antigenicity. The concentration of HBsAg present in preparations of DBY746 cell extracts was determined for the denatured form of the protein (section 4.2, figure 4.2.1). This estimation therefore gave an indication of the concentration of HBsAg irrespective of any antigenic property of the native protein, and allowed the determination of the effect of mutations in the a-region on a-antigenicity of HBsAg at standard concentrations.

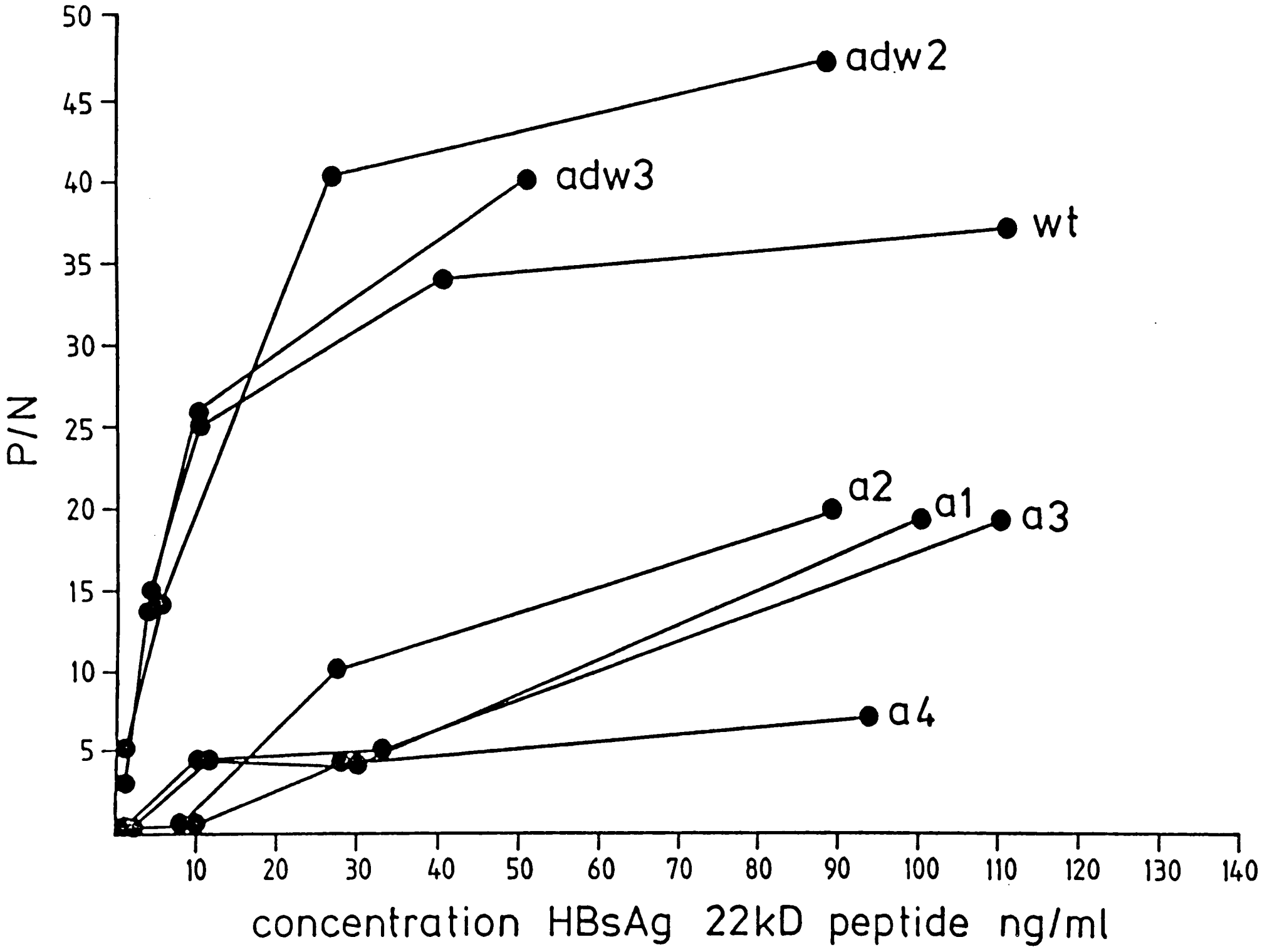
5.3.1 Effect on a-antigenicity of mutations in the a-region

Since the majority of antibodies present in an anti-HBsAg serum are directed against a-specific epitopes (McAuliffe et al., 1980), such an antiserum was used to estimate the relative antigenicity of a-regions in mutant and wild type HBsAg proteins. The Abbott Ausria II radioimmunoassay detection system for HBsAg was therefore used for this purpose.

Both amino acid substitutions at positions 142 (pro → ile, pro → gly) had a significant effect on a-antigenicity, reducing it compared with that of the unmutated antigen (figure 5.3.1). The isoleucine substitution had a greater effect than the glycine substitution. The cys → ser changes at positions 124 and 147 diminished antigenicity, with the 147 substitution all but abolishing a-antigenicity completely. The two subtype specific alterations ($122^{arg \rightarrow lys}$, $134^{tyr \rightarrow phe}$) although situated in the vicinity of the a-specific changes, did not significantly alter a-antigenicity, therefore

Figure 5.3.1 Effect of mutations on a-antigenicity

The Abbott Ausria II radioimmunoassay for the detection of human HBsAg was employed to assess the a-antigenicity of wild type (wt) HBsAg and mutant HBsAg - $124^{cys \rightarrow ser}$ (a3), $142^{pro \rightarrow gly}$ (a1), $142^{pro \rightarrow ile}$ (a2), $147^{cys \rightarrow ser}$ (a4), $122^{arg \rightarrow lys}$ (adw2), $134^{tyr \rightarrow phe}$ (adw3). The relative amount of HBsAg bound to polystyrene beads coated with anti-human HBsAg antibody (in a standard volume) was expressed as the ratio: amount of [^{125}I]-anti-HBsAg bound to immobilised antigen (P) to the amount of [^{125}I]-anti-HBsAg associated with the negative control (RIP buffer) (N). The concentration of HBsAg was estimated by immunoblotting (figure 4.2.1).


P/N
50
45
40
35
30
25
20
15
10
5
adw2
adw3
wt
a2
a1
a3
a4
10
20
30
40
50
60
70
80
90
100
110
120
130
140
concentration HBsAg 22kD peptide ng/ml

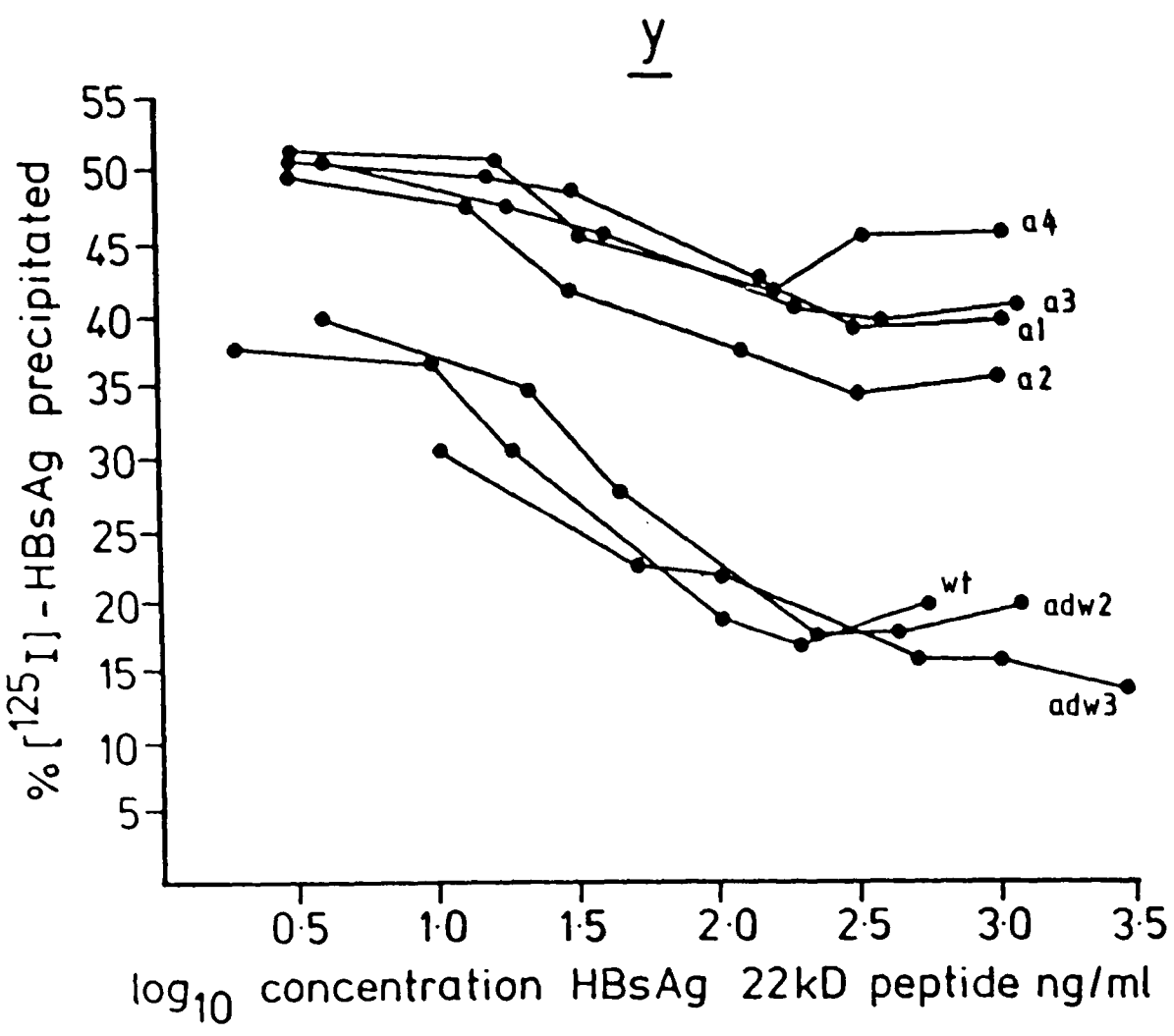
discounting the effects of non-specific amino acid changes on a-antigenicity.

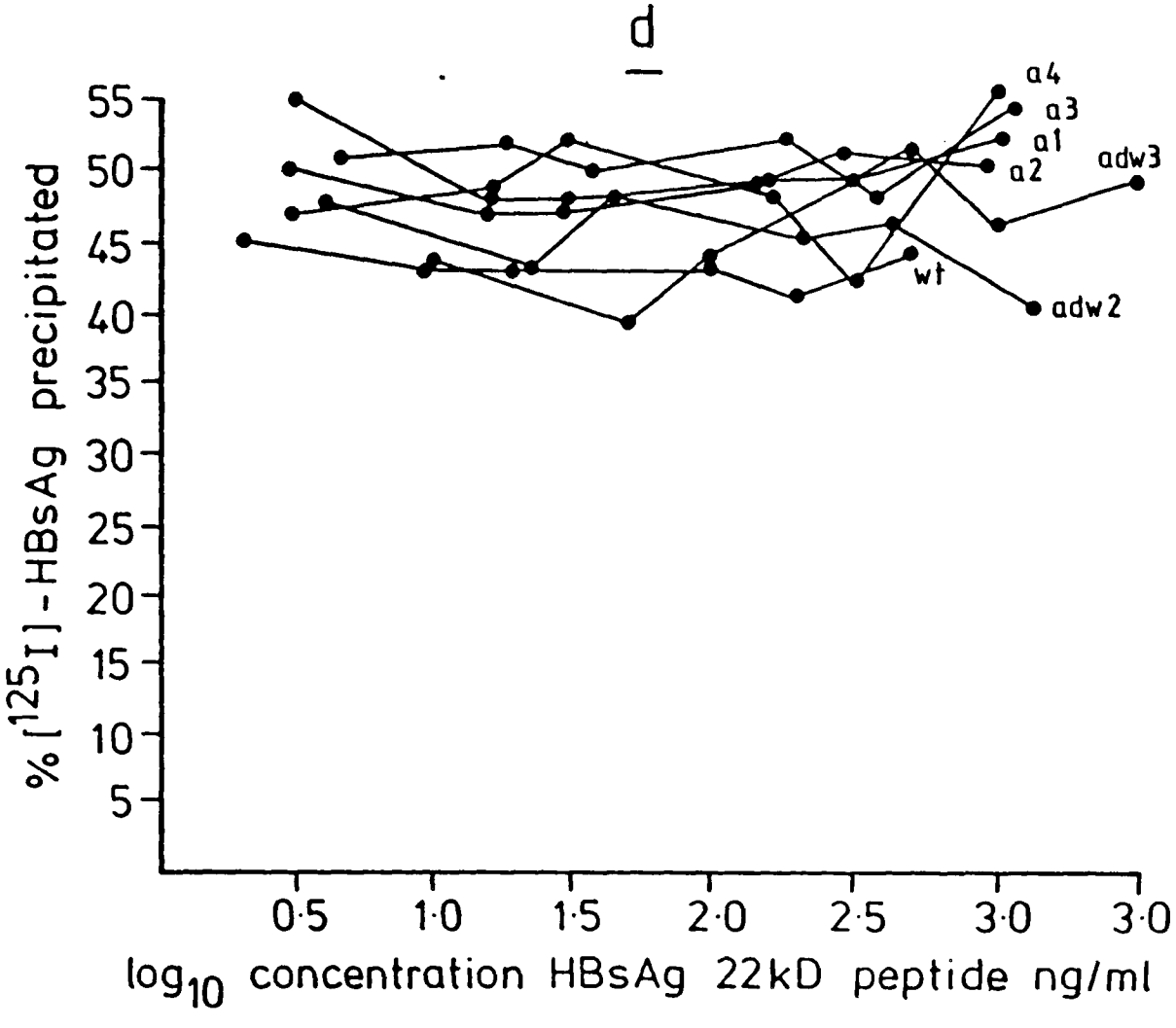
5.3.2 Influence of the a-region on d/y subtype

