Phage-Display Libraries and the Identification of Epitopes of Hepatitis B Virus Surface Antigen

by

Volker Germaschewski

A Thesis presented for the Degree of Ph.D
Institute of Cell and Molecular Biology
University of Edinburgh
July 1995
CONTENTS

Abstract v
Acknowledgements vii
Abbreviations viii
Amino acids single letter code x

CHAPTER I: INTRODUCTION

I.A) Fusion-Phage and Peptide Libraries: Tools for Exploring Molecular Interactions
   I.A.1) The filamentous phage fd .............................................................. 2
   I.A.2) Fusion phage: Phage that display foreign peptides or protein domains .......................................................... 3
   I.A.3) Phage-display libraries ............................................................... 6
   I.A.4) The fuse5 hexapeptide library .................................................... 8
   I.A.5) "Biopanning" and factors that influence selection ......................... 10
   I.A.6) Applications for phage-display libraries .................................... 12

I.B) Antibody-Antigen Interactions
   I.B.1) Antibodies: Structure and general aspects ..................................... 16
   I.B.2) Antigen recognition: What are epitopes ? ..................................... 17
   I.B.3) Epitope mapping ........................................................................... 19

I.C) Hepatitis B Virus: Studying the Immune Response to a Complex Antigen
   I.C.1) The morphology of hepatitis B virus (HBV) ................................ 23
   I.C.2) Clinical aspects ........................................................................... 24
   I.C.3) Replication and transcription of the HBV genome .......................... 25
   I.C.4) The genes of HBV and viral gene products .................................... 27
   I.C.5) The surface antigen of HBV (HBsAg) ............................................. 31
   I.C.6) The immune response to HBV
          General aspects ................................................................................. 34
          HBsAg is the prime target for the immune system ............................. 35
   I.C.7) The monoclonal antibody MA18/7 ................................................ 39
   I.C.8) The polyclonal serum sample of a vaccinated chimpanzee ........... 39
CHAPTER II: RESULTS AND CONCLUSIONS

II.A) Epitope Mapping of the Monoclonal Antibody MA18/7
   II.A.1) Amplification of the phage-display library and verification of the random composition of the displayed peptides................... 43
   II.A.2) Defining the epitope recognised by MA18/7............................. 45
   II.A.3) Three non-linear core residues, D, P and F, are sufficient for specific interaction of phage and antibody MA18/7................. 48
   II.A.4) Synthetic peptides compete with selected phage for interaction with MA18/7......................................................... 59

II.B) Investigating Antibody Specificities in the Serum of a Chimpanzee Vaccinated with Recombinant Surface Antigen
   II.B.1) Biopanning polyclonal IgG antibodies with the random phage-display library.................................................. 63
   II.B.2) Sequence comparison and alignment of pIII-fusion peptides with the primary sequence of HBsAg................................. 69
   II.B.3) Phage that preferentially recognise immune serum IgG share common sequence motifs and match in the same area of HBsAg. 77
   II.B.4) Binding studies with individual phage that preferentially recognise immune serum IgG.................................................. 83
   II.B.5) Competing binding of phage 6 and 7 to immune serum IgG with synthetic peptides..................................................... 87

CHAPTER III: DISCUSSION........................................................................ 91

CHAPTER IV: MATERIALS AND METHODS

IV.A) Materials and Equipment
   IV.A.1) Suppliers of reagents and laboratory equipment..................... 103
   IV.A.2) Bacteriological material and solutions...................................... 104
   IV.A.3) Standard solutions and buffers................................................ 105
   IV.A.4) Antibodies and anti-sera.......................................................... 106
   IV.A.5) Synthetic peptides.................................................................. 106
IV.B) Methods

IV.B.1) Protein purification and analysis........................................................................... 107

- Purification of antibodies (IgG) from sera
- Bradford assay for the determination of protein concentrations
- Biotinylation of antibody molecules
- SDS-polyacrylamide gel electrophoresis (PAGE)

IV.B.2) Phage amplification, purification and analysis...................................................... 109

- Large-scale preparation of phage particles
- Titration of phage as plaque forming unit (p.f.u.)
- Titration of phage as transforming units (t.u.)
- Preparation of single-stranded phage DNA for sequence analysis
- Sequence analysis of phage DNA
- Immuno dot blots with anti-fd phage rabbit serum

IV.B.3) Biopanning procedures...................................................................................... 112

- Amplification of the random hexapeptide phage-display library
- Biopanning with polyclonal serum IgG using petri dishes as solid support
- Biopanning with polyclonal serum IgG using paramagnetic Dynabeads® as solid support
- Biopanning with monoclonal antibodies using microwells as solid support
- Amplification and titration of phage contained in eluates of biopanning experiments (output phage)

IV.B.4) Analysis of individual phage................................................................................ 115

- Phage binding assay on solid phase (input-output relationship)
- Phage binding assay over time (time course experiments)
- Determination of relative dissociation constants ($K_D^{rel}$) for the interaction of phage with monoclonal antibodies in solution
- Indirect phage-sandwich enzyme-linked immunosorbent assay (ELISA) using anti-fd phage antibodies
- Competition experiments by indirect phage-sandwich ELISA using synthetic peptides
IV.B.5) Immunisation of rabbits with phage particles .............................. 118

CHAPTER V: REFERENCES ............................................................................. 119

CHAPTER VI: APPENDIX ............................................................................ 143


ABSTRACT

The aim of this project was to employ phage-display libraries to investigate epitopes recognised by antibodies directed against the hepatitis B virus surface antigen (HBsAg).

The experiments were based upon a library comprising a pool of filamentous phage (fusion-phage) which displayed random hexapeptides on the surface of the particles. Antibodies of monoclonal and polyclonal origin were used for selection of phage that bound the antibodies via their exposed peptides.

In the case of the monoclonal antibody, the phage ligands selected showed highly related amino acid sequences with a few distinctively conserved residues. These residues could be aligned with the primary sequence of the PreS1 domain of HBsAg thus precisely locating the epitope recognised by this monoclonal antibody. A quantitative assay, adapted specifically for use with fusion-phage particles, revealed binding constants in the nM range for the interaction of the fusion-phage with the antibody. Comparison of the binding constants of phage with divergent amino acid sequences and competition binding studies with synthetic peptides defined the epitope recognised by this antibody with high-resolution.

A similar approach was applied to polyclonal serum IgG from a chimpanzee that had been vaccinated with HBsAg. The complex situation of the polyclonal background made interpretation of the isolated phage ligands more difficult. However, consensus motifs were found among the phage analysed which could be aligned with short stretches of the primary sequence of HBsAg. Investigation of these consensus motifs and a screening procedure that selected for phage preferentially recognising immune serum IgG led to two related families of fusion-phage that displayed a sequence motif that could be aligned with residues 117-122 of HBsAg. Competition experiments with peptides representing this motif provided, in association with the other results, support for the argument that this area is recognised by this polyclonal serum raised against HBsAg.

It has been demonstrated that the phage-display approach can provide high-resolution epitope mapping for monoclonal antibodies. Quantitative values for the interaction in solution between fusion-phage representing this epitope and the antibodies have been
obtained readily by means of a simple assay based upon phage titre. Further, even within the complex context of a polyclonal serum individual epitopes or at least elements of them could be identified, thus indicating immunologically important areas of this antigen.
ACKNOWLEDGEMENTS

Foremost, I would like to thank Prof. Sir Ken Murray for his support and guidance throughout the whole project and his valuable advice on many issues connected with a "life in science".

I am also most grateful to Fiona Gray, Sandra Bruce, Heather Davidson, Dr. Ai-Li Shiau, Wen Siang Tan, Suling Zhao, Tina Dolle, Dr. Fiona Stewart and Jeanette Ferguson for their advice and support and for creating such a pleasant atmosphere to work in. Particularly, I would like to thank Dr. Mike Dyson for his patience while listening to my problems, for always providing sound and helpful advice and for having a good laugh on many occasions, also outside of work. There are many other people who I had a great time with and who supported me in one way or another particularly Sabine, Dieter, Thomas, Jim, Stefan, Marie-Christine and of course, Lisa.

Many thanks go to the people in the media room and the washing up kitchen and to everybody else in the ICMB who made these years so enjoyable.

I am thankful to Glynis Leadbetter for immunising the rabbits and providing us with valuable sera, to Prof. George Smith for providing us with the fuse5 library, Drs. Gerlich and Heermann for providing us with MA18/7 and to Prof. Richard Ambler and Margaret Daniels for carrying out the amino acid analysis on the synthetic peptides. I thank Dr. David Melton for being my second supervisor.

Finally, I would like to thank the Darwin Trust of Edinburgh for financing the official three year period and the Roseanne Campbell Trust for Hepatitis Research for supporting me for the remaining time.
**List of abbreviations (in alphabetical order):**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>Ab</td>
<td>antibody</td>
</tr>
<tr>
<td>Ag</td>
<td>antigen</td>
</tr>
<tr>
<td>AIDS</td>
<td>acquired immunodeficiency syndrome</td>
</tr>
<tr>
<td>APC</td>
<td>antigen presenting cell</td>
</tr>
<tr>
<td>BCIP</td>
<td>bromo-chloro-indolyl phosphate</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>C</td>
<td>Celsius</td>
</tr>
<tr>
<td>CDR</td>
<td>complementarity determining region</td>
</tr>
<tr>
<td>Ci</td>
<td>Curie</td>
</tr>
<tr>
<td>dATP</td>
<td>deoxyadenosine triphosphate</td>
</tr>
<tr>
<td>ddATP</td>
<td>dideoxy adenosine triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>deoxycytidine triphosphate</td>
</tr>
<tr>
<td>ddCTP</td>
<td>dideoxy cytidine triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>deoxyguanosine triphosphate</td>
</tr>
<tr>
<td>ddGTP</td>
<td>dideoxy guanosine triphosphate</td>
</tr>
<tr>
<td>DMF</td>
<td>dimethylformamide</td>
</tr>
<tr>
<td>DNA</td>
<td>desoxyribonucleic acid</td>
</tr>
<tr>
<td>dTTP</td>
<td>deoxythymidine triphosphate</td>
</tr>
<tr>
<td>ddTTP</td>
<td>dideoxy thymidine triphosphate</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetraacetic acid</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunoabsorbent assay</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>F&lt;sub&gt;ab&lt;/sub&gt;</td>
<td>monovalent Ag-binding fragment of immunoglobulin</td>
</tr>
<tr>
<td>F&lt;sub&gt;c&lt;/sub&gt;</td>
<td>constant sequence fragment of immunoglobulin</td>
</tr>
<tr>
<td>scFv</td>
<td>variable sequence fragments of heavy and light chains linked by spacer amino acids (&quot;single chain&quot;)</td>
</tr>
<tr>
<td>HBeAg</td>
<td>hepatitis B e-antigen</td>
</tr>
<tr>
<td>HBsAg</td>
<td>hepatitis B surface antigen</td>
</tr>
<tr>
<td>HBV</td>
<td>hepatitis B virus</td>
</tr>
<tr>
<td>HBxAg</td>
<td>hepatitis B x-antigen</td>
</tr>
<tr>
<td>HCC</td>
<td>hepatocellular carcinoma</td>
</tr>
<tr>
<td>hGH</td>
<td>human growth hormone</td>
</tr>
<tr>
<td>HIV</td>
<td>human immunodeficiency virus</td>
</tr>
</tbody>
</table>
IgE immunoglobulin E class
IgG immunoglobulin G class
IL interleukin
kb kilo bases
kD kilo Dalton
$K_d$ dissociation constant
L Luria
LB Luria broth
M Molar
MA18/7 monoclonal antibody against PreS1 domain of HBsAg
MAb monoclonal antibody
MHC major histocompatibility complex
NBT nitroblue tetrazolium
NMR nuclear magnetic resonance spectroscopy
OD optical density
ORF open reading frame
PAGE polyacrylamide gel electrophoresis
PBS phosphate buffered saline
PEG polyethylene glycol
p.f.u. plaque forming units
pNPP p-nitrophenyl phosphate
PreS1 domain of HBsAg (L) comprising 108-119 aa
PreS2 domain of HBsAg (L and M) comprising 55 aa
RF replicative form
RNA ribonucleic acid
rpm revolutions per minute
SDS sodium dodecyl sulphate
ss single stranded
SV 40 Simian virus 40
TB terrific broth
TBE tris-buffered EDTA solution
TBS tris-buffered saline
TEMED tetramethyl ethylenediamine
t.u. transforming units
U unit
v/v volume/volume
w/v weight/volume
**Single letter code (amino acids):**