To investigate whether or not a fully antigenic a-region is necessary for the presence of antigenic d/y regions, the mutants deficient in a-antigenicity were subjected to subtype analysis. The assay used was that described in section 5.1.1 (figure 5.1.1), being based on a competition double antibody-radioimmunoprecipitation procedure. All four mutants exhibiting impaired a-antigenicity (a1 - a4) showed substantial decreases in y-antigenicity compared with that of the unmutated antigen (figure 5.3.2). There was no change in d-antigenicity from the negligible level observed for unmutated (wt) antigen. As in the case of a-antigenicity, the mutation $147^{cys \rightarrow ser}$ had the greatest effect on y-antigenicity, completely abolishing any inhibition by the antigen of the anti-y:[^{125}I]-HBsAg (ayw) reaction. The mutations $122^{arg \rightarrow lys}$ and $134^{tyr \rightarrow phe}$ which had no effect on a-antigenicity did not alter y or d-antigenicity. From these results it is clear that perturbation of a-region antigenicity leads to changes in the antigenicity of adjacent but immunologically distinct peptide regions.

Figure 5.3.2 Effect of amino acid substitutions in the a-region on the d/y subtype of HBsAg

The subtype of HBsAg was defined using the experiment described in figure 5.1.1. The concentration of antigen was estimated by immunoblotting (figure 4.2.1).


y
% [125I] - HBsAg precipitated
55
50
45
40
35
30
25
20
15
10
5
a4
a3
a1
a2
wt
adw2
adw3
0·5
1·0
1·5
2·0
2·5
3·0
3·5
log10 concentration HBsAg 22kD peptide ng/ml

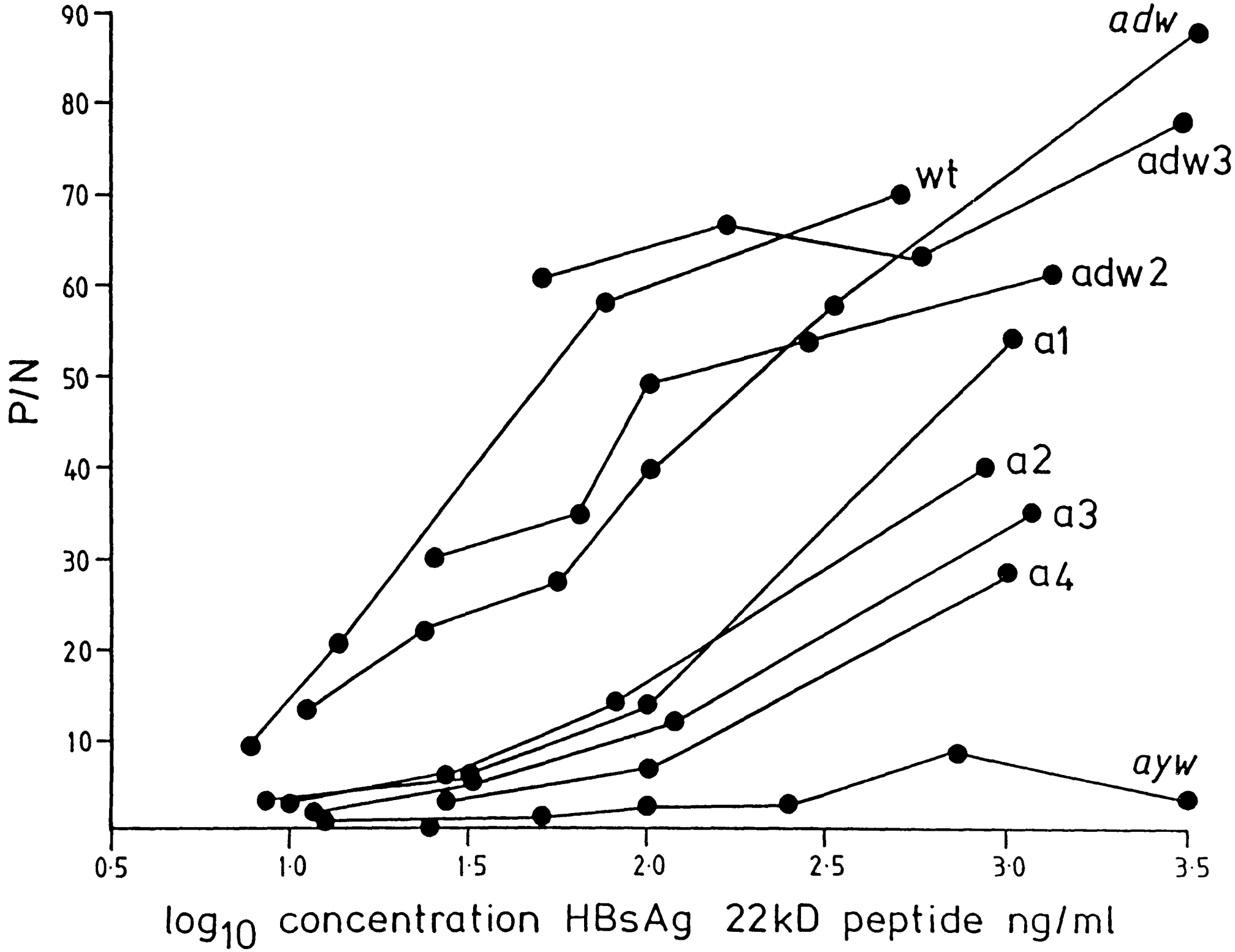

d
% [125I] - HBsAg precipitated
55
50
45
40
35
30
25
20
15
10
5
a4
a3
a1
a2
adw3
wt
adw2
0·5
1·0
1·5
2·0
2·5
3·0
3·0
log10 concentration HBsAg 22kD peptide ng/ml

5.3.3 Effect of mutations in the a-region on the binding of a d-specific monoclonal antibody

The radioimmunoassay described in section 5.1.3, was used to compare the relative affinities of HBsAg proteins mutated at positions 124, 142 and 147 and the wild type protein with the d-specific monoclonal antibody 1E6. All of the a-specific mutations identified in section 5.3.1 reduced the affinity of HBsAg for the d-specific monoclonal antibody (figure 5.3.3). This further indicates the importance of the immunodominant a-region in determining the antigenicity of adjacent peptide regions which harbour immunologically separate epitopes.

Figure 5.3.3 Effect of amino acid substitutions in the a-region on HBsAg reactivity to a d-specific monoclonal antibody

The reactivity of HBsAg to the monoclonal antibody was assessed by radioimmunoassay. The relative amount of HBsAg bound to polystyrene beads coated with monoclonal antibody (in a standard volume) was expressed as the ratio: amount of [^{125}I]-anti-HBsAg antibody bound to immobilised antigen (P) to the amount of [^{125}I]-anti-HBsAg antibody associated with the negative control (RIP buffer) (N). The concentration of HBsAg was estimated by immunoblotting (figure 4.2.1)


P/N
90
80
70
60
50
40
30
20
10
0·5
1·0
1·5
2·0
2·5
3·0
3·5
adw
wt
adw3
adw2
a1
a2
a3
a4
ayw
log10 concentration HBsAg 22kD peptide ng/ml

6. DISCUSSION

6. Discussion

This thesis has been concerned with an investigation into the factors responsible for determining the antigenicity and immunogenicity of a viral antigen. Since the serology of HBsAg is well defined and the location of antigenically important regions known, this protein provided an ideal system to study the amino acid residues involved in specific cross-reactions. The technique of oligonucleotide-mediated site-directed mutagenesis has been used to introduce specific nucleotide substitutions in the S gene of HBV. The gene was expressed in S. cerevisiae and gave rise to HBsAg harbouring predetermined mutations in the antigenic region.

To investigate the molecular basis of d/y subtype specificity, the effects of mutations at 113 (ser → thr), 122 (arg → lys) and 134 (tyr → phe) on antigenic and immunogenic subtypic activity were assessed. The contributions of amino acids proline 142, cysteine 124 and cysteine 147 in determining the a-antigenicity of HBsAg were also investigated, using a similar mutational approach. The influence of the a-region on the antigenicity of adjacent peptide regions was also assessed.

6.1 Mutational Analysis of the d/y Subtype of HBsAg

A summary of the immunological properties of d/y mutant-HBsAg can be found in table 6.1.

Table 6.1 Immunological Properties of HBsAg and its Mutants

Antigen or mutant	Cross-reactivity with Monospecific Antibody. Concentration of HBsAg giving 50% maximum inhibition of subtype specific antigen-antibody interaction		Cross-reactivity of Induced Antibody. Serum dilution at 50% maximum subtype specific antigen capture	
	Subtype of [^{125}I]-HBsAg		Subtype of [^{125}I]-HBsAg	
	y	d	y	d
Wild type	210ng.ml^{-1}	no inhibition	1:400	no reaction
y - 113	215	no inhibition	1:125	1:11
y - 122	230	no inhibition	1:22	1:22
y - 134	220	no inhibition	1:100	1:40
y - 133,122	235	230ng.ml^{-1}	1:32	1:562
y - 113,134	220	270	1:32	1:40
y - 113,122,134	260	250	no reaction	>1:1000

All the results are interpolated values from titration experiments in which the antigen or antibody under test competed with its counterpart in double antibody radioimmunoprecipitation (DARIP) assays using [^{125}I]-HBsAg of y or d subtype and monospecific antisera of the same subtype.

6.1.1 Distinction between the y and d antigenic subtypes is not attributable to a single amino acid difference

An approach that has been used to define the molecular basis of HBsAg subtype specificity involves the identification of amino acid positions which correlate with the presence of a specific subtype. The result of this comparative method implicated amino acid positions 68 and 122 in defining the d/y subtype (figure 1.4.2). Several groups of workers equipped with this information have focussed their attention on these residue positions. It has been established that differences in amino acid at position 122 determines the binding of d (lysine) and y (arginine) specific monoclonal antibodies (Peterson et al., 1984; Okamoto et al., 1986, 1987a). No group has investigated the contribution of residues at 68 on d/y subtype presumably since there is no evidence implicating the region around 68 in the participation of antigenic reactions (Prince et al., 1982; Bhatnager et al., 1982; Lerner et al., 1981; Gerin et al., 1983).

However, it was found that the change at 122 from the y specific to the d specific amino acid did not alter subtype, as defined by the avidity of the antigen for monospecific, polyclonal anti-d or anti-y sera (figure 5.1.2). To account for the discrepancies between the two results, it is suggested that the anti-d and anti-y monoclonal antibodies which have been used previously to define subtype, do not faithfully represent the whole population of subtype specific antibodies present in anti-d and anti-y sera.

6.1.2 The reaction of HBsAg with a subtype-specific monoclonal antibody does not define subtype

Unmutated HBsAg produced from pHIN/G-2 had a $\underline{y}^{+}\underline{d}^{-}$ subtype (figure 5.1.2). However, the antigen reacted with a $\underline{d}$-specific monoclonal antibody (1E6) to some extent as did $\underline{adw}$ HBsAg (figure 5.1.4A). This result highlights the fact that the designation of HBsAg depends on the interaction of several epitopes with several paratopes (Swenson et al., 1988). Therefore, the $\underline{d}$-specific monoclonal antibody used in these experiments did not accurately reflect the subtype of the antigen presumably because in a $\underline{d}$ specific polyclonal antiserum the antibody in question is present in low amounts or binds with low affinity. Furthermore, mutations in HBsAg that changed the subtype from $\underline{d}^{-}\underline{y}^{+}$ to $\underline{d}^{+}\underline{y}^{+}$ (adw2.1, adw3.1, adw2.1.3) did not increase the affinity of the $\underline{d}$-specific monoclonal antibody for the epitope present in HBsAg (figure 5.1.4B).

6.1.3 Substitution of three amino acids allows the partial switch from $\underline{y}$ to $\underline{d}$ antigenic subtype

The mutations $122^{arg \rightarrow lys}$ and $134^{tyr \rightarrow phe}$ when combined with $113^{ser \rightarrow thr}$ produced a dramatic increase in the avidity of HBsAg for monospecific $\underline{d}$-antiserum (figure 5.1.3). When all three mutations were present in HBsAg the avidity for monospecific $\underline{y}$-antiserum was noticeably decreased, with the avidity for the $\underline{d}$-antiserum similar to that of the double mutant. The $\underline{d/y}$ subtype is dependent on the presence or absence of several epitopes (Swenson et al., 1988), thus

amino acid substitutions that switch subtype must alter the structure of the region such that a set of epitopes are abolished and another set created. Therefore, both double mutants must have altered the structure of the antigen in such a way as to establish the presence of d-epitopes whilst preserving the existing y-epitopes. The three mutations present in adw2.1.3 facilitated the adoption of a d-specific structure, whilst at the same time diminished the y-nature of the area. This conformational basis for HBsAg subtype differences is consistent with the general features of the antigen.

Synthetic peptides corresponding to residues 110 - 137 of HBsAg as predicted from the S gene nucleotide sequences of HBV DNA from ayw and adw were tested for their ability to bind anti-ayw and anti-adw sera (Gerin et al., 1983). It was found that the peptides exhibited subtype specificity in reactions with the sera, thus implicating the positions of amino acid variability (110, 114, 122, 125, 127, 131, 134) in determining subtype specificity. Of the seven positions implicated by this study, substitutions at two (122, 134) have been shown sufficient to elicit the change $y^{+}d^{-}$ to $y^{+}d^{+}$ in the antigenicity of HBsAg when combined with the mutation $113^{ser \rightarrow thr}$. The contribution of 113 in d/y subtype was unexpected; the mutation at the position was made to serve as a negative control to the substitutions at 122 and 134, since variability at 113 is in accordance with w/r specificity. The most likely explanation is that the mutation $113^{ser \rightarrow thr}$ was sufficiently similar to the y → d specific change $114^{ser \rightarrow thr}$ (Gerin et al., 1983) to allow it to influence d/y antigenic behaviour.

6.1.4 Amino acid residues at three positions define the immunogenic d/y subtype of HBsAg

The single amino acid substitutions that had little effect on the d/y subtype of HBsAg resulted in substantial changes in the subtype specificity of anti-HBsAg sera (figure 5.2.3). The contribution of residues at 114, 122 and 134 in determining the immunogenic as well as the antigenic subtype of HBsAg has been reported (Gerin et al., 1983) in the case of peptides representing a region 110 - 139 of HBsAg.

The mutation at 122 had a more pronounced effect than the 134 or 113 mutations, when singly present or with 113 the mutation at (figure 5.2.4). This result would have been predicted from studies correlating the position of variable amino acid residues with subtype (figure 1.4.2), since residues at 122 correlate exactly with the appearance of the d/y subtype. The residues present at 122 (y - arg, d - lys) are likely to be immunogenic to B-cells and located at the surface of the antigen since they possess positively charged, hydrophilic side groups (Hopp and Woods, 1981). Residues at 122 would therefore participate to a large degree in determining the antigenic and immunogenic behaviour of HBsAg.

The three mutations present in adw2.1.3 changed the specificity of anti-HBsAg sera from y to d. The possibility that other areas of HBsAg contribute to the determination of the subtype specificity

still remains, however this study has identified the minimum structural changes necessary for the specificity switch.

The changes observed in the subtype of anti-HBsAg sera due to substitutions of amino acids at positions 113, 122 and 134 were more pronounced than changes observed in the subtype antigenicity of HBsAg protein. Furthermore the dominant effect of position 122 on subtype reactivity was best detected by the observation of changes in HBsAg immunogenicity rather than antigenicity. This greater influence of mutations on HBsAg immunoreactivity can be best explained by assuming that the in vivo detection of alterations in immunogen structure, afforded by the immune system of Balb/c mice, exhibited greater sensitivity than the in vitro reaction used to detect changes in antigenicity.

The contribution of T-cell specific epitopes to the immune response against HBsAg is well established (Roberts et al., 1975). A region (24 - 27) of HBsAg has been shown to be capable of inducing T-helper cell proliferation (Celis et al., 1988), however no amino acid variation occurs in this region, so it is unlikely that it contributes to immunogenic subtype differences. Like other viruses, HBV triggers the cell-mediated and humoral immune systems, a consequence being the priming of anti-HBsAg antibody production by activated T-cells specific for other viral immunogens (Vento et al., 1987; Milich et al., 1987a, b, c, 1988). Such an arrangement could determine the subtype specificity of anti-HBsAg sera through

variation in T-cell epitopes located in virion components other than HBsAg.

6.2 The a-Region of HBsAg

6.2.1 Mutations and a-antigenicity

The effects of the amino acid substitutions at positions 124 (cys → ser), 142 (pro → gly, pro → ile) and 147 (cys → ser) on HBsAg a-antigenicity are summarised in table 6.2. The residues at these positions, cysteine or proline, play an important role in the determination of polypeptide conformation in proteins. The cysteine residues will in all likelihood be involved in stabilisation through disulphide bridges, and the proline would also be conformationally significant, dictating a turn in the polypeptide chain. Substitution of the proline with isoleucine decreases antigenicity more than substitution with glycine (figure 5.3.1), which is consistent with the above postulate; the bulky hydrophobic isoleucine side chain may impede the turn more than the glycine which allows relatively free movement of the peptide (Garnier et al., 1978).

It has been shown that the cysteines at positions 124 and 147 are both required for full a-antigenicity, but that at 147 is virtually indispensible, suggesting that this residue participates in a disulphide bridge that is essential for maintenance of the correct arrangement of the antigen. The role of cysteines in maintaining HBsAg conformation and hence antigenicity is well established

Table 6.2 "Abbott Ausria II" RIA for human-HBsAg. P/N value with the 22kD polypeptide (50ng/ml)

Cross-reactivity with monospecific antibody Concentration of ($ng.ml^{-1}$) HBsAg (22kD polypeptide giving 50% maximum inhibition of subtype specific antigen-antibody interaction

Antigen or mutant	22kD-peptide	y	d
Wild type	34	16	no reaction
a - 124	8	no reaction	"
a - 142 gly	14	"	"
a - 142 ile	8	"	"
a - 147	5	"	"
y - 122	43	32	"
y - 134	40	10	"

(Dreesman et al., 1987; Sukeno et al., 1972; Vyas et al., 1972). Studies with synthetic peptides identified the regions between residues 122 and 137 (Dreesman et al., 1982) and residues 139 and 147 (Brown et al., 1984) as contributors to a-antigenicity and shows further that cyclisation through cysteine residues 124 and 137 or 139 and 147, as indicated in figure 1.5.2, enhanced a-immunoreactivity. Together these experiments support the view that cystine bridges involving residues 124 and 147 with one of the cluster of cysteines at 137 and 139 have a direct role in determining the conformational integrity of the a-antigenic region. It is interesting to note in contrast that substitution of cysteine 121 to serine had little effect upon reactivity of the antigen with its antibody in the "Ausria" radioimmunoassay, whereas against a panel of five monoclonal antibodies specific for the a-region reactivity was greatly reduced or eliminated in three and enhanced in two (Antoni and Peterson, 1988).

The production of further HBsAg proteins which have undergone substitution of cysteines for serines in the a-region will allow the further investigation of the contribution of cystine bridges in maintaining HBsAg antigenicity.

Experiments with synthetic peptides have been used in studies of subtype specificity and have implicated amino acid residues at 122 and 134 in the determination of the d/y subtype specificity, but not a-antigenicity (Gerin et al., 1983). These residues therefore provided a reference against which to assess the effect of other

alterations in local peptide structure on a-antigenicity. That these substitutions at these positions did not affect a-antigenicity reinforces the view that the changes observed with the four mutants at positions 124, 142 and 147 were attributable to significant deformations of the native structure of HBsAg.

6.2.2 Influence of a-region on d/y subtype

Mutations that diminished a-antigenicity also affected the extent of d/y antigenicity (figure 5.3.2). Furthermore, the affinity of HBsAg for a d-specific monoclonal antibody was reduced (figure 5.3.3). These results imply that the antigenic integrity of the immunodominant a-region (residues 138 - 147) is necessary for the optimal antigenic expression of the subtype determining regions of the protein. Any loss of integrity of the a-region would be likely to result in a change of conformation of adjacent epitopes not normally recognised by a-specific antibodies. This would probably drastically affect the antigenicity of non-cross reacting epitopes such as those which are subtype specific.

6.3 Limitations of the Thesis

The effects of each mutation made in HBsAg were assessed at the level of immunoreactivity, such that changes in the primary structure of the protein were related to a change in a biological property. As stated throughout the thesis, the B-cell immunoreactivity of protein antigens is largely dictated by the spatial arrangement of

constitutive polypeptides. However, in the study described no means to assess alteration in antigen conformation were employed. Any change in immunoreactivity observed gave rise to the proposal of hypothetical changes in conformation. The indirect nature of the approach used in the study described is shared by other methods employed to investigate B-cell-specific antigenic structure, such as experiments involving the use of synthetic peptides.

The use of X-ray diffraction-crystallography to investigate protein structure allows a comprehensive analysis of epitope structure; such that any changes in immunoreactivity can be directly rationalised in terms of altered protein conformation.

Another limitation of the study, and of all similar studies, is the extent of the mutational analysis conducted. In the case of the investigation of a-antigenicity, the residues chosen for mutagenesis were four out of several other residues of similar apparent importance in maintaining conformation. For example, the extent of mutational analysis could be increased to include other proline and cysteine residues present in the vicinity. Such work is in progress, and the results are eagerly awaited.

REFERENCES

Aggerbeck LP, Peterson DL. (1985) Electron microscopic and solution X-ray scattering observations on the structure of hepatitis B surface antigen. Virology 141, 155-161.

Alexander J, Bey E, Geddes E, Lecatsas G. (1976) Establishment of a continuously growing cell line from primary carcinoma of the liver. S Afr Med J 50, 2124-2128.

Amit AG, Mariuzza RA, Phillips SEU, Poljak RJ. (1986) 3D structure of an antigen-antibody complex at 2-8Å resolution. Science 233, 747-753.

Antoni BA, Peterson DL. (1988) Site-directed mutagenesis of the hepatitis B surface antigen gene: Creation of a free sulphydyl group and modification of the protein in the 22nm particle structure. In: Viral Hepatitis and Liver Disease. (Ed AJ Zuckerman) Alan R Liss Inc, New York.

Babinet C, Farza H, Morello D, Hadchouel M, Pourcell CH. (1985) Specific expression of hepatitis B surface antigen (HBsAg) in transgenic mice. Science 230, 1160-1163.

Bancroft WH, Mundan FK, Russell PK. (1972). Detection of additional antigenic determinants of hepatitis B antigen. J Immunol 109, 842-848.

Barlow DJ, Edwards MS, Thornton JM. (1986) Continuous and discontinuous protein antigenic determinants. Nature 322, 747-748.

Beasley RP, Huiang LY. (1984) In: Viral Hepatitis and Liver Disease. (Eds Vyas GW, Grune A, and Orlando FL) Franklin Institute Press, Philadelphia, Pennsylvania.