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Single Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>A</td>
</tr>
<tr>
<td>Arginine</td>
<td>R</td>
</tr>
<tr>
<td>Asparagine</td>
<td>N</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>D</td>
</tr>
<tr>
<td>Cysteine</td>
<td>C</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>E</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Q</td>
</tr>
<tr>
<td>Glycine</td>
<td>G</td>
</tr>
<tr>
<td>Histidine</td>
<td>H</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>I</td>
</tr>
<tr>
<td>Leucine</td>
<td>L</td>
</tr>
<tr>
<td>Lysine</td>
<td>K</td>
</tr>
<tr>
<td>Methionine</td>
<td>M</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>F</td>
</tr>
<tr>
<td>Proline</td>
<td>P</td>
</tr>
<tr>
<td>Serine</td>
<td>S</td>
</tr>
<tr>
<td>Threonine</td>
<td>T</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>W</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Y</td>
</tr>
<tr>
<td>Valine</td>
<td>V</td>
</tr>
</tbody>
</table>

**Single letter code (nucleotide bases):**

<table>
<thead>
<tr>
<th>Nucleotide Base</th>
<th>Single Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenine</td>
<td>A</td>
</tr>
<tr>
<td>Cytosine</td>
<td>C</td>
</tr>
<tr>
<td>Guanine</td>
<td>G</td>
</tr>
<tr>
<td>Thymine</td>
<td>T</td>
</tr>
<tr>
<td>Purine</td>
<td>R</td>
</tr>
<tr>
<td>Pyrimidine</td>
<td>Y</td>
</tr>
<tr>
<td>G or T</td>
<td>K</td>
</tr>
<tr>
<td>Nonspecific base</td>
<td>N</td>
</tr>
</tbody>
</table>
DECLARATION

I hereby declare that the work presented in this thesis is the result of my own independent investigation unless otherwise stated and that this thesis has been composed by myself. The experiments presented were devised in collaboration with my supervisor Prof. Sir Kenneth Murray.

The work has not been or is not being concurrently submitted in candidature for any other degree.

Edinburgh, July 1995

(Volker Germaschewski)
CHAPTER I: INTRODUCTION
I. Introduction

I.A) Fusion-Phage and Peptide Libraries: Tools for Exploring Molecular Interactions

I.A.1) The filamentous phage fd

Fd is a filamentous bacteriophage very similar to the well-studied M13 bacteriophage. Their genomes consist of single-stranded DNA molecules and M13 in particular has been extensively used as a vector for molecular cloning since the late seventies (Messing et al., 1977). The reasons for its suitability are first, these bacteriophage do not kill their host cells after infection i.e. they are not lytic. Virions are exported from the host cells by extruding the circular plus strand, the one which is packaged into the virions, through the cell membrane where it associates with phage capsid proteins that were integral parts of the host membrane prior to assembly. Second, these phage allow the packaging of DNA molecules that are longer than the unit length of the phage genome (Messing, 1981) and third, phage DNA which has been subjected to in vitro manipulation can easily be reintroduced into *Escherichia coli* cells by standard transfection methods.

The principal structural component of the phage particle is pVIII, the product of gene VIII which is present in almost 2700 copies per virion. pVIII has a molecular weight of 5.2 kD and consists of only 50 amino acids. Other minor coat proteins are the products of genes III, VI, VII and IX (pIII, pVI, pVII and pIX, respectively). pIII and pVI and pVII are present in about five copies per phage and pIX in 10 (Lin et al., 1980; Simons et al., 1981). The other gene products pI, pII, pIV and pV are non-structural and provide functions for phage proliferation and assembly.

The N-terminal domain of the coat protein pIII is exposed on the surface at one tip of the phage particle and is essential for the infection process. pIII interacts with the F-pili of susceptible *E.coli* cells to initiate infection (Armstrong et al., 1981). Its C-terminal domain is buried in the capsid and necessary for correct assembly of normal phage particles (Crissman and Smith, 1984).

Following uptake of the particle and unpacking, the single stranded DNA molecule present in the virion is converted into the double stranded RF (replicative form) DNA form by host cell polymerase activities. Replication of RF DNA serves two purposes.
First, it provides the sense strand for transcription of the viral gene products and second, as the name might imply, it generates templates for further synthesis of progeny DNA (plus strands). This synthesis is thereby performed in a continuous fashion around the circular minus strand template. The resulting multimeric copies are then cleaved by the gene II product in order to separate the unit length progeny DNA molecules from each other.

RF DNA accumulates to 100-200 copies per host cell before a regulatory phage-derived protein, the gene V product, reaches levels that prevent further accumulation of RF DNA and promotes export of progeny single strands by interacting with newly synthesised plus strand molecules. This ssDNA-protein complex then migrates to the membrane where it becomes engulfed with membrane associated coat proteins and is extruded into the periplasmic space. However, when phage morphogenesis is impaired but phage genome replication still continues, as it is for example in the case of mutants defective for correct assembly, the host cell is finally killed by the accumulation of phage DNA and phage gene products which cannot be packaged and exported. This phenomenon is known as cell killing (Pratt et al., 1966).

RF DNA can easily be prepared from cultures of infected *E. coli* cells in a similar way to plasmid DNA, but usually with lower yields due to the relatively low copy number present per cell. For sequencing purposes it is also possible to prepare ssDNA from phage particles contained in the supernatant of cultures of infected cells.

**I.A.2) Fusion phage: Phage that display foreign peptides or protein domains**

The term "fusion phage" was first introduced by George Smith in 1985 for filamentous phage particles that had foreign proteins fused to their pIII coat proteins and therefore displayed these moieties on the virion surface. In this pioneering work Smith constructed fusion-phage which displayed fragments of the EcoRI restriction endonuclease. Using petri dishes with adsorbed anti-EcoRI antibodies he was able to demonstrate the feasibility of an affinity purification that achieved the isolation of these fusion-phage from a large background of unspecific phage.
I. Introduction

Since their introduction fusion-phage vectors have been modified in various ways to optimise the system or to tailor it according to specific applications. One important modification was the introduction of an antibiotic resistance gene against tetracycline (Parmley and Smith, 1988) as had been done earlier with the non-fusion phage vectors (Zacher et al., 1980). The propagation of these altered fusion-phage does not depend on continuous reinfection of new E. coli cells but can be propagated in selective medium in which the DNA genome is passed on like a plasmid. A single infection therefore leads to a visible clone even when infectivity of the phage is reduced, as is often the case for fusion-phage involving the infectivity-sensitive pIII coat protein. Since the tetracycline gene was inserted into the origin of minus strand replication these phage were also partially defective for minus strand synthesis. This avoids the accumulation of progeny ssDNA in bacteria infected with virions which are impaired in assembly or export and as a consequence cell killing is prevented.

It also became apparent that inserting sequences at the 5' end of gene III and consequently displaying the foreign peptide at the very N-terminus of pIII was preferable to internal sites used previously (Parmley and Smith, 1988), probably because the pIII-fusion was more surface-accessible and the negative effects on infectivity were reduced.

After the idea of fusion phage was established, other groups picked up the principle and generated fusion-phage by linking the foreign peptide to pVIII, the major coat protein which is present in about three thousand copies per particle. This allows a higher degree of multivalency and also circumvents a potential infectivity problem since pVIII is not involved in the infection process (Felici et al., 1991; Kang et al., 1991; Markland et al., 1991). Recently, fusions to pVI, another minor coat protein have been reported (Jesper et al., 1995). The main difference of this system is that, in contrast to pIII, the C-terminus of pVI is exposed on the phage surface which allows foreign proteins to be fused with their N-termini leaving the C-termini of the fusions free to interact with specific ligands. Further, fusions with not naturally secreted proteins or domains seem to be more efficiently displayed than in the current systems which could be of particular importance for the display of cDNA libraries.

At least in the case of pVIII fusions the inserts seem to be somewhat limited in size since larger inserts appear to interfere with assembly of the phage particle. This problem can be overcome by using so-called phagemid vectors where wild type
pVIII is provided by an external helper phage (Felici et al., 1991; Kang et al., 1991). Competition between the wild type pVIII and the manipulated pVIII (from the phagemid) during assembly of new phage particles in the cell will result in incorporation of a relatively low number of mutant pVIII molecules. This usually assures correct assembly of particles and increases the tolerance for larger fragments.

This system can also be applied to the p111-fusion phage resulting in incorporation of fewer than five copies per virion. Since wild type p111 from the helper phage is in vast excess a near monovalency of the displayed peptide or domain can be achieved which may have advantages for certain applications (Bass et al., 1990; Lowman et al., 1991; Marks et al., 1991; Hoogenboom et al., 1991). This way a range of whole proteins or domains have been displayed on filamentous phage which would otherwise have caused serious problems with infectivity (p111-fusions) or assembly (pVIII-fusions). They include enzymes such as alkaline phosphatase (McCafferty et al., 1992), trypsin (Corey et al., 1993) and \textit{Staphylococcus} nuclease (Chiswell and McCafferty, 1992), growth factors and cytokines such as human growth hormone hGH (Bass et al., 1990) and IL-3 (Gram et al., 1993), and receptors like the IgE-binding extracellular domain of FcεRIα (Robertson et al., 1993). The display of antibody fragments (Fab) or so called single chain Fv fragments (scFv) has been extremely successful since they show full binding activity and can be used essentially like free antibodies (McCafferty, 1990; Hoogenboom et al., 1991).

Most of the displayed proteins and domains are extracellular proteins and naturally secreted and therefore do not cause problems when exported from the host cell for phage assembly. The export problem seems to be a more important factor than the size of the fused polypeptide which does not seem to be limiting \textit{per se}. However, the cell internal DNA-binding transcriptional regulators Zif268 (Rebar and Pabo, 1994; Wu et al., 1995) and Jun/Fos (Crameri et al., 1993) have also recently been displayed on phage.