Beggs JD. (1978) Transformation of yeast by a replicating hybrid plasmid. Nature 275, 104-109.

Benjamin DC, Berzofsky JA, East IJ, Gurd FRN, Hannum C, Leach SJ, Margoliash E, Michael JG, Miller A, Prager EM, Reichlin M, Sercarz EE, Smith-Gill SJ, Todd PE, Wilson AC. (1984) The antigenic structure of proteins: an appraisal. Ann Rev Immunol 2, 67-101.

Bhatanagar PK, Papas E, Blum HE, Milich DR, Nitecki D, Karels MJ, Yvas GN. (1982) Immune response to synthetic peptide analogues of HBsAg for the a determinant. Proc Natl Acad Sci USA 79, 4400-4404.

Bichko V, Pushko P, Dreilina D, Pumpen P, Gren E. (1985) Subtype ayw variant of HBV DNA primary structure analysis. FEBS Lett 185, 208-212.

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

Blum HE, Haase AT, Harris JD, Walker D, Vyas GN. (1984) Asymmetric replication of hepatitis B virus DNA in human liver: demonstration of cytoplasmic minus-strand DNA by blot analysis and in situ hybridization. Virology 139, 87-96.

Blum HE, Stowring L, Figus A, Montgomery C, Haase A, Vyas GN. (1983) Detection of hepatitis B virus DNA in hepatocytes, bile duct, epithelium, and vascular elements by in situ hybridization. Proc Natl Acad Sci USA 80, 6685-6688.

Blumberg BS, Alter HJ, Visnich S. (1965) A "new" antigen in leukemia sera. J Am Med Assoc 191, 541-546.

Blumberg BS, Gerstly BJS, Hungerford DA, London WT, Sutnick AI. (1967) A serum antigen (Australia antigen) in Down's syndrome, leukemia and hepatitis. Ann Int Med 66, 924-931.

Boyer HW, Roulland-Dissoix D. (1969) A complementation analysis of the restriction and modification of DNA in Escherichia coli. J Mol Biol 41, 459-472.

Brady J, Radonovich M, Vodkin M, Natorajan V, Thoren M, Das G, Janik J, Salzman NP. (1982) Site-directed base substitution and deletion mutations that enhance or supress transcription of the SV40 major late RNA. Cell 31, 625-633.

Brown SE, Howard CR, Zuckerman AJ, Steward MW. (1984) Determination of affinity of antibodies to HBsAg in human sera. J Immunol Meth 72, 41-48.

Burrell CJ. (1975) Host components in hepatitis B antigen. J Gen Virol 27, 117-126.

Burrell CJ, Leadbetter G, Black SH, Hunter WM. (1978) Rapid detection of hepatitis B surface antigen by double antibody radioimmunoassay. J Med Virol 3, 19-26.

Burrell CJ, Proudfoot E, Keen GA, Marmion PB. (1973) Carbohydrates in hepatitis B antigen. Nature New Biol 243, 260-262.

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

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

Celis, E, Ou D, Otuos L. (1988) Recognition of hepatitis B surface antigen by human T lymphocytes. J Immunol 140, 1808-1815.

Chakrabarty P, Ruiz-Opaza N, Shoual D, Shafritz, D. (1980) Identification of human hepatocellular carcinoma cell lines. Nature 286, 531-533.

Chang C, Enders G, Sprengel R, Peters N, Varmus HE, Ganem D. (1987) Expression of the precore region of an avian hepatitis B virus is not required for viral replication. J Virol 61, 3322-3325.

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

Chisari FV, Filippi P, McLachlan A, Milich DR, Riggs M, Lee S, Palmiter RD, Pinkert CA, Brinster RL. (1986) Expression of hepatitis B virus large envelope polypeptide inhibits hepatitis B surface antigen secretion in transgenic mice. J Virol 60, 880-887.

Chou PY, Fasman GD. (1978) Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzym 47, 45-148.

Christman J, Gerber M, Price P, FlordellsC, Edelman J, Acs G. (1982) Amplification of expression of hepatitis B surface antigen in 3T3 cells cotransfected with a dominant-acting gene and cloned viral DNA. Proc Natl Acad Sci USA 79, 1815-1819.

Cohen BJ, Richmond JE. (1982) Electron microscopy of hepatitis B core antigen synthesized in E. coli. Nature 296, 677-679.

Colman PM, Lower WG, Vanghese JN, Baker AT, Tulloch PA, Air GM, Webster PG. (1987) 3D structure of a complex of antibody with influenza virus neuramidase. Nature 326, 358-363.

Cossart YE. (1971) Australia antigen and hepatitis: a review. J Clin Pathol 24, 394-403.

Coulson A, Collins J, Lyall A. (1987) Protein and nucleic acid sequence database searching: a suitable case for parallel processing. The Computer Journal 30, 420-424.

Dane DS, Cameron CH, Briggs M. (1970) Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. Lancet i, 695-698.

DeJean A, Bouguelenet L, Grzeschile K, Tiollais P. (1986) Hepatitis B virus DNA integration in a sequence homologous to v-erb-A and steroid receptor genes in a hepatocellular carcinoma. Nature 322, 70-72.

DeJean A, Sonigo P, Wain-Hobson S, Tiollais P. (1984) Specific hepatitis B virus integration in hepatocellular carcinoma DNA through a viral 11-base pair direct repeat. Proc Natl Acad Sci USA 81, 5350-5384.

DeLisi C, Berzofsky JA. (1985) T-cell antigenic sites tend to be amphipathic structures. Proc Natl Acad Sci USA 82, 7048-7052.

DeThé H, Marchio A, Tiollais P, DeJean A. (1987) A novel steroid thyroid hormone receptor-related gene inappropriately expressed in human hepatocellular carcinoma. Nature 330, 667-670.

Dreesman GR, Hollinger RB, MacCombs RM, Melnick JL. (1973) Alteration of HBsAg determinants by reduction and alkylation. J Gen Virol 19, 129-134.

Dreesman GR, Sanchez Y, Ionescu-Matiu I, Sparrow JT, Six HR, Peterson DL, Hollinger BL, Melnick JL. (1982) Antibody to HBsAg after a single inoculation of uncoupled synthetic HBsAg peptide. Nature 295, 158-160.

Dubois M-F, Pourcell CH, Rousset S, Chang C, Tiollais P. (1980) Excretion of hepatitis B surface antigen particles from mouse cells transformed with cloned viral DNA. Proc Natl Acad Sci USA 77, 4549-4553.

Eble BE, Lingappa VR, Ganem D. (1986) Hepatitis B surface antigen: an unusual secreted protein initially synthesized as a transmembrane polypeptide. Mol Cell Biol 6, 1454-1463.

Eble BE, Macrae D, Lingappa VR, Ganem D. (1987) Multiple topogenic sequences determine the transmembrane orientation of HBsAg. Mol Cell Biol 7, 3591-3601.

Edman JC, Gray P, Valenzuela P, Rall LB, Rutter WJ. (1980) Integration of hepatitis B virus sequences and their expression in a human hepatoma cell. Nature 286, 535-538.

Emini E, Hughes J, Perlow D, Boger J. (1985) Induction of hepatitis A virus-neutralizing antibody by a virus specific synthetic peptide. J Virol 55, 836-839.

Enders GH, Ganem D, Varmus HE. (1985) Mapping the major transcripts of ground squirrel hepatitis virus. The presumptive template for reverse transcriptase is terminally redundant. Cell 42, 297-308.

Farza H, Salmon AM, Hadchouel M, Moreau JL, Babinet C, Tiollais P, Pourcell CH. (1987) Hepatitis B surface antigen gene expression is regulated by sex steroids and glucocorticoids in transgenic mice. Proc Natl Acad Sci USA 84, 1187-1191.

Feinberg AP, Vogelstein BC. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Analyt Biochem 132, 6-13.

Feinberg AP, Vogelstein BC. (1984) Addendum: A technique for radiolabelling DNA restriction endonuclease fragments to high specificity activity. Analyt Biochem 137, 266-267.

Francis MJ, Hastings LZ, Syred AD, McGinn B, Brown F, Rowlands DJ. (1987) Non-responsiveness to a foot and mouth disease virus peptide overcome by addition of foreign helper T cell determinants. Nature 330, 168-170.

Fujiyama A, Miyanohara A, Nozaki C, Yoneyama, Oktamo N, Matsubara K. (1983) Cloning and structural analysis of HBV DNA adr. Nucleic Acids Res Com 11, 4601-4610.

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

Gan R, Shen L, Chur M, Li Z. (1984) The nucleotide sequence of HBsAg subtype adr. Scientia sinica 27, 926-935.

Ganem D, Varmus HE. (1987) The molecular biology of the hepatitis B viruses. Ann Rev Biochem 56, 651-693.

Garnier J, Osguthorpe DJ, Robson B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 120, 97-120.

Gavilanes F, Gonzalez-Ros JM, Peterson DL. (1982) Structure of hepatitis B virus surface antigen. J Biol Chem 257, 7770-7777.

Geisse S, Scheider C, Westphal HM, Hynes WE, Groner B, Beato M. (1982) Glucocorticoid receptors recognize DNA sequences in and around murine mammary-tumour virus-DNA. EMBO J 1, 1613-1619.

Gerber MA, Hadziyannis S, Vissoulis C, Schaffer F, Paranetto F, Popper H. (1974) Electron microscopy and immunoelectronmicroscopy of cytoplasmic HBsAg in hepatocytes. Am J Pathol 75, 489-496.

Gerin JL, Alexander H, Shih JW-K, Purcell RH, Dopolito G, Engle R, Green N, Sutcliffe JG, Shinnick TM, Lerner RA. (1983) Chemically synthesised peptides of HBsAg duplicate the d/y subtype specificities and induced subtype specific antibodies in chimpanzees. Proc Natl Acad Sci USA 80, 2365-2369.

Gerlich W, Robinson WS. (1980) A protein covalently attached to 5' of an incomplete DNA strand. Cell 21, 801-809.

Geysen HM, Tainer JA, Rodda SJ, Mason TJ, Alexander H, Getzoff ED, Lerner RA. (1987) Chemistry of antibody binding to a protein. Science 235, 1184-1190.

Glaser T, Lewis WH, Bruns GAP, Watkins PC, Rogler CE, Shows TB, Powers UE, Willard HF, Goguen JM, Simola KOJ, Housman DE. The β-subunit of follicle-stimulating hormone is deleted in patients with aniridia and Wilms' tumour, allowing further definition of the WAGR locus. Nature 321, 882-886.

Gough NM. (1983) Core and e antigen synthesis in rodent cells transformed with hepatitis B virus DNA is associated with greater than genome length viral messenger RNAs. J Mol Biol 165, 683-699.

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

Gough JA, Murray NE. (1983) Sequence diversity amoung related genes for recognition of specific targets in DNA molecules. J Mol Biol 166, 1-20:

Gronenborn B. (1976) Overproduction of phage lambda repressor under control of the lac promoter of E. coli. Mol Gen Genet 148, 243-250.

Grunstein M, Hogness DS. (1975) Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc Natl Acad Sci USA 72, 3961-3965.

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

Hess G, Arnold W, Meyer zum Buschenfelde KH. (1981) Demonstration and partial characterization of an intermediate HBcAg (Dane particle) population. J Med Virol 7, 241-250.

Hitzeman R, Chen C, Hugie F, Potezer E, Chung-Cheng L, Estell D, Miller JU, Yaffe A, Kleid D, Levinson A, Oppermann H. (1983) Expression of hepatitis B surface antigen in yeast. Nucleic Acids Res Com 11, 2745-2763.

Hopp TP, Woods KR. (1981) Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci USA 78, 3824-3828.

Imai M, Yanase Y, Nojiri T, Miyakowa Y, Mayumi M. (1979) A receptor for poly-human and chimpanzee albumins on HBV particles co-occurring with HBeAg. Gastroenterology 76, 242-247.

Janin J, Wodak S, Levitt M, Maigret B. (1978) Conformation of amino acid side-chains in proteins. J Mol Biol 125, 357-386.

Junker M, Galle P, Schaller H. (1987) Expression and replication of the HBV genome under foreign promoter control. Nucleic Acids Res Com 15, 10117-10132.

Kaplan PM, Greenman RL, Gerin JL, Purcell RH, Robinson WS. (1973) DNA polymerase associated with human hepatitis B antigen. J Virol 12, 995-1005.