Some phagemid vectors contain amber stop codons between the foreign gene and gene III which allow the production of soluble fusion fragments in non-suppressor strains (Hoogenboom et al., 1991; Lowman et al., 1991; Robertson et al., 1993). This can be helpful if the fused domain is to be used in subsequent studies.
I. Introduction

I.A.3) Phage-display libraries

After it had been demonstrated that fusion-phage can be isolated from large backgrounds of unspecific phage on the basis of their displayed protein moiety the next breakthrough was to display random peptides and create large pools of such fusion-phage in order to search for any phage that bind specifically to molecules of interest. Since these libraries were initially used in connection with antibodies to perform epitope mapping studies the classical term "epitope library" was introduced (Scott and Smith, 1990) but strictly speaking this term does not apply if other interactions are studied. Therefore the more general term "phage-display library" is preferable and widely used now.

The inspiration for creating libraries on a random basis came from earlier work by Geysen et al. (1986) who developed a synthetic peptide library on plastic pins to search for specific ligands to antibodies. They isolated peptides mimicking a discontinuous epitope and consequently called their library a "mimotope library". Despite this obvious success, however, biological libraries have distinct advantages over synthetic libraries. First, sequence analysis of the ligands is much easier since the selection for binding activity also means selecting for the gene encoding it because the coding information and the protein displayed are physically linked (in the phage particle). Second, biological libraries are cheaper to make and can be propagated indefinitely. Third, multiple rounds of selection followed by amplification enhance the chances of finding very rare ligands that might have existed only once in the original library. On the other hand synthetic libraries offer the advantage of not having a biological bias towards certain sequences since there is no in vivo selection pressure (Clackson and Wells, 1994).

Random phage-display libraries can be based upon either pIII- or pVIII fusions and usually display between 6 and 15 amino acids. For reasons explained above pVIII-libraries are usually phagemid libraries and therefore require superinfection with helper phage (Felici et al., 1991). Both libraries offer distinct advantages depending on the application. pVIII-libraries have the point of a higher degree of multivalency for them which can be helpful if very weak interactions are studied that would otherwise be lost. In the same way the pIII-library can have advantages if stronger interactions are pursued since too many interactions per virion could result in too tight binding and failure to recover the bound phage during the selection procedure.
I. Introduction

Both fd and the close relative M13 have been used successfully for the construction of combinatorial display libraries.

Ideally, a random library covers all possible amino acid sequences using all 20 amino acids. For a random hexapeptide library this means that there have to be at least 20^6 or 6.4 x 10^7 different transformants to have every possible amino acid combination represented in theory. In reality this will of course not be precisely the case due to biological preference for some amino acid sequences and the abortive effects of others. Further, some amino acid sequences are likely to be thermodynamically less favourable than others and therefore less likely to be present. It is also probable that some sequences will occur more than once further reducing the overall complexity. Consequently, even higher numbers of transformants are required, in practice about 2.5 x 10^9 transformants if one would like to achieve 90% confidence for having all possible peptides represented (Hoess, 1993). However, the transformation efficiency of the E. coli host cells limits library sizes to 10^7-10^8 transformants if still manageable amounts of competent cells are to be used. This means that libraries consisting of more than five random residues will be underrepresented to various extents (Hoess, 1993). Although a complete representation might be impossible in many cases it has been proposed to increase the number of possible short peptides by cloning larger random peptides. The idea is that, for example, a 38- residue random peptide offers 33 "windows" for unique hexapeptides thus significantly reducing the number of independent transformants required for a hexapeptide library (Kay et al., 1993; McLafferty et al., 1993).

Several cloning strategies have been developed for the construction of phage-display libraries, but essentially all use chemically synthesised random oligonucleotides which are converted into double stranded fragments by annealing or polymerase chain reaction (PCR) and ligated into the linearised vector RF DNA while maintaining the reading frame for the phage fusion protein. Often the third base of the triplets coding for the random amino acids are restricted to G or T (Scott and Smith, 1990). This still provides all 20 amino acids but reduces the degeneracy of triplets used to code for a single amino acid since several amino acids are represented by more codons than others in the genetic code. Furthermore two of the three stop codons are avoided which would lead to non-functional nonsense fusions. After ligation E. coli cells are transformed with the ligation mixture, usually by electroporation. To increase library size several independent transformations are
I. Introduction

performed and the transformants pooled. The library is then amplified by growing up the transformed host cells in large scale and then purified by precipitating the phage particles. Aliquots of the transformed cells are plated out prior to amplification to check transformation efficiency and provide colonies for sequence analysis which is an indicator of the quality of the library. For the integrity of the library it is important that there is no preference for certain amino acids or positions within the displayed peptide.

I.A.4) The fuse5 hexapeptide library

The fuse5 library is an epitope library of the first generation of libraries and is probably one of the most commonly used libraries at present mainly attributable to the generous distribution policy of Prof. George Smith.

The library consists of random hexapeptides fused to the N-terminus of the minor coat protein pIII of a modified fd-tet (Zacher et al., 1980) based phage vector (fuse5; Scott and Smith, 1990). It is based on 2x10^8 independent transformants in the original primary library which represent around 69% of all possible hexapeptides (6.4x10^7; Scott and Smith, 1990). However, in the amplified library, which is the one distributed, each original clone is represented by approximately 500 subclones. Amplification of this amplified stock further reduced the diversity of this library in a given volume due to expansion of subclones, but was obviously needed in order to have enough material to perform several screening experiments.

Titration of the amplified culture showed that it contained 8x10^10 infected cells in total. In relation to the input phage (10^{12} particles) the infectivity was around 8%. After purification of the phage particles the library had a titre of 1x10^{13} t.u./ml and 1.5x 10^{14} physical particles/ml.

The original primary library was constructed by using PCR for the conversion of the single stranded degenerate oligonucleotide into a clonable double stranded fragment. The triplets coding for the random amino acids were NNK with K standing for either G or T which, as previously mentioned, excludes two of the three stop codons and reduces the number of codons leaving 32 possible triplets. The annealed PCR products were cut with the restriction endonuclease Bgl I in order to generate sticky
I. Introduction

ends. Through this method theoretically about 12000 of the 6.4x10^7 possible hexapeptides are lost due to internal Bgl I restriction sites created within the random sequence. Alternative cloning procedures for generating random phage libraries use universally compatible inosine nucleotides (Devlin et al., 1990) or a single stranded gap opposite the degenerate sequence (Cwirla et al., 1990) to avoid this problem of losing random sequences.

The fragments were then ligated into compatible Sfi I restricted fuse5 vector DNA molecules. E.coli host cells were transformed with aliquots of the ligation product in 64 independent electroporations in order to increase the number of total transformants generated. The transformed cells were pooled and amplified. Phage particles can be purified from the supernatant of infected cultures by precipitation.

![Phage library diagram]

**Figure I.1: The general principle of "biopanning"**
I. Introduction

I.A.5) "Biopanning" and factors that influence selection

The general principle for screening a library, often referred to as "biopanning", is to attach the molecule of interest to a solid support and then present the immobilised molecules to the phage pool or, alternatively, to present the molecules first and then capture them on a solid support (Fig. I.1). Phage that display peptides that bind to the molecules presented will adhere to the solid support. Subsequent washing procedures remove all unbound phage. Phage that stay bound are then recovered by elution, usually with low pH buffers. These eluted phage will be enriched for phage that bind the molecule of interest and the whole eluate is then amplified creating many subclones of each phage in this selected population. Several rounds of binding-washing-eluting-amplifying are performed to select for very rare sequences and obtain pools consisting to a large degree, if not exclusively, of phage that bind the molecule of interest. Methodical aspects have proven to be a crucial and often unpredictable factor in biopanning experiments and recent modifications of the method focus on the following parameters:

- **Solid support**
  The streptavidin-biotin link is often used to immobilise the substrate. This is largely due to the fact that proteins can be easily biotinylated and the bond is virtually irreversible. This way molecules can be linked to plastic dishes (Scott and Smith, 1990), microtiter wells (Barret et al., 1992; Blond-Elguindi et al., 1993) and beads (Bass et al., 1990; Hawkins et al., 1992). The molecules are either first incubated with the library in solution and then captured on streptavidin-coated matrices or presented to the library after being immobilised. It is not even necessary to biotinylate the molecule of interest since the capturing can be achieved via biotinylated secondary antibodies (Scott and Smith, 1990). Even simpler is adsorption on columns (McCafferty et al., 1990; Dedman et al., 1993) or unspecific adsorption on plastic (Jellis et al., 1993; Ames et al., 1994; Germaschewski and Murray, 1995) and nitrocellulose (Dyson and Murray, 1995) which all have been used for attachment. For some purposes it might be important to present the molecule in a directed fashion so that certain domains become better accessible.
1. Introduction

- **Incubation step**
  Most incubations are performed for 4 h or overnight so that equilibrium is approached. Temperature has effects on binding and incubation at room temperature will probably select for tighter ligands. Some cases are reported where libraries were pre-incubated with other substrates prior to biopanning in order to deplete the library of certain phage. This can be important if the molecules of interest are present in a mixture or contain unwanted impurities (Dybwad et al., 1993; Ames et al., 1994). In any case the selection procedure will not distinguish between different molecules and potentially provides ligands to all molecules that are presented. Blocking of all unspecific binding sites therefore is essential, but despite this, problems can occur due to the enormous sensitivity of the combined selection and amplification procedure (Christian et al., 1992; G.P. Smith, personal communication; see also Results section, p.68). The concentration of the molecules presented to the library is also crucial and can be adjusted in order to promote competition between different phage resulting in selection for ligands with higher affinities (Scott and Smith, 1990). This can easily be achieved if the molecules are incubated with the library in solution but it has also been found to be valid for solid matrices in which case lower coating concentrations will to some degree eliminate interactions of one phage particle with several ligand molecules at the same time due to larger distances between binding sites (Barret et al., 1992).

- **Washing procedure**
  The stringency of washing is crucial for the selection procedure and determines the affinity threshold of the ligands selected for (Barret et al., 1992; Hawkins et al., 1992). It also has to be considered that, although to different extents, pIII- and pVIII libraries are both multivalent systems and that this may affect the avidity of a potential ligand. The avidity being the combination of the affinity of a ligand, specified by its chemical properties, and the number of sites involved in the interaction, even low affinity ligands may have high avidities due to multiple interactions of one phage with one or several molecules. Phage vectors that allow multiple interactions are therefore thought to bias towards the selection of low to moderate affinity phage with strong affinity phage being too tightly bound to be recovered. To increase stringency pH or salt concentrations can be varied. Sequential washes with increasing stringency or different specificity can be applied and phage contained in the different washes separately analysed (O'Neil
et al., 1992; Koivunen et al., 1993; Dyson and Murray, 1995). Temperature is also an important factor and washes with cold buffers in the cold are likely to be less stringent.

- **Elution process**

  Most commonly low pH buffers are used since phage are relatively stable and pH as low as pH 2 can be used. Natural ligands of the molecule used for biopanning have been used to specifically select for phage that are only released on binding of the ligand (O'Neil et al., 1992; Dedman et al., 1993; Blond-Elguindi et al., 1993) or displace the phage from the same binding site (Koivunen et al., 1993).

I.A.6) Applications for phage-display libraries

As mentioned before phage-display libraries have, until recently, been used mainly for epitope mapping purposes. Many examples of this have been reported, but the number of completely newly defined epitopes that can be located precisely on the antigen remains low (Stephen and Lane, 1992; Yayon et al., 1993; Böttger and Lane, 1994). Most of the successes are confined to continuous epitopes where the contributing amino acids follow each other in the same sequence as in the primary sequence of the antigen against which the antibodies were raised thereby making identification of the epitope straight forward (Scott and Smith, 1990; Felici et al., 1991; Stephen and Lane, 1992; Jellis et al., 1993; Grosso et al., 1993; Yayon et al., 1993; Böttger and Lane, 1994; Miceli et al., 1994; Böttger et al., 1995; Germaschewski and Murray, 1995).

One of the key advantages of phage-display libraries is that no structural information about the protein of interest is required prior to biopanning. Since the random population will theoretically cover all possible binding sites it was assumed that phage might also mimic conformational epitopes or binding sites consisting of amino acids that are far apart in the primary sequence of the protein but in close proximity in the folded form. In many cases phage binding such antibodies could be isolated but it was difficult or impossible to identify the corresponding residues, i.e. the epitope, in the antigen unless the structure was known. As a result, numerous attempts with antibodies recognising discontinuous epitopes have been unsuccessful in terms of epitope mapping purposes.
I. Introduction

One example where structural information was available, however, has been reported and here the authors isolated phage with a weak consensus motif that bound a monoclonal antibody to human ferritin. The conserved residues could be located in the structural model and point mutation of these residues abolished recognition of the antibody (Luzzago et al., 1993). Another similar experiment by the same group was less successful however, since no strong consensus sequence was found and structural information was not available, although the selected phage again recognised the specific antibody (Felici et al., 1993).

In most cases it will be possible to isolate specific ligands with the biopanning procedure, but the sequences isolated will often bear little or no sequence similarity with the natural ligand (e.g. the antigen in the case of antibodies being used for screening). These so called mimotopes are not useful for epitope mapping purposes since for this application the exact identity of the amino acids is desired (Cwirla et al., 1990; Christian et al., 1992; Felici et al., 1993; Keller et al., 1993). In many other cases however, mimotopes proved extremely valuable due to their ability to specifically recognise a ligand about which very little prior information was available. These mimotopes can have importance if they recognise molecules of biological or medical interest and bear the potential of being lead compounds for the development of drugs (Ames et al., 1994; Christian et al., 1992; O'Neil et al., 1992; Balass et al., 1993; Keller et al., 1993; Koivunen et al., 1993; Dyson and Murray, 1995). It is also possible to identify sequences with higher binding constants than the ones of naturally occurring ligands (Lowman and Wells, 1991; Barret et al., 1992; O'Neil et al., 1992) and the method is capable of distinguishing subtle single amino acids changes (O'Neil et al., 1992; Koivunen et al., 1993). In these cases binding affinity and specificity are of particular interest, and not necessarily the amino acid sequence. For example Balass et al. (1993) isolated phage that recognised a monoclonal antibody directed against the chicken acetylcholine receptor. Although there was no sequence similarity with any region of the receptor molecule free peptides of the same sequence inhibited the interaction of the antibody with the receptor in vivo.

As these examples illustrate, applications for phage-display have recently moved away from classic epitope mapping and all kinds of interactions are studied involving many other protein-protein interactions and protein-nucleic acid interactions (Reba and Pabo, 1994; Wu et al., 1995).
1. Introduction

Libraries themselves have also undergone changes and recently so called constrained libraries have been successfully used in some cases (O'Neil et al., 1992; McLafferty et al., 1993; Luzzago et al., 1993). It is assumed that in many cases peptides that are displayed in a constrained conformation may be recognised better by ligands since target sequences often also occur in a more or less constrained environment (Koivunen et al., 1993). However, the range of restricted conformations engineered during library construction might not include the conformation suitable for interaction with the molecule, in which case it may be impossible to select for a ligand with this specific library. Conformational constraints often result from cysteine residues that flank the random peptide so that the peptide forms a loop on the surface of the phage by intramolecular disulphide bridge formation.

It is worthwhile mentioning alternative recombinant libraries which have been developed from the idea of the phage-display system and sometimes rely on different expression or display vehicles. Libraries of displayed antibody fragments [F\textsubscript{ab} or scF\textsubscript{v}(single chain)], generated by cloning genes for the heavy and light chains of the v-region into bacteriophage vectors have been mentioned already and have gained considerable importance since they can be prepared from cells of unimmunised animals or humans and offer the prospect of selecting monoclonal antibodies specific for virtually any target molecule including, self-antigens (Griffith et al., 1993) or whole cell surfaces (Marks et al., 1993). Further, it has been demonstrated that the affinity maturation process that usually occurs during an immune response in vivo to generate high affinity ligands by somatic mutation can be mimicked in this in vitro system by introducing random mutations and subsequent screening. (Gram et al., 1992; Hawkins et al., 1992). Also, by mixing different heavy and light chain libraries, chain combinations can be shuffled creating new specificities and greater affinities (Marks et al., 1992).

Synthetic peptide libraries have also been mentioned earlier and remain important through their advantages already discussed and the possibility of including residues that do not occur in nature, but have in many instances been replaced by the cheaper biological libraries. In fact the identification of reactive peptides with these libraries is often a mass screen rather than a selection procedure (Geysen et al., 1984).

Peptides have also been expressed as fusions to surface proteins of E.coli cells (Little et al., 1993; Lu et al., 1995) or even directly attached to the plasmids that encode
them (peptide-on-plasmid approach; Cull et al., 1992). Some of these techniques offer advantages in very specific situations and were often developed for these purposes and will probably not gain universal success like the bacteriophage approach or the synthetic libraries. An interesting variation of the pIII-fusion phage has been described recently where the coat-anchored and the exposed infection-mediating domain of pIII are split into two independent molecules and infectivity is only restored in cases where the two domains interact via protein-protein interactions. If one component of the interface is kept constant while the other one is variant like a library similar screening tasks can be carried out (Gramatikoff et al., 1994).
I.B) Antibody-Antigen Interactions

I.B.1) Antibodies: Structure and general aspects

Antibodies consist of two identical light (25 kD) and two identical heavy chains (55 kD) forming the familiar Y-shaped bivalent structure by interchain disulphide crosslinking. The N-terminal domains of the heavy and light chains (V domains) each contain three hypervariable loops (CDRs, complementarity determining regions) which directly contact the antigen and therefore determine the specificity of the antibody. The six loops that form the antigen binding site (paratope) are in close proximity to each other in the folded V domains which provide a scaffold for correct presentation. The size and the shape of the antigen binding site can be very different depending on the amino acid sequence although the β-sheet framework in which the loops are presented is very conserved. At least four of the six loops, termed L1-3 and H1-3, respectively, are involved in any given interaction (Wilson and Stanfield, 1993) and heavy chains are believed to contribute slightly more to the interaction (Kabat and Wu, 1991). This is supported by a larger promiscuity of the heavy chains for many different light chains observed in chain shuffling experiments (Collet et al., 1992; Zebedee et al., 1992).

The C-terminal domains of the heavy chains (C domains) form a conserved framework and determine the subclass of the antibody molecule. This Fc region also mediates the biological effector functions on binding to various receptors of the immune system. For the IgG class there are three basic receptor types termed FcyRI-III and two receptors of the complement system (CR1 and CR3) which are located on different cell types mediating different functions.

The CDR loops are encoded by the v-genes (for variable) which provide a diversity of about $10^7$ which is then greatly increased by somatic recombination and hypermutation of the CDR areas generating in excess of $10^{10}$ specificities. Hypermutation also accounts for an effect called affinity maturation which leads to mutants with improved affinity during secondary responses.

Antibody fragments are often used for binding studies particularly the Fab fragment (for antibody binding) consisting of the truncated heavy and light chain antigen
binding site, but lacking the conserved Fc region (for constant) of the heavy chains. Since each truncated F_{ab} fragment only consists of one antigen binding site the fragments are monovalent in comparison to bivalent for the complete antibody.

**I.B.2) Antigen recognition: What are epitopes?**

The term "epitope" was established as early as 1960 by Niels Jerne describing an epitope as the area on the surface of a native antigen accessible to the immune system. Areas that were only recognised after denaturation or breakage were termed "cryptotopes" since they were cryptic in the native form.

Epitopes are classically divided into continuous, or linear, and discontinuous, or conformational, epitopes. Linear epitopes are supposed to involve a contiguous stretch of amino acids of the primary sequence of the antigen. It is therefore likely that linear epitopes are recognised on the native and denatured form of the antigen. "Epitopes" recognised by antibodies raised against denatured antigens are mostly linear and so are T-cell "epitopes" recognised by T-cell receptors since they are located on short fragments of the proteolytic degraded antigen which are presented by cells of the immune system such as B-cells and antigen presenting cells.

Discontinuous epitopes are not represented in the primary sequence as such but assembled from various parts of the amino acid sequence brought together in the folded structure of the native antigen. They are therefore dependent on the conformation of the antigen and usually disrupted on denaturation.

Strictly speaking there are no linear epitopes since the original definition was restricted to native functional epitopes and there is experimental evidence that all epitopes of Ab-Ag complexes studied by X-ray diffraction analysis so far are conformational and involve several parts of the primary sequence (Amit et al., 1986; Sheriff et al., 1987; Padlam et al., 1989; Tulip et al., 1992).

Comprehensive mutational analysis of the amino acid sequence of hGH (human growth hormone) and its effects on recognition by 21 monoclonal antibodies also revealed that all epitopes recognised on hGH were conformational involving at least two areas of the primary sequence and were, although overlapping, recognising
I. Introduction

slightly different epitopes (Jin et al., 1992). However, the authors also argued that although many residues were involved in these "contact epitopes" (i.e. residues making contact at the interface), the functional epitope (i.e. the residues crucial for binding) maybe constituted of only a few of these residues and that certain amino acids dominated this role. In this case an average of three amino acids dominated the functional epitopes and eight residues contributed noticeably to binding with arginine (R) and proline (P) being the amino acids most commonly involved.

Despite these findings it is assumed that linear components can be part of larger more complex conformational determinants. Linear target sequences of antigens have been shown in numerous examples to be sufficient for specific recognition by antibodies. There is also structural data supporting this, for example, a 15mer peptide of the capsid protein VP2 of human rhinovirus representing a linear epitope has been shown to assume a similar conformation to that of the native antigen on binding the antibody (Tormo et al., 1994).