Katz DH, Hamaoka T, Dorf ME, Maurer PH, Benacerraf BJ. (1973) Cell interactions between histoincompatible T and B lymphocytes. J Exp Med 138, 734-739.

Kay A, Mandart E, Trepo C, Galibert F. (1985) The HBV HBx gene expressed in E. coli is recognised by sera from hepatitis patients. EMBO J 4, 1287-1272.

Kobayashi M, Kaike K. (1984) Complete nucleotide sequence of hepatitis B virus DNA of subtype adr and its conserved gene organization. Gene 30, 227-232.

Kramer W, Drusta V, Jansen H-W, Kramer B, Pfugfelder M, Fritz H-J. (1984a) The gapped duplex DNA approach to oligo-directed mutation control. Nucleic Acids Res Com 12, 9441-9456.

Kramer B, Kramer W, Fritz H-J. (1984b) Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of E. coli. Cell 38, 879-887.

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

Le Bouvier GL. (1971) The heterogeneity of Australia antigen. J Infect Dis 123, 671-675.

Lerner RA, Green N, Alexander H, Liu F-T, Sutcliffe JG, Shinnick TM. (1981) Chemically synthesised peptides predicted from nucleotide sequence of hepatitis B virus genome elicit antibodies reactive with the native envelope protein of Dane particles. Proc Natl Acad Sci USA 78, 3403-3407.

Lien JM, Aldrich CE, Mason WS. (1986) Evidence that a capped oligonucleotide is the primer for duck hepatitis B virus plus-strand DNA synthesis. J Virol 57, 229-237.

Lo K-M, Jones SS, Hackett NR, Khorara HG. (1984) Specific amino acid substitutions in bacterioopsin: replacement of a restriction fragment in the structural gene by synthetic DNA fragments containing altered codons. Proc Natl Acad Sci USA 81, 2285-2289.

Lowry OH, Rosebrough NF, Farr AL, Randall RL. (1951) Protein measurement with the folin-phenol reagent. J Biol Chem 193, 265-275.

Machida A, Kishimoto S, Ohnuma H, Baba K, Ito Y, Miyamato H, Funastu G, Oda K, Usuda S, Togami S, Nakamura T, Miyahawa Y, Mayumi M. (1984) A polypeptide containing 55 amino acids coded by the pre-S of HBV bears receptor for human as well as chimpanzee albumins. Gastroenterology 86, 910-916.

MacKay P, Lees J, Murray K. (1981a) The converstion of hepatitis B core antigen synthesised in E. coli into e antigen. J Med Virol 8, 237-243.

MacKay P, Pasek M, Magazin M, Kovacic T, Allet B, Stahl S, Gilbert W, Schaller H, Bruce SA, Murray K. (1981b) Production of immunologically active HBsAg by E. coli. Proc Natl Acad Sci USA 78, 4510-4514.

Margalit H, Spouge JL, Cornette KB, Cease C, DeLisi C, Berzofsky JA. (1987) Prediction of immunodominant helper T cell antigenic sites from the primary sequence. J Immunol 38, 2213-2224.

Marion PL, Oshiro LS, Regnery DC, Scullard GH, Robinson WS. (1980) A virus in beechy ground squirrels which is related to hepatitis B of man. Proc Natl Acad Sci USA 77, 2941-2943.

Mason WS, Aldrich C, Summers JW, Taylor JM. (1982) Asymmetric replication of duck hepatitis B virus DNA in liver cells: Free minus-strand DNA. Proc Natl Acad Sci USA 79, 3997-4001.

Mason WS, Halpern MS, England JM, Seal G, Egan J, Coates L, Aldrich C, Summers JW. (1983) Experimental transmission of duck hepatitis B virus. Virology 131, 375-384.

Mason, WS, Seal G, Summers JW. (1980) Virus of pekin ducks with structure and biological relatedness to human hepatitis B virus. J Virol 36, 35-96.

Maupas P, Guesry P (eds). (1981) Hepatitis B vaccine. Inserm Symposium no 18. Elsevier/North Holland Biomedical Press, Amsterdam.

Mazzur S, Burgent S, Le Bouvier GL. (1975) Compound $\underline{d}^{+}\underline{y}^{+}$ particles of Australia antigen. J Immunol 114, 1510-1512.

McAuliffe VJ, Purcell RH, Gerin JL. (1980) Type-B hepatitis - a review of current prospects for a safe and effective vaccine. Rev Infect Dis 2, 470-492.

McGlynn E, Murray K. (1988) The hepatitis B virus polymerase: Expression of its gene in Escherichia coli, and detection of antibodies to the product in convalescent sera. In: Hepatitis and Livers Disease (Ed AJ Zuckerman) Alan R Liss Inc, New York.

McLachlan A, Milich DR, Raney AK, Riggs MG, Hughes JL, Sorge J, Chisari FV. (1987) Expression of HBV core and surface antigens; influences of pre S and precore sequences. J Virol 61, 683-692.

Messing J, Viera J. (1982) A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene 19, 269-276.

Meyers ML, Trepo LT, Nath N, Sninsky JL. (1986) Hepatitis B virus polypeptide X: Expression in Escherichia coli and identification of specific antibodies in sera from hepatitis B virus-infected humans. J Virol 57, 101-109.

Milich DR, Leroux-Roels GG, Louie RE, Chisari FV. Genetic regulation of the immune response to hepatitis B surface antigen (HBsAg). (1984) IV Distinct H-2-linked Ir genes control antibody responses to different HBsAg determinants on the same molecule and map to the I-A and I-C subregions. J Exp Med 159, 41-56.

Milich DR, 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 DR, McLachlan A, Moriarty A, Thornton GB. (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 DR, McLachlan A, Thornton GB, Hughes JL. (1987b) Antibody production to nucleocapsid and envelope of HBV primed by a single T-cell site. Nature 329, 547-549.

Milich DR, McLachlan A, Moriarty A, Thornton GB. (1987c) A single 10-residue pre-S (1) peptide can prime T cell help for antibody production to multiple epitopes within the pre-S (1), pre-S (2) and S regions of HBsAg. J Immunol 138, 4457-4465.

Milich DR, Hughes JL, McLachlan A, Thornton GB, Moriarty A. (1988) Hepatitis B synthetic immunogen comprised of nucleocapsid T-cell sites on an envelope B-cell epitopes. Proc Natl Acad Sci USA 85, 1610-1614.

Miller RH. (1987) Proteolytic self-cleavage of hepatitis B virus core protein may generate serum e antigen. Science 236, 722-725.

Miller RH, Robinson WS. (1984) Hepatitis B virus DNA forms in nuclear and cytoplasmic fractions of infected human liver. Virology 137, 390-399.

Mishiro S, Imai K, Takahashi K, Machida T, Gotanda Y, Mayumi M. (1980) A 49kD glycoprotein bearing all antigenic determinants and full immunogenicity of 22nm HBsAg particles. J Immunol 124, 1589-1593.

Mitchison NA. (1971) The carrier effect in the secondary response to hapten-protein conjugates. Eur J Immunol 1, 10-17.

Miyanohara A, Toke EA, Nozaki C, Fukusaburo H, Oktamo N, Matsubara K. (1983) Expression of hepatitis B surface antigen gene in yeast. Proc Natl Acad Sci USA 80, 1-5.

Molnar-Kimber KL, Summers JW, Mason WS. (1984) Mapping of the cohesive overlap of duck hepatitis B virus DNA and of the site of initiation of reverse transcription. J Virol 51, 181-191.

Mondelli M, Miele Vergoni G, Alberti A, Vergani D, Portmann B, Eddleston ALWF, Williams R. (1982) Specificity of T lymphocyte cytotoxicity to autologous hepatocytes are directed against HBV core antigen expressed on hepatocytes. J Immunol 129, 2773-2778.

Moriarty AM, Alexander H, Lerner RA, Thornton GB. (1985) Antibodies to peptides detect new hepatitis B antigen: serological correlation with hepatocellular carcinoma. Science 277, 420-433.

Moriarty AM, Hayer B, Shih J, Gerin W-K, Hamer D. (1981) Expression of the hepatitis B virus surface antigen gene in cell culture using Simian Virus 40 vector. Proc Natl Acad Sci USA 78, 2606-2610.

Moroy T, Etiemble J, Trepo C, Tiollais P, Buendia MA. (1985) Transcription of woodchuck hepatitis virus in the chronically infected liver. EMBO J 4, 1507-1514.

Murphy BL, Maynard JE, Le Bouvier GL. (1974) Viral subtypes and cross-protection in hepatitis B virus infection of chimpanzees. Intervirology 3, 378-381.

Murray K. (1987) The Leeuwenhoek Lecture, 1985. A molecular biologist's view of viral hepatitis. Proc R Soc Lond 230, 107-146.

Murray K, Bruce SA, Wingfield P, van Eerd PMCA, de Reus A, Schellekens H. (1987) Protective immunisation against hepatitis B with an internal antigen of the virus. J Med Virol 23, 101-107.

Murray K, Bruce SA, Hinnen A, Wingfield P, van Eerd PMCA, de Reus A, Schellekens H. (1984) Hepatitis B virus antigens made in microbial cells immunise against viral infection. EMBO J 3, 645-650.

Neurath AR, Kent SBH, Parker K, Prince A, Strick N, Brotman B, Sproul P. (1986a) Antibodies to a synthetic peptide from the pre S 120-145 region of HBV *env* are virus neutralizing. Vaccine 4, 35-37.

Neurath AR, Kent SBH, Stick N, Taylor P, Stevens, C. (1985) Hepatitis B virus contains pre-S gene encoded domains. Nature 315, 154-156.

Neurath AR, Kent SBH, Strick N, Parker K. (1986b) Identification and chemical synthesis of a host cell receptor binding site on HBV. Cell 46, 429-436.

Neurath AR, Kent SBH, Strick N. (1984) Location and chemical synthesis of a pre S gene coded immunodominant epitope of HBV. Science 224, 392-395.

Nonderfelt E, Le Bouvier GL. (1973) HBsAg with both d and y specificites on the same particle. Intervirology 2, 65-74.

Novotny J, Handschumacher M, Haber E, Bruccoliri PE, Carlson WB, Farning DW, Smith JA, Rose GD. (1986) Antigenic determinants in proteins coincide with surface regions accessible to large probes (antibody domains). Proc Natl Acad Sci USA 83, 226-230.

Okamoto H, Imai M, Miyakawa Y, Mayumi M. (1987b) Site-directed mutagenesis of HBsAg at $160^{arg \rightarrow lys}$ for the conversion of subtypic determinant r → w. Biochem Biophys Res Comm 148, 500-504.

Okamoto H, Imai M, Shimazaki M, Hashi Y, Iizuha H, Gotarda G, Tsuda F, Mihahawa Y, Mayumi M. (1986) Nucleotide sequence of a cloned HBV genome, subtype ayr: comparisons of genomes of other 3 subtypes. J Gen Virol 67, 2305-2314.

Okamoto H, Imai M, Tsuda F, Tanaka T, Miyakawa Y, Mayumi M. (1987a) Point mutation in S gene HBV for d/y subtype change in two blood donors carrying a surface of compound subtype adyr adwr. J Virol 61, 3030-3034.

Ono Y, Onda H, Sasada R, Igorashi K, Sugino Y, Nishioka K. (1983) The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype adr and adw. Nucleic Acids Res Com 11, 1747-1757.

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

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

Patzer E,Nakamura G, Yaffe A. (1984) Intracellular transport and secretion of HBsAg in mammalian cells. J Virol 51, 346-353.

Patzer E, Nakamura G, Simonsen, C, Levinson A, Brands R. (1986) Intracellular assembly and packaging of HBsAg particles occur in the endoplasmic reticulum. J Virol 58, 884-892.

Paul DA, Purcell RH, Peterson DL. (1986) Use of monoclonal antibodies to determine HBsAg of mixed subtype is on one particle or two. J Virol Methods 13, 43-53.

Persing DH, Varmus HE, Ganem D. (1985) A frameshift mutation in the pre-S region of human hepatitis B virus genome allows production of surface antigen particles by eliminating binding to polymerised albumin. Proc Natl Acad Sci USA 82, 3440-3444.

Persing DH, Varmus HE, Ganem D. (1986b) Inhibition of secretion of HBsAg by a related pre S polypeptide. Science 234, 1388-1392.

Persing DH, Varmus HE, Ganem D. (1986a) Antibodies to pre S and X determinants arise during natural infection with ground squirrel hepatitis virus. J Virol 60, 177-184.

Persing DH, Varmus HE, Ganem D. (1987) The pre S1 protein of HBV is acylated at its N-terminus with myristic acid. J Virol 61, 1672-1677.

Peterson DL. (1981) Isolation and characterization of the major protein and glycoprotein of HBsAg. J Biol Chem 256, 6975-6983.

Peterson DL, Nath N, Gavilanes F. (1982) Structure of HBsAg. Correlation of subtype with amino acid sequence and location of carbohydrate moiety. J Biol Chem 257, 10414-10420.

Peterson DL, Paul DA, Lam J, Tribby IIE, Achard DT. (1984) Antigenic structure of hepatitis B surface antigen: Identification of the "d" subtype determinant by chemical modification and use of monoclonal antibodies. J Immunol 132, 920-927.

Peterson DL, Roberts IM, Vyas GN. (1977) Partial amino acid sequence of two major component polypeptides of hepatitis B surface antigen. Proc Natl Acad Sci USA 74, 1530-1534.

Pourcel CH, Louise A, Gervais M, Cheneirer N, Dubois M-F, Tiollais P. (1982) Transcription of hepatitis B surface antigen gene in mouse cells transformed with cloned viral DNA. J Virol 42, 100-105.

Prince AM, Ikram H, Hopp TP. (1982) Hepatitis B virus vaccine: Identification of HBsAg/a and HBsAg/d but not HBsAg/y subtype antigenic determinants on a synthetic immunogenic peptide. Proc Natl Acad Sci USA 79, 579-582.

Pugh JC, Weber C, Houston H, Murray K. (1986) Expression of the X gene of hepatitis B virus. J Med Virol 20, 229-246.

Rall LB, Standring DN, Laub O, Rutter WJ. (1983) Transcription of hepatitis B virus by RNA polymerase II. Mol Cell Biol 3, 1766-1773.

Rees AR, de la Paz P. (1986) Investigating antibody specificity using computer graphics and protein engineering. TIBS 11, 144-148.

Roberts IM, Bernard CC, Vyas GN, MacKay I. (1975) T-cell dependence of immune response to HBsAg in mice. Nature 154, 606-607.

Roberts S, Chetham JC, Rees AR. (1987) Generation of an antibody with enhanced affinity and specificity for its antigen by protein engineering. Nature 328, 731-734.

Robinson WS. (1977) The genome of hepatitis B virus. Am Rev Microbiol 13, 357-377.

Robinson WS, Greenman RL. (1974) DNA polymerase in the core of the human hepatitis B virus candidate. J Virol 13, 1231-1236.

Rogler CE, Sherman M, Su CY, Shafritz D, Summers JW. (1985) Deletion in chromosome 11p associated with a hepatitis B integration site in hepatocellular carcinoma. Science 230, 319-322.

Roitt I, Brostoff J, Male D (eds). (1985) Immunology. Gower Medical Publishing, London.

Roossinck M, Jameel S, Loukin S, Siddiqui A. (1986) Expression of HBcAg in mammalian cells. Mol Cell Biol 6, 1393-1400.

Roossinck M, Siddiqui A. (1987) In vivo phosphorylation and protein analysis of HBcAg. J Virol 61, 955-961.

Rothbard JB. (1986) Peptides and the cellular immune response. Ann Inst Pasteur Virol 137E, 518-528.

Sarger F, Nicklen S, Coulson AR. (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74, 5463-5467.

Schaeffer E,Snyder R, Sninsky J. (1986) Location and function of pre S polypeptide in woodchuck hepatitis virus. J Virol 57, 173-182.

Schlicht HJ, Salfeld J, Schaller H. (1987) DHBV pre C encodes a signal sequence which is essential for synthesis and secretion of processed core proteins but not for virus formation. J Virol 61, 3701-3709.

Schurr A, Yagil E. (1971) Regulation and characterization of acid and alkaline phosphatase in yeast. J Gen Microbiol 65, 291-301.

Seeger C, Ganem D, Varmus HE. (1986) Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science 232, 477-483.

Serfling E, Jasin M, Schaffner W. (1985) Enhancers and eukaryotic gene transcription. Trends Genet 1, 224-231.

Shaul Y, Ben-Levy R. (1987) Multiple nuclear proteins in liver cells are bound to hepatitis B virus enhancer and its upstream sequences. EMBO J 6, 1913-1920.

Shaul Y, Rutter WJ, Laub O. (1985) A human hepatitis B viral enhancer element. EMBO J 4, 427-430.

Sheriff S, Silverton EW, Padlan EA, Cohen GH, Smith-Gill SJ, Finzel BC, Davies DR. (1987) Three-dimensional structure of an antibody antigen complex. Proc Natl Acad Sci USA 84, 8075-8079.

Siddiqui A, Jameel S, Mapoles J. (1986) Transcriptional control elements of hepatitis B surface antigen gene. Proc Natl Acad Sci USA 83, 566-570.

Skelly J, Howard E, Zuckermann AJ. (1978) The labelling of galactose residues in HBsAg glycoprotein. J Gen Virol 41, 447-457.

Sobeslavsky O. (1978) In: Viral hepatitis. (Eds GN Vyas, Cohen SN, Schmid R). The Franklin Institute Press, Philadelphia, Pennsylvania.

Spandau DF, Lee C-H. (1988) trans-Activation of viral enhancers by HBx protein. J Virol 62, 427-434.

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

Standring D, Ou J, Rutter WJ. (1986) Assembly of viral particles in Xenopus oocytes: pre S-antigens regulates secretion of the HBsAg env particle. Proc Natl Acad Sci USA 83, 9338-9342.

Standring DN, Rall LB, Laub O, Rutter WJ. (1983) Hepatitis B virus encodes an RNA polymerase III transcript. Mol Cell Biol 3, 1774-1782.

Steiner S, Huebner MT, Dreesman GR. (1974) Major polar lipids of HBsAg preparations: evidence for presence of a glycophospholipid. J Virol 14, 572-577.

Stibbe W, Gerlich WH. (1983) Structural relationships between minor and major proteins of HBsAg. J Virol 46, 626-628.

Sukeno N, Shirachi R, Yamaguchi N, Ishida N. (1972) Reduction and reoxidation of Australia antigen: loss and reconstitution of particle structure and antigenicity. J Virol 9, 182-183.

Summers JW, Mason WS. (1982) Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate. Cell 29, 403-415.

Summers JW, Smolec J, 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.

Sureau C, Romet-Lemonne J-L, Mulling JI, Essex M. (1986) Production of hepatitis B virus by a differentiated human hepatoma cell line after transformation with cloned circular HBV DNA. Cell 47, 37-47.

Swenson PD, Peterson DL, Piesheng H. (1988) Antigenic analysis of HBsAg with monoclonal antibodies specific for S-protein and pre-S2 sequences. In: Viral Hepatitis and Liver Disease (Ed AJ Zuckerman) Alan R Liss Inc, New York.

Szmuness W, Alter HJ, Maynard JE. (eds) (1982) Viral hepatitis 1981. International symposium. Franklin Institute Press, Philadelphia, Pennsylvania.

Szmuness W, Stevens, CE, Harley EJ, Zang EA, Oleszko WE, William DC, Sadovsky R, Morrison J, Kelner A. (1980) Hepatitis B vaccine. Demonstration of efficacy in a controlled clinical trial in a high risk population in the United States. New Engl J Med 203, 833-841.

Tagawa M, Omata M, Okuda K. (1986) Appearance of viral RNA transcripts in the early stage of duck hepatitis B virus infection. Virology 152, 477-482.

Takahashi K, Michada A, Funatsu G, Nomura M, Usuda S, Aoyagi S, Tachibana K, Miyamoto H, Imai M, Nakamura T, Miyakawa Q, Mayumi M. (1983) Immunochemical structure of HBeAg in the serum. J Immunol 130, 2903-2908.

Taylor JW, Ott J, Eckstein F. (1985a) Generation of oligonucleotide-directed mutations using phosphorothioate-DNA. Nucleic Acids Res Com 13, 8765-8785.

Taylor JW, Schmidt W, Cosstick R, Okruszek A, Eckstein F. (1985b) The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA. Nucleic Acids Res Com 13, 8749-8764.

Tedder RS, Yao JL, Ferns RB. (1984) In: Proceedings of an International Hepatitis Workshop. Nuclear Enterprises Ltd., Edinburgh, Scotland.

Theilmann L, Klinkert M-Q, Gmelin K, Kommerell B, Pfaff E. (1987) Detection of antibodies against pre S1 proteins in sera of patients with HBV infection. Hepatology 4, 22-28.

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

Thornton JM, Edwards MS, Taylor WR, Barlow DJ. (1986) Location of 'continuous' antigenic determinants in the protruding regions of proteins. EMBO J 5, 409-413.

Thung SN, Gerber MA. (1981) Presence of receptors for polyalbumin in HBsAg containing hepatocytes and hepatoma cell line. Hepatology 1, 132-136.

Tiollais P, Charnay P, Vyas GN. (1981) Biology of hepatitis B virus. Science 213, 406-411.

Tognani A, Cattaneo R, Serfling E, Schaffner W. (1985) A novel expression selection approach allows precise mapping of hepatitis B virus enhancer. Nucleic Acids Res Com 13, 7457-7472.

Towbin H, Staebelin T, 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.

Townsend ARM, Gotch FM, Davey J. (1985) Cytotoxic T cells recognise fragments of the influenza nucleoprotein. Cell 42, 457-467.

Townsend ARM, Rothbard JB, Gotch FM, Bahadur G, Wraith D, McMichael AJ. (1986) The epitopes of influencae nucleprotein recognised by cytotoxic T lymphocytes can be defined with short synthetic peptides. Cell 44, 959-968

Treinin M, Laub O. (1987) Identification of a promoter element located from the hepatitis B virus X gene. Mol Cell Biol 7, 545-548.

Tur-Kaspa R, Bank RD, Shaul Y, Shafritz, DA. (1986) Hepatitis B virus DNA contains a glucocorticoid responsive element. Proc Natl Acad Sci USA 83, 1627-1631.

Tuttleman J, Pourcel CH, Summers JW. (1986) Formation of the pool of covalently closed circular viral DNA in hepadna virus-infected cells. Cell 47, 451-460.

Twu J-S, Schloemer RH. (1987) Transcriptional trans-activating function of hepatitis B virus. J Virol 61, 3448-3453.

Valenzuela P, Coit D, Kuo CH. (1985a) Synthesis and assembly in yeast of hepatitis B surface antigen particles containing the polyalbumin receptor. Biotechnology 3, 317-320.

Valenzuela P, Coit D, Medina-Selby M, Kuo C, Van Nest G, Burke R, Bull P, Undea M, Graves P. (1985b) Antigen engineering in yeast: synthesis and assembly of hybrid hepatitis B surface antigen-herpes simplex IgD particles. Biotechnology 3, 323-326.

Valenzuela P, Gray P, Quiraga M, Zaldivr J, Goodman HM, Rutter WJ. (1979) Nucleotide sequence of the gene coding for the major protein of hepatitis B virus surface antigen. Nature 280, 815-819.

Valenzuela P, Medina A, Rutter WJ, Ammerer G, Hall B. (1982) Synthesis and assembly of hepatitis B virus surface antigen particles in yeast. Nature 298, 347-350.

van Kooten Kok-Doorschadt JH, van den Alcher R, Gisjen R. Deterimination and distribution of two types of hepatitis-associated antigen. J Infect Dis 126, 117-122.

Varmus HE. (1984) The molecular genetics of cellular oncogenes. Am Rev Genet 18, 553-612.

Varmus HE, Swanstrom R. (1985) In: Molecular biology of tumour viruses. (Ed J Tooze) Cold Spring Harbor Press, New York.

Vento S, Chen S-H, Giuliani-Piccari G, O'Brien C, Dal Monte P, Eddleston ALWF, Howard C, Williams, R. (1987) Cellular immunity to nucleocapsid and pre S determinants in asymptomatic hepatitis B virus carriers. Immunology 62, 593-598.

Vy A, Bruss V, Gerlich WH, Köchel H, Thomssen R. (1986) Precore sequences of hepatitis B virus inducing e antigen and membrane association of the viral core protein. Virology 155, 89-96.