In contrast with more recent findings by X-ray diffraction of crystallised Ab-Ag complexes it was initially thought that the areas of interaction between antibodies and antigens involved only a few residues and were quite rigid binding pockets which gave rise to the lock-and-key theory of antibody interaction (Saul et al., 1978; Satow et al., 1986). Through detailed studies on more Ab-Ag and Ab-peptide complexes during the last few years it became evident though that i) conformational changes occur in the antibody on binding of the antigen (Herron et al., 1991; Rini et al., 1992; induced fit theory), ii) depending on the size of the antigen, the contact areas are quite flat and also larger than anticipated involving 10-25 amino acids (Amit et al., 1986; Sheriff et al., 1987; Padlam et al., 1989; Tulip et al., 1992), iii) numerous Van der Waals and salt bridges stabilise the binding (same references as ii)) and iv) water molecules can play an important role in mediating binding (Fischman et al., 1991; Bhat et al., 1994).

Potentially the whole surface of an antigen is immunogenic, but in many cases immunodominant areas will exist which are highly accessible and often flexible allowing a perfect fit with a wide range of antibody paratopes especially in view of the induced fit theory (Davies and Padlam, 1992). Charged and hydrophobic residues also are more likely to be the target of antibody binding due to the nature of interactions occurring at antibody-antigen interfaces (Jin et al., 1992).
I. Introduction

I.B.3) Epitope mapping

Detailed pictures of the residues involved in antibody-antigen interactions and the precise structures of the molecular interfaces have been provided in the last few years by X-ray diffraction analysis of Ab-Ag crystals and defined epitopes as conformationally defined areas recognised on the surface of the folded protein in the classical sense of the term. However, obtaining useful crystals is not trivial and often elusive. Further, the native conformation of neither the antigen nor the antibody in the crystal can be guaranteed and is difficult to assess when no other structural determination in solution is available for comparison. Structural determination by NMR, which could achieve this goal, is currently restricted to molecules much smaller than antibodies (~25kD or 200 amino acids, depending on labelling techniques and technical equipment) and far smaller than any complexes formed with them.

The term "epitope mapping" is widely used for a range of techniques which all reveal various aspects of a particular interaction in question, but have, for example, very different degrees of resolution. Investigating whether two monoclonal antibodies have overlapping binding sites is also epitope mapping and can be very helpful in the individual case. In some cases the identification of the approximate site of interaction on the surface of the antigen, for example identifying a certain subdomain as the target, can be sufficient, in others the identification of the crucial amino acids involved in binding the antibody is the principal interest.

The following list gives an overview of the most commonly used epitope mapping techniques and their degree of resolution. Other advantages or problems are indicated:

- **Cross-reactivity between monoclonal antibodies against the same antigen**
  As already mentioned this technique is often used if the interest is solely whether two antibodies have overlapping binding sites or not, but when the binding sites are in close proximity steric hindrance can also lead to apparently overlapping binding sites. This method provides a very rough idea about the epitopes recognised.
I. Introduction

• Cross-reactivity of anti-sera with fragments or peptides representing various portions of the antigen (antigenicity)

Anti-sera are tested for reactivity with peptides representing various regions of the antigen. The regions are linear stretches of the primary sequence of the antigen and therefore a true conformational representation of the fragment is not guaranteed and conformational epitopes can be lost. Resolution depends on the size of the peptides used but although shorter pieces improve resolution, they jeopardise the conformational authenticity of the fragment.

• Cross-reactivity of anti-peptide and anti-antigen fragment antibodies with the native antigen (immunogenicity)

Although raising antibodies against peptides is not the major problem the outcome of these experiments cannot be predicted and even a high titre of anti-peptide antibodies does not mean that the antibodies will recognise the native antigen. Conformational integrity of the fragments is as questionable as with the previous method, but nevertheless some successes have been achieved (Geysen et al., 1987; Van Regenmortel et al., 1989).

• Screening antibodies with overlapping synthetic peptides

Methods such as PepScan have been quite successful since they only present peptides to the antibody that are based on the primary sequence of the antigen thus excluding the appearance of so-called mimotopes (Geysen et al., 1984; Hastings et al., 1990; Coursaget et al., 1991; Howard et al., 1991; Sällberg et al., 1990, 1991). Linear target sequences are more likely to be detected since short peptides usually give poor representation of native conformations. Again shorter sequences improve resolution, but the technique does not define the contribution of individual residues within the peptide "window" (i.e. the length of the peptides) used for screening since shifting the window one position basically represents a N-terminal deletion and a C-terminal addition in comparison to the peptide of the position before. Deletion of just one crucial residue may thus abolish binding to a degree that makes it difficult or impossible to assess the contribution of the other neighbouring residues (Coursaget et al., 1991; Howard et al., 1991).
I. Introduction

- **Random peptide library screening**
  These libraries often mimic conformational epitopes and peptides that bind to a wide range of antibodies can be isolated, but since mimotopes are often isolated (Geysen et al., 1986; Cwirla et al., 1990; Christian et al., 1992; Felici et al., 1993; Keller et al., 1993), the application for mapping purposes is limited. However, even if the amino acids that form the conformational epitope displayed on the phage are represented in the primary sequence of the antigen the antibodies were raised against, it is virtually impossible to identify them unless structural information about the antigen is available (Luzzago et al., 1993). Linear parts of antigen determinants are easier to identify and have been successfully determined in many cases revealing new epitopes and specificities.

Libraries can be entirely synthetic (Geysen et al., 1986) or biological, for example presented on bacteriophage (Scott and Smith, 1990).

- **Mutations of the antigen that affect antibody binding**
  Deletions, insertions or point mutations can be introduced in the antigen and the effects on recognition by the antibody investigated. Although this method allows one to study the effects of subtle changes in the context of the whole protein it usually requires some prior information that highlights hot spots for mutational analysis. Further, it is difficult to establish whether the effects correlated with the mutation occur because the mutated residue is a contact residue with the paratope or whether the overall conformation of the antigen has been changed so dramatically that the epitope was also disrupted, possibly at a different site of the antigen. These problems may be reduced by alanine substitution studies where one by one the crucial residues are substituted by alanine which is expected to have neutral effects on the actual conformation of the antigen but may reveal essential side chains for antibody binding (Jin et al., 1992). Nevertheless, this approach provides high resolution and usually gives a good picture of the epitope.

- **Structural predictions based on sequence data**
  These provide a highly hypothetical approach that can help to target certain residues for mutational analysis or highlight areas for peptide studies. Facts about the primary sequence of the antigen based on the chemical properties of the amino acids and other known protein sequences are gathered and structural
predictions for the antigen are made. However, the method is not directly suitable as a mapping technique and still requires experimental verification. Molecular modelling of antigen-antibody binding sites has also been used to suggest crucial residues, but this also requires prior information (Padlam and Kabat, 1991).

- **Protection assays**

Chemical derivatisation of amino acids or proteolytic digestion of the peptide chain can be prevented in areas where the antibody interacts with the antigen and thus leads to regions in the primary sequence that are participating in the interaction. This method assures correct conformation of the antigen during treatment and large fragments generated often retain the conformation of the native antigen, but identification of individual residues is difficult and microanalysis of digested fragments is required.

Whether mapping techniques based on the use of linear peptide fragments truly represent epitopes or whether in most cases they just represent parts of larger overall conformational structures is a matter of debate. The fact is that in many cases mapping studies involving techniques based upon linear peptides have been extremely successful, but in others no region could be conclusively identified as the target. It should be kept in mind when using peptide based techniques however, that other residues (or areas) may also contribute to binding and that to some extent the residues identified may represent a somewhat artificial epitope. On the other hand linear target sequences have been demonstrated to be highly specific, antigenic and sometimes immunogenic. This has led to examples where immunisation with peptides representing parts of the pathogen gave rise to neutralising antibodies which could provide the basis for vaccines against a number of diseases (diphtheria: Audibert et al., 1982; cholera: Jacob et al., 1984; malaria: Patarroyo et al., 1988; common cold: Francis et al., 1987b; Hastings et al., 1990; influenza: Muller et al., 1982; foot-and-mouth disease: Parry et al., 1989; Francis et al., 1987a; and AIDS: Javaherian et al., 1990; Rini et al., 1993). The idea of using only synthetic parts of an antigen as a vaccine is not new and already developed in the 1960s when antibodies against a peptide representing parts of the tobacco mosaic virus coat protein inhibited assembly of the virus and infectivity (Anderer, 1963)
1. Introduction

I.C)  Hepatitis B Virus: Studying the Immune Response to a Complex Antigen

I.C.1)  The morphology of hepatitis B virus (HBV)

The hepatitis B virion, or Dane particle as it is sometimes called after its discoverer Dane (1970), is found in the blood and the liver of infected individuals as a spherical particle of ~42 nm diameter. The outer envelope consists of three forms of the viral surface antigen (HBsAg) called S (short), M (middle) and L (long), embedded in a lipid bilayer membrane derived from the host cell since the virions leave the cells through a budding process (Ganem, 1991). The inner capsid, or core particle, is ~27 nm diameter and occurs in two sizes assembled of 180 or 240 monomers of the core antigen (HBcAg), respectively (Crowther et al., 1994). A truncated form of HBcAg, named HBeAg exists as a non-particulate secreted antigen in the serum of infected individuals (Magnius and Espmark, 1972) and is correlated with the number of viral particles in the blood and relative infectivity, and its disappearance has therefore been used as an indicator for viral clearance at the end of an infection process.

The partially double stranded viral DNA genome is packaged inside the capsid together with a viral DNA polymerase activity and a protein bound to the 5' end of the minus sense DNA-strand.

In addition to complete infectious HBV virions two other particulate non-infectious structures can be found in the serum of infected individuals which are basically "empty" envelopes not containing any viral nucleic acids. They may contain empty nucleocapsids which is rare in the serum but can be found in infected liver tissue. The two forms occurring are either small spherical particles or long rod-shaped filaments of variable length which both are around 22 nm diameter. They consist almost exclusively of the S form of HBsAg and only very few molecules of M and L. Both these antigens are produced in large quantities in infected host cells which has the effect that the subviral particles are actually present in vast excess (10^3-10^6-fold) over complete virions in the serum (Simon et al., 1988). The biological function, if there is any at all, may be to act as decoy molecules for neutralising antibodies generated by the immune response of the infected host (Ganem and Varmus, 1987). Subviral particles are non-infectious and highly immunogenic resulting in a
I. Introduction

A protective immune response which made them a successful vaccine against hepatitis B. Originally the particles were isolated from the sera of patients which made the vaccine expensive and also potentially risky due to other pathogens associated with a blood derived product. With the development of molecular cloning and genetic engineering in the 1970s and 80s the vaccine is now produced in large quantities and extremely pure by recombinant DNA methodology in microorganisms.

I.C.2) Clinical aspects

HBV infection is considered to be one of the most important virus infections in the world. Typically, an infection with HBV ranges from very mild, inapparent disease to fulminant hepatitis, severe chronic liver disease and liver cirrhosis (Ganem, 1982). At present it is estimated that there are more than 300 million people world-wide who are chronic HBV carriers, many of whom may suffer from liver failure or hepatocellular carcinoma (HCC) eventually. In this respect HBV has been shown to be associated with HCC and two possible mechanisms for tumour formation have been advanced. Integration of viral DNA into host chromosomal DNA may cause activation of cellular oncogenes (Hino et al., 1986; Fourel et al., 1990) or disruption of tumour suppressor functions (Zhou et al., 1988). Alternatively, the continuous event of cell death and regeneration in chronic liver inflammation could increase the probability of critical mutational events leading to subsequent tumour formation (Chisari et al., 1989). The latter theory is favoured by many scientists because patients with liver diseases that are not connected to HBV but also cause chronic liver inflammation also show high incidence of HCC.

The spread of HBV is linked to social and environmental circumstances. In developing countries, where there are high rates of HBV carriers, the virus spreads mainly maternal-neonatally, i.e. vertically from one generation to the next. In industrialised countries however, virus spread is mainly horizontally by sexual contact, drug abuse and direct contact with infected blood products or blood samples by people that work in the health sector or in science. Interestingly, the epidemiology of HBV and HIV, the virus which causes AIDS are very similar and both viruses affect the same high risk groups among the population, but it should be stressed that HBV is more contagious than HIV in terms of the likelihood of contracting the disease on exposure to the virus.
In many cases infection with HBV leads, after an incubation time of two to six months, to acute hepatitis with jaundice and other symptoms. HBV infection can then be detected by identifying HBsAg in the serum. In the classical form of acute hepatitis patients recover completely with an immunity against reinfection for many years. In some cases, however, the disease remains permanently asymptomatic. In a minority of cases, infection will develop into fulminant hepatitis, a form of the disease which is usually fatal. It is caused by a very strong immune response of the host which destroys large parts of the liver by killing all cells bearing HBV antigens as markers by cytotoxic T-cells.

In chronic carriers no antibodies against HBsAg are ever detected and the viral antigens can be detected in the serum for a lifetime. The infection is therefore never cleared and the virus persists in the liver, causing various states of chronic disease patterns again ranging from more or less asymptomatic to liver cirrhosis and HCC. The risk of developing HCC among chronic carriers is at least 500-fold higher than normal, probably for the same reasons discussed above (Beasley et al., 1989).

I.C.3) Replication and transcription of the HBV genome

After the virion has successfully entered a host cell by binding to liver cell-specific receptor molecules in a process which is still poorly understood, the viral DNA migrates to the nucleus. The first step during a primary infection is the conversion of the asymmetric open circular (oc) DNA of the virion into a covalently closed circular (ccc) replicationally active form. This is the starting point for both viral replication, via reverse transcription of a pregenomic full length RNA transcript by the virally encoded DNA polymerase activity, and gene expression involving additional subgenomic mRNA transcripts. The various stages can be summarised as follows:

1. Conversion of the ocDNA originating from the virion to cccDNA by enzymes of the host cell.

2. Transcription of cccDNA by host cell RNA polymerase II generates a 3.5 kb full length pregenomic RNA which represents the plus strand and serves as template for the viral reverse transcriptase to create DNA minus strands and as mRNA for the core antigen (C gene product) and the viral polymerase (P gene product; Will
et al., 1986, 1987). Other transcripts of 2.4 kb and 2.1 kb length function as mRNAs for the remaining viral gene products. All transcripts are unspliced and capped. Since neither of the two shorter mRNAs cover the C and preC open reading frame (ORF) or the P ORF, the pregenomic RNA has to fulfil the mRNA function for the C and P gene. The 2.4 kb mRNA covers the ORFs for the surface antigen domains PreS1 and PreS2 (S gene) and the viral transactivator HBx (X gene) but is believed to function mainly as template for the expression of the L form of HBsAg which includes PreS1 and PreS2 (Will et al., 1987; Fig.1.2). The 2.1 kb mRNA serves as template for the translation of the M and S form of HBsAg (Cattaneo et al., 1983, 1984). Although it is now clear that the P gene, like the C gene, is expressed from the pregenomic 3.5 kb RNA, no mRNAs that contain the message for the P or the X gene alone have been found in natural infections. However, several additional mRNA species have been identified in HBV-transfected human (Saito et al., 1986; 0.7 kb) and rat (Gough, 1983; 1.0 kb) hepatocytes and in WHV-transfected woodchuck hepatocytes (Kaneko and Miller, 1988).

![Diagram](image)

**Figure 1.2:** Transcripts of HBV genes: A) the full length genomic RNA serves as mRNA for the expression of PreC (HBeAg) and C (HBcAg) by starting at two different transcription sites (note: transcript also serves as mRNA for P gene but this ORF is not shown here); B) the 2.4 kb transcript serves as mRNA for the expression...
of the L form of HBsAg; heterogeneous 5' ends of the 2.1 kb transcript generate mRNAs for expression of the M and S form of HBsAg.

3. The pregenomic RNA and the viral polymerase activity (P gene product) are packaged into capsids which readily assemble in the cytoplasm from core antigen monomers by largely unknown mechanisms, possibly involving high-molecular weight intermediates and cellular chaperones (Lingappa et al., 1994). During encapsidation the first minus strand DNA molecules are synthesised by the viral polymerase, or reverse transcriptase, using a protein as primer. Simultaneously with the DNA minus strand synthesis the pregenomic RNA is degraded.

4. The complementary plus strand is synthesised within the capsid, during envelopment and export of the virion, and utilises an oligonucleotide from the degraded pregenomic RNA as primer. Pores in the capsid might allow small molecules including nucleoside triphosphates to enter the capsid for polymerisation of the extending plus strand (Crowther et al., 1994). Plus strand synthesis stops as soon as the virion buds off the host cell. As a result of this the length of the plus strands and hence the single stranded gap in the genome is variable among virions depending on the time elapsed until budding. Before release from the host cell the capsid harbouring the polymerase and the pregenomic RNA is enveloped by budding into the lumen of the ER where it then follows the secretory pathway to the surface of the cell.

I.C.4) The genes of HBV and viral gene products

The nucleotide sequence of HBV was determined by two groups in 1979, Galibert et al., and Pasek et al., which led to two different numbering systems which are both used today. Galibert et al., specified the unique EcoRI restriction site in the genome as nucleotide No.1 which bears the inconvenience that not all sequenced HBV genomes have an EcoRI restriction site and that nucleotide No.1 is within two coding sequences. Pasek et al. therefore chose the first coding nucleotide of the C gene as No.1. Most of the four open reading frames C, S, P and X overlap substantially and so about half of all nucleotides are part of more than one reading frame (Fig.I.3). Furthermore, some genes within the same reading frame possess multiple start codons which give rise to several gene product subspecies of the same gene. This is the case for both structural antigens HBsAg and HBcAg.
I. Introduction

Figure 1.3: Organisation of HBV genes: The four major genes are shown in relation to the total length of the HBV genome (~3.2 kb); the S gene (HBsAg) consists of three subgenes coding for the PreS1, PreS2 and S domains; the C gene also has two transcription start sites giving rise to HBeAg (including PreC) and HBcAg.

The four gene products of HBV are:

The S gene gives rise to three related surface antigens (HBsAg) using three different in frame start codons. The functional parts of the gene can be divided into PreS1, PreS2 and S and code accordingly for the PreS1; the PreS2 and the S domain of HBsAg. The shortest transcript codes for the S (short) domain consisting 226 amino acids. The medium length transcript codes for the S and the PreS2 domains which form the M (middle) form of HBsAg and adds another 55 amino acids to the N-terminus of the S domain. Both transcripts are driven by the same promotor. The longest gene product is L (long) which contains all three domains encoded by PreS1, PreS2 and S. The PreS1 domain adds another 108-119 amino acids, depending on the virus subtype, to the N-terminus of M (Fig.1.4).
The C gene codes for the core antigen (HBcAg), a 21kD protein with 184 amino acids which is the subunit of the nucleocapsid. It has two distinct domains with the N-terminal 144 amino acids being sufficient for self-assembly of capsids. The C-terminus is extremely arginine rich due to multiple SPRRR(R) motif repeats and is probably involved in non-specific RNA binding during virion assembly. A short region named PreC proximal to the C gene encodes a signal sequence that directs translated transcripts bearing this signal region into the secretory pathway of the cell. During post-translational processing parts of the N-terminal PreC region and the C-terminus are cleaved off leaving a 10 amino acid N-terminal extension to the truncated primary sequence that the gene product has in common with HBcAg. This resulting 17kD viral antigen, termed HBeAg, is secreted and only found in a soluble form in the serum (Fig. 1.5).

Although HBcAg and HBeAg share in large part an identical amino acid sequence, they have a very different conformation resulting in different antigenicity, which results in little cross-reaction between anti-HBcAg and anti-HBeAg sera (Salfeld et al., 1989). HBe epitopes can be detected however, on denaturation of core antigen (i.e. particles; MacKay et al., 1981) which implies that HBe epitopes are cryptic on core antigen and only become exposed on denaturation. The conformational differences observed are related to the few amino acids which are different between the two antigens, including cysteine residues in particular which are responsible for
the folding framework of these antigens (Nassal and Rieger, 1993; Wasenauer et al., 1993). In the case of HBeAg an additional cysteine residue is found in the short N-terminal extension (C₇) whereas C₁₈₃ is deleted due to its truncation in comparison to HBcAg.

The **P gene** which codes for the DNA polymerase, or reverse transcriptase, activity is, considering the length of the entire genome, relatively long and overlaps with all other open reading frames.

The **X gene** and its product HBxAg was the last to be discovered since detection of the gene product was hampered by the low expression levels of this protein which had, until recently, no known function. It was finally detected by antibodies raised against the recombinant protein in transfected cell cultures (Pugh et al., 1986) and infected liver tissue (Moriarty et al., 1985). Antibodies that cross-reacted with HBxAg were detected in the sera of infected individuals (Kay et al., 1985; Weber et al., 1988), providing evidence that the antigen was expressed at some stage during the viral cycle. It is clear now that HBxAg acts as a transcriptional activator for HBV gene expression (Colgrove et al., 1989; Rossner et al., 1990) but it can also serve as a transactivator for genes under control of foreign viral promotors such as SV40, Rous Sarcoma and HIV. It is also known that HBxAg induces maximal activation of genes under control of the transcriptional activator NF-κB in hepatocytes (Lucito and
Schneider, 1992). However, the precise mechanism for HBxAg transactivation is still unknown but is believed to be mediated through interaction with other transcriptionally active proteins (Rossner, 1992).

I.C.5) The surface antigen of HBV (HBsAg)

The three related products of the surface antigen gene which share identical protein domains (Fig.1.4), provide a good example of the extremely efficient use of the small HBV genome to generate several gene products from one open reading frame by using various translation start codons. HBsAg not only occurs in three differently sized forms but also in various glycosylation states and investigators became interested in the question whether the different forms of HBsAg fulfilled different functions. However, it was not until recently that details about the structure and also about the involvement of the various forms of HBsAg in the biology of HBV began to emerge. Some of the reasons that made progress comparatively slow included: i) in vivo the three different forms occur in very different concentrations and always as a mixture, ii) they are conformationally sensitive transmembrane proteins with a complex structure that is essential for their functional integrity, iii) different HBV subtypes have variable amino acids in crucial positions of HBsAg affecting cross-reactivity with anti-HBsAg sera which probably reflects differences in conformation, and iv) expression in E.coli gives only poor yields (MacKay et al., 1981) so that expression in yeast or animal cells in culture is preferable (Valenzuela et al., 1982; Miyanohara et al., 1983; Hitzeman et al., 1983). Recently, HBsAg has been successfully expressed in vitro (Ostapchuk et al., 1994; Prange and Streeck, 1995; Dyson and Murray, 1995).