Vyas GN, Cohen SN, Schmid R. (eds) (1978) Viral hepatitis. Franklin Institute Press, Philadelphia, Pennsylvania.

Vyas GN, Rao KR, Ibrahim AB. (1972) Australia antigen (hepatitis B antigen). A conformation antigen dependent on disulphide bonds. Science 178, 1300-1301.

Wallace RB, Schaffer J, Murphy RF, Bonner J, Hirose T, Itakura K. (1979) Hybridization of synthetic oligodeoxyribonucleotides to ϕx1974 DNA: The effect of single base pair mismatch. Nucleic Acids Res Com 6, 3543-3556.

Wampler DE, Lehman ED, Boger J, McAleer WJ, Scolnick EM. (1985) Multiple chemical forms of HBsAg produced in yeast. Proc Natl Acad Sci USA 82, 6830-6834.

Weiser B, Ganem D, Seeger C, Varmus HE. (1983) Closed circular viral DNA on asymmetrical heterogeneous forms in livers from animals infected with ground squirrel hepatitis virus. J Virol 48, 1-9.

Will H, Salfeld J, Pfaff E, Manso C, Theilman L, Schaller H. (1986) Putative reverse transcriptase intermediates of human hepatitis B virus in primary liver carcinomas. Science 231, 594-596.

Yaginuma K, Shirakat Y, Kobayashi M, Kioke 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, 2678-2682.

Yamada G, Sakamoto Y, Mizuno H, Nishihara T, Kobayashi T, Takohashi T, Nagashima H. (1982) Electron and immunoelectron microscopic study of Dane particle formation in chronic hepatitis B virus infection. Gastroenterology 83, 348-356.

Yu M, Finlayson J, Shih J. (1985) Interaction between various polymerized human albumins and HBsAg. J Virol 55, 736-743.

Ziegler K, Unanue ER. (1981) Identification of a macrophage antigen-processing event required for I-region-restricted antigen presentation to T-lymphocytes. J Immunol 127, 1869-1875.

Zoller MJ, Smith M. (1983) Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. Meth Enzymol 100, 468-500.

APPENDIX A

Nucleotide sequence of the HBV genome

Shown is the DNA sequence of the HBV genome (subtype ayw) as cloned by Pasek et al. (1979), 3182 nucleotides long. The numbering of nucleotides is based on the start of the core gene being 1. Below the gene sequences are the corresponding translation products in all three phases (a, b, c). The starts of HBV ORFs map to the following positions: Core - 1a, polymerase - 407b, pre Surface 1 - 948c, pre Surface 2 - 1272c, Surface - 1437c, X ORF - 2656a.

```
                ATGGACATTGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCCGTACGAGATCTTCTAGATACCGCCGCAGCTCTGTATCGGGAT
          1     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+    120
a:              MetAspIleAspProTyrLysGluPheGlyAlaThrValGluLeuLeuSerPheLeuProSerAspPhePheProSerValArgAspLeuLeuAspThrAlaAlaAlaLeuTyrArgAsp

                GCCTTAGAGTCTCCTGAGCATTGTTCACCTCACCATACTGCACTCAGGCAAGCAATTCTTTGCTGGGGAGACTTAATGACTCTAGCTACCTGGGTGGGTACTAATTTAGAAGATCCAGCA
        121     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+    240
a:              AlaLeuGluSerProGluHisCysSerProHisHisThrAlaLeuArgGlnAlaIleLeuCysTrpGlyAspLeuMetThrLeuAlaThrTrpValGlyThrAsnLeuGluAspProAla

                TCTAGGGACCTAGTAGTCAGTTATGTCAACACTAATGTGGGCCTAAAGTTCAGACAATTATTGTGGTTTCACATTTCTTGTCTCACTTTTGGAAGAGAAACGGTTCTAGAGTATTTGGTG
        241     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+    360
a:              SerArgAspLeuValValSerTyrValAsnThrAsnValGlyLeuLysPheArgGlnLeuLeuTrpPheHisIleSerCysLeuThrPheGlyArgGluThrValLeuGluTyrLeuVal

                TCTTTTGGAGTGTGGATTCGCACTCCTCCAGCTTATAGACCACCAAATGCCCCTATCCTATCAACACTTCCGGAGACTACTGTTGTTAGACGACGATGCAGGTCCCCTAGAAGAAGAACT
        361     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+    480
a:              SerPheGlyValTrpIleArgThrProProAlaTyrArgProProAsnAlaProIleLeuSerThrLeuProGluThrThrValValArgArgArgCysArgSerProArgArgArgThr
b:                                                                                  MetProLeuSerTyrGlnHisPheArgArgLeuLeuLeuLeuAspAspAspAlaGlyProLeuGluGluGluLe

                CCCTCGCCTCGCAGACGAAGATCTCAATCGCCGCGTCGCAGAAGATCTCAATCTCGGGAATCTCAATGTTAGTATCCCTTGGACTCATAAGGTGGGAAATTTTACTGGGCTTTATTCTTC
        481     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+    600
a:              ProSerProArgArgArgArgSerGlnSerProArgArgArgArgSerGlnSerArgGluSerGlnCys
b:              uProArgLeuAlaAspGluAspLeuAsnArgArgValAlaGluAspLeuAsnLeuGlyAsnLeuAsnValSerIleProTrpThrHisLysValGlyAsnPheThrGlyLeuTyrSerSe

                TACTGTACCTGTCTTTAACCCTCATTGGAAAACGCCCTCTTTTCCTAATATACATTTACACCAAGATATTATCAAAAAATGTGAACAGTTTGTAGGGCCGCTCACAGTCAATGAGAAAAG
        601     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+    720
b:              rThrValProValPheAsnProHisTrpLysThrProSerPheProAsnIleHisLeuHisGlnAspIleIleLysLysCysGluGlnPheValGlyProLeuThrValAsnGluLysAr

                AAGGTTAAAATTGATCATGCCTGCTAGGTTTTATCCTAATTTTACCAAATATTTGCCCTTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAATCATTACTTCCAAACTAG
        721     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+    840
b:              gArgLeuLysLeuIleMetProAlaArgPheTyrProAsnPheThrLysTyrLeuProLeuAspLysGlyIleLysProTyrTyrProGluHisLeuValAsnHisTyrPheGlnThrAr

                ACACTATTTACACACTCTATGGAAGGCGGGTGTTTTATATAAGAGAGTATCAACACATAGCGCCTCATTTTGTGGGTCACCATATTCTTGGGAACAAGAGCTACAGCATGGGGCAGAATC
        841     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+    960
b:              gHisTyrLeuHisThrLeuTrpLysAlaGlyValLeuTyrLysArgValSerThrHisSerAlaSerPheCysGlySerProTyrSerTrpGluGlnGluLeuGlnHisGlyAlaGluSe
c:                                                                                                                  MetGlyGlnAsnL

                TTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGATCCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCA
        961     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+   1080
b:              rPheHisGlnGlnSerSerGlyIleLeuSerArgProProValGlySerSerLeuGlnSerLysHisGlnGlnSerArgLeuGlyLeuGlnSerGlnGlnGlyHisLeuAlaArgArgGl
c:              euSerThrSerAsnProLeuGlyPhePheProAspHisGlnLeuAspProAlaPheArgAlaAsnThrAsnAsnProAspTrpAspPheAsnProAsnLysAspThrTrpProAspAlaA

                ACAAGGTAGGAGCTGGAGCATTCGGGCTAGGGTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAATGCAAACCTTGCCAGCAAATCCGCCTCCTGCCT
       1081     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+   1200
b:              nGlnGlyArgSerTrpSerIleArgAlaArgValHisProThrAlaArgArgProPheGlyValGluProSerGlySerGlyHisAsnAlaAsnLeuAlaSerLysSerAlaSerCysLe
c:              snLysValGlyAlaGlyAlaPheGlyLeuGlyPheThrProProHisGlyGlyLeuLeuGlyTrpSerProGlnAlaGlnGlyIleMetGlnThrLeuProAlaAsnProProProAlaS
```

```
          CTACCAATCGCCAGTCAGGACGGCAGCCTACCCCGCTGTCTCCACCTCTGAGAACCACTCATCCTCAGGCCATGCAGTGGAACTCCACAACCTTCCACCAAACTCTGCAAGATCCCAGAG
     1201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1320
b:        uTyrGlnSerProValArgThrAlaAlaTyrProAlaValSerThrSerGluAsnHisSerSerSerGlyHisAlaValGluLeuHisAsnLeuProProAsnSerAlaArgSerGlnSe
c:        erThrAsnArgGlnSerGlyArgGlnProThrProLeuSerProProLeuArgThrThrHisProGlnAlaMetGlnTrpAsnSerThrThrPheHisGlnThrLeuGlnAspProArgV

          TGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGGACAGTAAACCCTGTTCCGACTACTGCCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGG
     1321 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1440
b:        rGluArgProValPheProCysTrpTrpLeuGlnPheArgAspSerLysProCysSerAspTyrCysLeuSerHisIleValAsnLeuLeuGluAspTrpGlyProCysAlaGluHisGl
c:        alArgGlyLeuTyrPheProAlaGlyGlySerSerSerGlyThrValAsnProValProThrThrAlaSerProIleSerSerIlePheSerArgIleGlyAspProAlaLeuAsnMetG

          AGAACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTC
     1441 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1560
b:        yGluHisHisIleArgIleProArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrAlaGluSerArgLeuValValAspPheSerGlnPheSe
c:        luAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePheLeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsnPheL

          TAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAATCTCCAATCACTCACCAACCTCCTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCA
     1561 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1680
b:        rArgGlyAsnTyrArgValSerTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHi
c:        euGlyGlyThrThrValCysLeuGlyGlnAsnSerGlnSerProIleSerAsnHisSerProThrSerCysProProThrCysProGlyTyrArgTrpMetCysLeuArgArgPheIleI

          TCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCATCAACCACCAGCACGGGACCCTGCA
     1681 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1800
b:        sLeuProLeuHisProAlaAlaMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSerAsnSerArgIleIleAsnHisGlnHisGlyThrLeuGl
c:        lePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTyrGlnGlyMetLeuProValCysProLeuIleProGlySerSerThrThrSerThrGlyProCysA

          GAACCTGCACGACTCCTGCTCAAGGAATCTCTATGTATCCCTCCTGTTGCTGTACAAAACCTTCGGATGGAAACTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCC
     1801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1920
b:        nAsnLeuHisAspSerCysSerArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyTrpLysLeuHisLeuTyrSerHisProIleIleLeuGlyPheArgLysIlePr
c:        rgThrCysThrThrProAlaGlnGlyIleSerMetTyrProSerCysCysCysThrLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheGlyLysPheL

          TATGGGAGTGGGCCTCAGCCCGTTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCATTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGG
     1921 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2040
b:        oMetGlyValGlyLeuSerProPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCysLeuAlaPheSerTyrMetAspAspValValLeuGl
c:        euTrpGluTrpAlaSerAlaArgPheSerTrpLeuSerLeuLeuValProPheValGlnTrpPheValGlyLeuSerProIleValTrpLeuSerValIleTrpMetMetTrpTyrTrpG

          GGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGCATACATTTAAACCCTAACAAAACAAAAAGATGGGGTTATTCTCTAAATTTCAT
     2041 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2160
b:        yAlaLysSerValGlnHisLeuGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMe
c:        lyProSerLeuTyrSerIleLeuSerProPheLeuProLeuLeuProIlePhePheCysLeuTrpAlaTyrIle
```

```
          GGGCTATGTCATTGGATGTTGGGGATCATTGCCACAAGATCACATCATACACAAAATCAAAGAATGTTTTAGAAAACTCCCTGTTCACAGGCCTATTGATTGGAAAGTGTGTCAACGAAT
     2161 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2280
b:        tGlyTyrValIleGlyCysTrpGlySerLeuProGlnAspHisIleIleHisLysIleLysGluCysPheArgLysLeuProValHisArgProIleAspTrpLysValCysGlnArgIl

          TGTGGGTCTTCTGGGTTTTGCTGCCCCTTTCACACAATGTGGTTATCCTGCTTTAATGCCCTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTTTCGCCAACTTACAAGGCCTT
     2281 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2400
b:        eValGlyLeuLeuGlyPheAlaAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLysGlnAlaPheThrPheSerProThrTyrLysAlaPh

          TCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCAGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGATGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCG
     2401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2520
b:        eLeuCysLysGlnTyrLeuAsnLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrProThrGlyTrpGlyLeuValMetGlyHisGlnArgMetAr

          TGGAACCTTTCTGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTCTCGGAACTGACAACTCTGTTGTCCTCTCCCGCAA
     2521 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2640
b:        gGlyThrPheLeuAlaProLeuProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAsnIleLeuGlyThrAspAsnSerValValLeuSerArgLy

          ATATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGCCGCTTGGG
     2641 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2760
a:                          MetAlaAlaArgLeuCysCysGlnLeuAspProAlaArgAspValLeuCysLeuArgProValGlyAlaGluSerCysGlyArgProPheSerGlyProLeuGly
b:        sTyrThrSerPheProTrpLeuLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAsnProAlaAspAspProSerArgGlyArgLeuGl

          GCCCTGTCGTCCTCTTCTCTGCCTGCCGTTCCGGCCGCCCACGGGGCGCACCTCTCTTTACGCGGCCTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCT
     2761 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2880
a:        AlaLeuSerSerSerSerLeuProAlaValProAlaAlaHisGlyAlaHisLeuSerLeuArgGlyLeuProValCysAlaPheSerSerAlaGlyProCysAlaLeuArgPheThrSer
b:        yProCysArgProLeuLeuCysLeuProPheArgProProThrGlyArgThrSerLeuTyrAlaAlaSerProSerValProSerHisLeuProAspArgValHisPheAlaSerProLe

          GCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATCTTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCTGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGT
     2881 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3000
a:        AlaArgArgMetGluThrThrValAsnAlaHisGlnIleLeuProLysValLeuHisLysArgThrLeuGlyLeuSerAlaMetSerThrThrAspLeuGluAlaTyrPheLysAspCys
b:        uHisValAlaTrpArgProPro

          TTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGATTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAA
     3001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3120
a:        LeuPheLysAspTrpGluGluLeuGlyGluGluIleArgLeuLysValPheValLeuGlyGlyCysArgHisLysLeuValCysAlaProAlaProCysAsnPhePheThrSerAla

          TCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGC
     3121 ---------+---------+---------+---------+---------+---------+--  3182
```

APPENDIX B

Nucleotide sequence of the S gene

Shown is the DNA sequence (1437 - 2114) of the S gene of HBV, as cloned by Pasek *et al*. (1979), subtype *ayw*. The numbering of nucleotides is as in Appendix A. Below the nucleotide sequence is the corresponding translation product.

```
          1450                1470                1490
ATGGAGAACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTTACAGGCGGGGTTTTTC
1437 ---+---------+---------+---------+---------+---------+------ 1496
MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe

          1510                1530                1550
TTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAAT
1497 ---+---------+---------+---------+---------+---------+------ 1556
LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn

          1570                1590                1610
TTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAATCTCCAATCAC
1557 ---+---------+---------+---------+---------+---------+------ 1616
PheLeuGlyGlyThrThrValCysLeuGlyGlnAsnSerGlnSerProIleSerAsnHis

          1630                1650                1670
TCACCAACCTCCTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTT
1617 ---+---------+---------+---------+---------+---------+------ 1676
SerProThrSerCysProProThrCysProGlyTyrArgTrpMetCysLeuArgArgPhe

          1690                1710                1730
ATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTAT
1677 ---+---------+---------+---------+---------+---------+------ 1736
IleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTyr

          1750                1770                1790
CAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCATCAACCACCAGCACGGGACCC
1737 ---+---------+---------+---------+---------+---------+------ 1796
GlnGlyMetLeuProValCysProLeuIleProGlySerSerThrThrSerThrGlyPro

          1810                1830                1850
TGCAGAACCTGCACGACTCCTGCTCAAGGAATCTCTATGTATCCCTCCTGTTGCTGTACA
1797 ---+---------+---------+---------+---------+---------+------ 1856
CysArgThrCysThrThrProAlaGlnGlyIleSerMetTyrProSerCysCysCysThr

          1870                1890                1910
AAACCTTCGGATGGAAACTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAA
1857 ---+---------+---------+---------+---------+---------+------ 1916
LysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheGlyLys

          1930                1950                1970
TTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTT
1917 ---+---------+---------+---------+---------+---------+------ 1976
PheLeuTrpGluTrpAlaSerAlaArgPheSerTrpLeuSerLeuLeuValProPheVal

          1990                2010                2030
CAGTGGTTCGTAGGGCTTTCCCCCATTGTTTGGCTTTCAGTTATATGGATGATGTGGTAT
1977 ---+---------+---------+---------+---------+---------+------ 2036
GlnTrpPheValGlyLeuSerProIleValTrpLeuSerValIleTrpMetMetTrpTyr

          2050                2070                2090
TGGGGGCCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTT
2037 ---+---------+---------+---------+---------+---------+------ 2096
TrpGlyProSerLeuTyrSerIleLeuSerProPheLeuProLeuLeuProIlePhePhe

          2110
TGTCTTTGGGCATACATT
2097 ---+---------+---- 2114
CysLeuTrpAlaTyrIle
```

Antigenic Index: This feature indicated the probability that a region be antigenic based on a summation of several, weighted, measures of secondary structure, such as flexibility, hydrophilicity and surface probability. The method used was that of Jameson and Wolf (CABIOS in press).

The methods of Chou and Fasman (CF) (1978) and Garnier, Osguthorpe and Robson (GOR) (1978) were used to predict helices, β-sheets and peptides turns.

Glycosylation sites: These were predicted for sites where the residues have the composition N X T or N X S, when X is D, W or P the site is taken to be a weak glycosylation site, otherwise it is taken as a strong glycosylation site, i.e. highly probable.

APPENDIX C

Biophysical secondary structure prediction for the S gene product of HBV

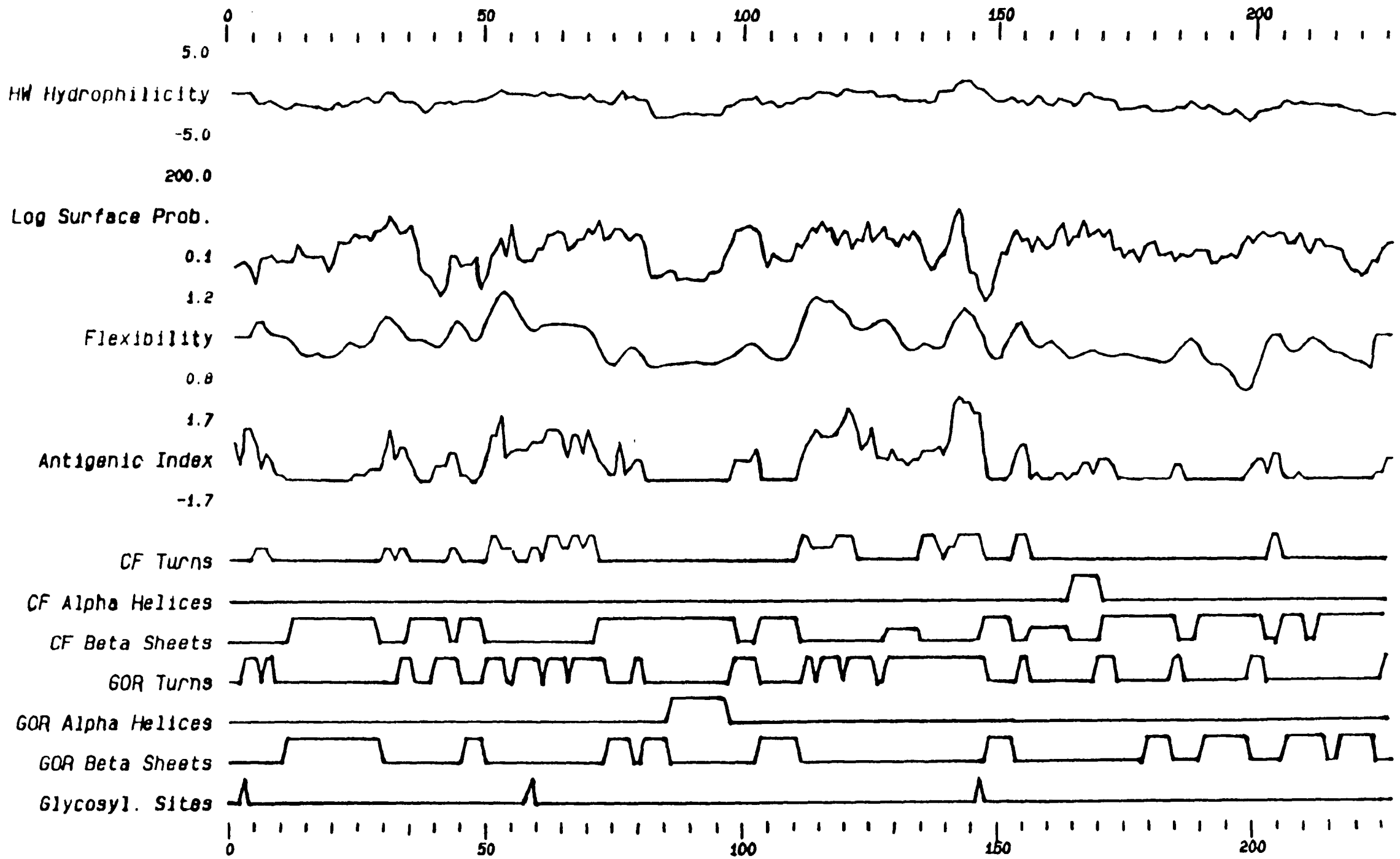
Using the peptide sequence of HBsAg, deduced from the nucleotide sequence of the S gene (Pasek et al., 1979; Appendix B), predictions of secondary structure were undertaken.

The UWGCG peptide structure program was used to perform the following predictions:

HW Hydrophilicity: The Hopp and Woods (1981) method for assessing local hydrophilicity of peptide regions was used. The window was set to 7 residues for this study.

Log Surface Probability: Using the method described by Emini et al. (1985) the probability of a short peptide region appearing on the hydrophilic surface of a protein was determined. Each amino acid is assigned a value relating to the degree of likelihood of it appearing at the surface of the molecule. These values are combined in the method to give an overall propensity of surface appearance for the molecule.

Flexibility: The predicted ease of movement about the alpha-carbon of amino acids in a peptide region was calculated according to the Karpus-Schulz method (unpublished).


0
50
100
150
200
5.0
HW Hydrophilicity
-5.0
200.0
Log Surface Prob.
0.1
1.2
Flexibility
0.8
1.7
Antigenic Index
-1.7
CF Turns
CF Alpha Helices
CF Beta Sheets
GOR Turns
GOR Alpha Helices
GOR Beta Sheets
Glycosyl. Sites
0
50
100
150
200

APPENDIX D

Homology based secondary structure prediction for the S gene product of HBV

The peptide sequence of HBsAg, deduced from the nucleotide sequence of the S gene cloned by Pasek et al. (1979) (Appendix B) was used in the method of Coulson et al. (1987) to predict the secondary structure of HBsAg.

The peptide sequence was compared with a database of protein sequences of known crystal structure; for every point of homology recorded, the relevant secondary structure present at that point was recorded. E - β-sheet, H - α-helix, T - turn.

An amino acid with no corresponding prediction showed no homology to any of the sequences stored in the database.

```
1   MENITSGFLG PLLVLQAGFF LLTRILTIPQ SLDSWWTSLN RLGGTIVCLG
    EEEEEEHT  T EEEEEETHHHH EE  EEE  E EE    EEEE HHHHEEEEEE
51  QNSQSPISNH SPTSCPPTCP GYRWMCLRRF IIFLFILLLG LIFLLVLLDY
    EEEEEEEEEEE EEEEEETT  T TEEEEHHHHH EEEEHHEEEH HHHEEEEEEE
101 GMLPVCPLI PSGGTTSTGP CRTCTTPAQG ISMYPSCCCT KRSDGNCTCI
    EEEE  T   TTTTEEE EE EEEEEETTEEE HHEETEEEEE ETEET EEEE
151 PIPSSWAFGK FLWEWASARF SWLSLLVPFV QWFVGLSPIV WLSVIWMMWY
    EEE EEEEEE HHHHHHHHHE EEEEEEEHHH HHHEEEEEEE EE EEEEEEE
201 WGPSLYSILS PFLPLLPIFF CLWAYI
    EE EEEEEEE EEHHHHHEEH HHHHHH
```