The shortest polypeptide S (226 aa) is by far the most abundant form and constitutes the major fraction of the envelope of mature virions and subviral particles. M (226 aa + 55 aa) can also be found in the viral envelope and, in small amounts, in subviral particles. The largest form L (226 aa + 55 aa + 109-119 aa) is mainly present in complete virions and therefore gives it unique properties. Both PreS-containing forms M and L comprise only up to 10% of the total HBsAg.

The surface antigens are co-translationally integrated in the ER membrane and L plays a crucial role in aggregating other S molecules into patches to interact with
core particles to form complete enveloped virions by budding into the ER lumen (Ganem, 1991). Several factors seem to influence this process and the expression levels of the three forms of HBsAg S, M and L probably play a role in regulating the export of complete virions or subviral particles (Bruss and Ganem, 1991a). Subviral particles also bud off the ER membrane but independent of an interaction with core particles. Both complete virions and subviral particles then follow the secretory pathway to the cell surface where they are released.

Domains pointing into the ER lumen after integration into the membrane are generally believed to become external on the released virion whereas regions located on the cytosolic side of the ER become buried inside the envelope i.e. not accessible from outside. Recent evidence, however, suggests that this simplistic view is not entirely true for HBsAg.

Synthetic peptides representing parts of the PreS domains of L and M had been shown to be highly immunogenic, generating antibodies cross-reactive with native HBsAg and even neutralising antibodies in chimpanzees; they were also recognised by anti-HBsAg antibodies (Neurath et al., 1984, 1985, 1986; Okamoto et al., 1985). In order to be recognised by B-cells the identified sequences required an external location on the virion. It was also found that PreS1 sequences were involved in cell attachment of HBV particles to HepG2 cells (a hepatoma cell line) in culture (Neurath et al., 1986) and liver cells (Pontisso et al., 1989; Budkowska et al., 1993) and that a MAb directed against a linear epitope at the N-terminus of PreS1 inhibited attachment (Pontisso et al., 1989). This observation also implied an external orientation of the PreS1 domain. Sensitivity of the PreS sequences to trypsin digestion further supported an external location (Heermann et al., 1987).

In contrast with these observations, PreS1 had also been linked with viral assembly, mainly for the following reasons:

L is mainly present in complete virions and is essential for secretion of virions in transfected cell cultures (Bruss and Ganem, 1991b). L alone is not secreted as subviral particles (Persing et al., 1986), in contrast to M and S (Sheu and Lo, 1992), which points towards a crucial function for L during assembly. Furthermore, the glycosylation site within the PreS2 domain of L is not glycosylated in virions (Heermann et al., 1984) which could mean that the PreS sequences may initially not
be located in the lumen of the ER where they could undergo post-translational modifications.

Work by Dyson and Murray (1995) recently showed that L but not S interacted with core particles in an *in vitro* system. This also suggested a crucial involvement of PreS sequences in the interaction between HBsAg and core particles.

Collectively, these results indicate that at least PreS1 maybe located on the cytoplasmic side of the ER after integration, but how could the same domain be orientated to the lumen of the ER (i.e. externally on the virion) at the same time to fulfil the functions outlined above? Recent studies which compared the topology of newly synthesised L in the ER membrane with that of L in complete virions secreted from transfected cell cultures (Bruss et al., 1994) or studied the topology of L in *in vitro* expression systems (Ostapchuk et al., 1994; Prange and Streeck, 1995) indicated that HBsAg may undergo unique post-translational translocation events. According to this almost all PreS domains are initially orientated cytoplasmically allowing interaction with core particles and thus triggering off the budding process. During virion maturation the domains would then become translocated across the envelope membrane so that on mature virions about 50% of all L molecules would have their PreS domains exposed on the surface. This scheme satisfies the involvement of L on both sides of the envelope membrane and solves the apparent contradiction.

Further analysis is required to identify the precise contact sites between the surface antigen and core particles which may depend on a specific conformation of L. N-terminal deletion studies of L have indicated that residues 1-102 may be deleted without effects on virion secretion but that deletion beyond residue 102 abolished secretion (Bruss et al., 1994). However, it is difficult to draw conclusions from cell culture transfection studies which use the secretion of virions as an indicator for the interaction of HBsAg with core particles. Evidences of *in vitro* studies of the interaction between N-terminal deletion mutants of L and core particles indicate that deletion of up to position 60 of PreS1 has no negative effect on the interaction but that further deletion results in reduction of binding and is completely abolished when the whole PreS domains are deleted (Wen Siang Tan, unpublished observations). The S domain may also contribute to the interaction of L with core particles since *in vitro* studies on the topology of HBsAg in the membrane recently suggested a cytosolic
I. Introduction

I. C. 6) The immune response to HBV

General aspects
After infection with HBV an individual usually mounts an immune response against the virus which in most cases leads to viral clearance and full recovery from hepatitis. The wide range of symptoms observed for HBV infections however, are a result of the varying types and intensities of responses in different individuals.

During a normal immune response a balanced interplay of the humoral, i.e. antibody based, and cellular, i.e. lymphocyte-mediated, immune responses is observed. Antibodies are produced by B-cells which recognise epitopes on the surface of antigens circulating in the body fluids. Most B-cells require helper functions from activated T-helper cells (Th) to undergo clonal expansion and produce antibodies. The Th cells in turn become activated by interacting specifically with antigen presenting cells (APC). B-cells usually encounter antigens in their native conformation as they occur in the blood and the lymph fluids, in this case complete virions and subviral particles. Consequently, the antibodies produced by these B-cells, or plasma cells once they are activated, are directed against the outside of the virion i.e. the parts of the surface antigen which are exposed on the surface. In contrast, cytotoxic (Tc) and helper T-cells recognise parts of antigens which result from degradation of the antigen and subsequent display of viral peptides on the surface of infected host cells, APCs and B-cells. T-cell epitopes are therefore short linear peptide segments of the primary sequence of the viral antigen whereas the majority of B-cell epitopes are conformational since they bind to native structures on the surface of the complete particle. T-cell epitopes are only recognised in connection with proteins encoded by the major histocompatibility complex (MHC) on the cell surface of infected host cells (MHC class I) or antigen presenting cells (MHC class II).

Internal antigens of HBV, such as the core antigen, can also confer protection since vaccination with recombinant core antigen resulted in partial protection of
chimpanzees (Murray et al., 1984). It was thought that activated $T_h$-cells could provide helper functions to B-cells producing virus neutralising antibodies directed against HBsAg which was later confirmed by studies by Milich et al. (1987b). It was shown that it is crucial that HBcAg T-cell epitopes and HBsAg epitopes are physically linked, as for example in an HBV virion. B-cells specific for HBsAg epitopes take up whole virions and process the antigens displaying many HBcAg T-cell epitopes that can be recognised by primed $T_h$-cells which then induce production of the HBsAg-specific antibodies. It was also demonstrated that immunisation with a synthetic peptide carrying a $T_h$-cell epitope from HBcAg and a B-cell epitope from PreS2 induced antibodies which cross-reacted with native HBsAg (Milich et al., 1988).

Neutralising antibodies aggregate and mark virions for uptake and degradation by APCs and lymphocytes, or prevent binding of the virions to receptors on the surface of target hepatocytes. Despite the fact that potentially the whole surface of an antigen is antigenic most antibodies will be directed against a few so called immunodominant areas of the antigen which are readily accessible to the immune system and elicit neutralising antibodies.

HBsAg is the prime target for the immune system

Early serological studies revealed a major immunodominant area on HBsAg which was common to all HBV subtypes and consisted of more than one epitope. This area, termed the $a$ determinant, is located in a hydrophilic region between residues 122-147, which conforms with an external positioning of this determinant (Bhatnager et al., 1982; Prince et al., 1982). Other determinants, not related to the $a$ determinant, were also identified which led to the classification of the subtypes we know today. In an attempt to further elucidate the elements of these determinants, and maybe draw conclusions for the design of synthetic vaccines, numerous studies were carried out with fragments of HBsAg or, more often, synthetic peptides representing the $a$ determinant and various other parts of HBsAg, covering almost the whole primary sequence (Bhatnager et al., 1982; Brown et al., 1984; Dreesman et al., 1982; Gerin et al., 1983; Howard et al., 1988; Lerner et al., 1981; MacKay et al., 1981; Prince et al., 1982; Neurath et al., 1982). Many of these peptides elicited antibodies in rabbits or mice which recognised native HBsAg (immunogenicity) or the peptides cross-reacted with antibodies (mono- or polyclonal) specific for native HBsAg (antigenicity). In
some instances the latter was not investigated, but in others peptides were both immunogenic and antigenic.

Taken as a whole the results obtained from these experiments were of great value and demonstrated well the complexity of the immunological properties of HBsAg. In some cases they led to the identification of virus neutralising epitopes which even conferred partial protection when injected into chimpanzees (Gerin et al., 1983; Purcell et al., 1985), but generally they did not contribute in a large degree to our understanding of the folding and the three-dimensional orientation of HBsAg in the envelope. This lies mainly in the nature of using short linear peptides for studying folded antigen structures and their possible failure to reproduce correct conformations. However, peptide studies were very successful in identifying T-cell epitopes within surface and core antigen since they occur naturally as short linear sequences (Milich et al., 1985, 1986, 1987a).

Point mutational analysis identified residues which contributed to the correct folding of HBsAg and which had profound effects on the conformation if changed or deleted (Ashton-Rickard and Murray, 1989a,b; Antonini and Peterson, 1981). Most of these residues were cysteines which could form disulphide bridges with other cysteines thus creating a secondary structure framework. This is the reason why some peptides were more effective in inducing HBsAg-specific antibodies or being recognised by anti-HBsAg antibodies when they were presented in a cyclic form which was achieved by oxidising internal cysteine residues (Brown et al., 1984; Dreesman et al., 1982; Howard et al., 1988).

Fig. II.8 (p.71) gives an overview of most of the peptides used in various studies of the immunogenicity and antigenicity of HBsAg (S). Residues 1-160 are well covered by a range of peptides, but C-terminal areas have been relatively poorly investigated and peptides 212-226 and 221-226 (not shown) failed to induce antibodies at all (Lerner et al., 1981). Peptides covering residues 140-160 depended strongly on the inclusion of crucial cysteine residues which probably forced the peptides to adopt a conformation similar to possible loop structures which have been suggested to create a basic framework for the conformational organisation of the a determinant.
I. Introduction

Figure I.6: Possible loop structures predicted for the "a determinant" of S (after Carman, 1994); individual amino acids are shown between residues 116 and 149; square boxes indicate positions of naturally occurring point mutations; possible disulphide bridges are indicated.

According to this model C\textsubscript{124} and C\textsubscript{137} and C\textsubscript{139} and C\textsubscript{147} or C\textsubscript{149} form disulphide bridges which create two loops (Carman, 1994; Fig.I.6). Accordingly, peptide 140-148 (not shown) which excluded C\textsubscript{139} was not immunogenic whereas cyclic peptides representing loop 1 (124-137) or loop 2 (139-147) where able to induce anti-S antibodies and were more potent than their linear counterparts in doing so. However, other structural arrangements are equally possible.

The region near the N-terminus seemed to be very immunogenic and not dependent on complex secondary structure because linear peptides induced antibodies that recognised native S (Fig.II.8, p.71). Some peptides not only induced antibodies (Lerner et al., 1981; Gerin et al., 1983) but also cross-reacted with anti-S antibodies (Howard et al., 1988). Three peptides between residues 31 and 52 (not shown) failed to induce antibodies against themselves (31-52) or failed to induce antibodies reactive with HBsAg (38-52 and 47-52) (Lerner et al., 1981). Although some of the failures may be attributed to the use of extremely short peptides which are usually not considered good immunogens since they are not very effective in reproducing a given conformation, it may however indicate that some areas, such as 31-52 or 212-226 cannot induce antibodies which recognise native HBsAg because these sequences may not be exposed (i.e. recognisable) on the surface of the virus particle \textit{in vivo}.

Area 110-140 has been investigated intensively, mainly because of interest in the \textit{a} determinant, the interesting cysteine residues in the region, and the subtype specific
amino acids exchanges which are located mostly in this region. Various peptides, all including C\textsubscript{137} but with variable N-termini, starting at residues 110, 117, 122, 124 and 125 respectively were all able to induce antibodies that reacted with HBsAg, or were recognised by anti-S antibodies (124-137; Howard et al., 1988), indicating that the variable N-termini did not have a great effect on immunogenicity (Fig.II.8, p.71). Cyclic peptides were more potent than their linearised counterparts. Surprisingly, peptide 125-139 (Gerin et al., 1983) was immunological active despite the deletion of C\textsubscript{124} thought to be involved in formation of the postulated loop 1 structure (Fig.I.6, p.37).

Point mutational analysis of S was mentioned earlier in connection with cysteine residues, but other residues which are either exchanged in different subtypes or occur in natural escape mutants have also been investigated. Mutant polypeptides were expressed and effects on the reactivity with polyclonal subtype-specific sera or monoclonal subtype-specific antibodies tested. Mutation of P\textsubscript{142} had negative effects on recognition by y-specific sera but change of R\textsubscript{122} and Y\textsubscript{134} had no such effect (Ashton-Rickard and Murray, 1989a). However, double amino acid replacement of S\textsubscript{113} and R\textsubscript{122} or S\textsubscript{113} and Y\textsubscript{134} resulted in HBsAg being recognised by both anti-y and anti-d-specific sera (original subtype y\textsuperscript{+}d\textsuperscript{-}) and induced antibodies reactive with both d and y subtype HBsAg. Triple mutants of S\textsubscript{113}, R\textsubscript{122} and Y\textsubscript{134} were also recognised by both anti-sera but more importantly, induced antibodies that reacted with only the d subtype, i.e. completely switching the subtype from y to d in terms of immunogenicity (Ashton-Rickard and Murray, 1989b). Recent studies by Bruce and Murray (1995) showed that mutation of the central cysteine residue of the three cysteine stretch 137-139, C\textsubscript{138}, had no effect whereas changing the two others were deleterious, consistent with their involvement in the predicted two loop structure. Changes of C\textsubscript{149} to serine also affected cross-reaction with anti-HBsAg sera which offers several possibilities for disulphide bridges between these residues (Fig. I.6).

These studies show that subtype specificities are also complex and probably determined by several amino acids which all contribute to the overall conformation and therefore cause changes in antigenicity when mutated. Escape mutants like G\textsubscript{145} to R (Carman et al., 1990; Bruce and Murray, 1995) probably have similar effects, so that neutralising antibodies can no longer recognise the altered antigen and thus escape the immune response (Waters et al., 1992).
I. Introduction

I.C.7) The monoclonal antibody MA18/7

The monoclonal antibody MA18/7 is widely used in research on hepatitis B virus (HBV) mainly because of its well characterised properties and its interesting effects on HBV biology. It interferes with attachment of HBV to liver cells (Pontisso et al., 1989) and is specific for the PreS1 domain of the surface antigen (Heermann et al., 1984). This domain has been connected with cell attachment and therefore binding of MA18/7 to PreS1 interferes with this process. This observation prompted numerous investigations into the precise target site of MA18/7 on PreS1. By using the PepScan method linear amino acids 18-25 of PreS1 (QLDPAfRA; Coursaget et al., 1991) and 17-20 (HQLD; Howard et al., 1991) were reported as binding sites for MA18/7. Studies involving the screening of deletion mutants of the coding sequence of PreS1 cloned into fr bacteriophage expression vectors (Sominskaya et al., 1992) identified four amino acids consisting of residues 20-23 (DPAF).

This monoclonal antibody was chosen as a model system to investigate the potentials of phage-display technology for high-resolution epitope mapping, and to compare the results obtained with the results of the mapping techniques applied previously, which had led to somewhat incompatible views on the target sequence for MA18/7 (e.g. compare residues identified by Howard et al., 1991 and Sominskaya et al., 1992).

It was particularly important to distinguish between internal residues in respect of their contribution towards binding of this antibody. None of the methods previously applied was able to address this question satisfactory. The approach chosen here was to immobilise the MAb on microtiter wells of an ELISA plate and then screen the pool of phage contained in the random phage-display library by multiple rounds of selection and amplification (biopanning).

I.C.8) The polyclonal serum sample of a vaccinated chimpanzee

The animal Ianthe was one of six chimpanzees inoculated with a recombinant HBV antigen in an immunisation evaluation (Murray et al., 1984). Ianthe and Dolf received at least two injections of hepatitis B virus surface antigen (HBsAg) produced by recombinant DNA methods in yeast (adyw subtype, coding sequence derived from pHBV130, Gough and Murray, 1982) and Gwen and Rinka received core antigen particles (HBcAg) that had been produced similarly in Escherichia coli.
I. Introduction

(Stahl et al., 1982). Two animals served as controls and were not vaccinated. After challenge with HBV (ayw subtype) the two controls developed hepatitis four weeks after infection with a normal disease pattern. Gwen and Rinka were partially protected and suffered very mild infections. The protections were believed to be due in part to anti-HBe specificities contained in the HBcAg preparation, but possibly also to anti-HBc to some extent. It has been demonstrated in the meantime that immunisation with these internal antigens of HBV leads to protection against HBV infection via cell-mediated immunity by priming T-cells to provide helper function for proliferating B-cells including those producing virus-neutralising anti-HBs antibodies (see also I.C.6, p.35).

<table>
<thead>
<tr>
<th>name</th>
<th>sex</th>
<th>vaccinated with:</th>
<th>injections</th>
<th>status after infection with HBV:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dolf</td>
<td>m</td>
<td>HBsAg</td>
<td>4</td>
<td>protected</td>
</tr>
<tr>
<td>lanthe</td>
<td>f</td>
<td>HBsAg</td>
<td>2</td>
<td>protected</td>
</tr>
<tr>
<td>Gwen</td>
<td>f</td>
<td>HBcAg</td>
<td>2</td>
<td>partially protected</td>
</tr>
<tr>
<td>Rinka</td>
<td>f</td>
<td>HBcAg</td>
<td>3</td>
<td>partially protected</td>
</tr>
<tr>
<td>Peter</td>
<td>m</td>
<td>-</td>
<td>-</td>
<td>developed hepatitis</td>
</tr>
<tr>
<td>Fraukje</td>
<td>f</td>
<td>-</td>
<td>-</td>
<td>developed hepatitis</td>
</tr>
</tbody>
</table>

Ianthe and Dolf were both protected against the challenge as expected since both animals showed high anti-HBs titres before the time of infection. Protection of Ianthe and Dolf was not surprising per se since inoculation with HBsAg was known to protect against the disease and had been used previously as a vaccine in the form of inactivated serum-derived subviral particles purified from patients. However, the use of recombinant surface antigen acting as an efficient vaccine was demonstrated by this experiment. The reason for protection is obviously the primed status of the immune system of an vaccinated individual leading to an efficient and quick secondary response on exposure to the surface antigens of virus particles. Ianthe was chosen for the investigation of the antibody repertoire, primarily for the high anti-HBsAg titre she developed on challenge with HBV. The serum samples were from day 301 which was 117 days post-infection (day 184) and 196 days after the first
I. Introduction

inoculation, and from day 105, one day before the first inoculation (pre-immune serum).

After viral challenge the vaccinated animals showed a typical secondary response with immunoglobulins of the G class being the major fraction in the blood. IgG molecules of this stage (day 301) and the pre-immune stage (day 105) were prepared from Ianthe serum by two ammonium sulphate precipitations followed by an ion-exchange column. The antibody preparations were then biotinylated to provide linker groups for immobilising the antibodies on solid supports via biotin-streptavidin interactions.
II. Results and Conclusions

II.A.) Epitope Mapping of the Monoclonal Antibody MA18/7

II.A.1) Amplification of the phage-display library and verification of the random composition of the displayed peptides

Before the library could be used in biopanning experiments it was important to verify that the hexapeptides displayed on the phage were truly random. 15 phage clones were picked at random and the DNA inserts analysed by sequencing. Peptide sequences of the pIII-fusions were derived by translating the DNA sequence by using the genetic code (Table II.1). All analysed phage displayed different hexapeptides and ~50% of them consisted of six different amino acids.

<table>
<thead>
<tr>
<th>clone no.</th>
<th>hexapeptide displayed</th>
<th>number of different aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GMSYYS</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>HAERYL</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>LFATCG</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>HLAATG</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>GTQIFS</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>SRMDYL</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>AQLSGF</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>LETALN</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>VKRITG</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>YGGFML</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>LSHHRD</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>KLSSLNA</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>FVSSSHV</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>MNPQGF</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>FVTARG</td>
<td>6</td>
</tr>
</tbody>
</table>

Table II.1: Amino acid sequences of the pIII-fusions analysed

Table II.2 demonstrates two important findings: First, all six positions showed a similar diversity although position 2 was somewhat above and position 6 somewhat below the average. Second, the amino acids occurred in frequencies roughly
II. Results and Conclusions

according to the number of possible triplets that coded for them (bearing in mind that the library was constructed in a way that excluded 50% of all triplets). For example, leucine was the most abundant amino acid in accord with three triplets that could be used to code for a Leu residue. However, two exceptions could be mentioned which occurred either above (F) or under (P) their expected codon representation but this could be attributed to the relatively small number of phage analysed.

Table II.2: Frequency and distribution of amino acids between positions 1-6 and codon usage

<table>
<thead>
<tr>
<th>aa</th>
<th>frequency</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>codons possible</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>11</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>-</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>S</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>-</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>G</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>-</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>T</td>
<td>6</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>R</td>
<td>5</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>H</td>
<td>5</td>
<td>2</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Y</td>
<td>5</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>3</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>M</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>V</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Q</td>
<td>3</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>N</td>
<td>2</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>K</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>P</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>W</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>stop</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
</tbody>
</table>

Total number of different aa/position: 10 13 11 10 9 8 average: 10

Total number of aa/position: 15 15 15 15 15 15 = 90
II.A.2) Defining the epitope recognised by MA18/7

The intention here was to apply the phage-display library to a well studied antibody, presumably recognising a linear epitope and compare the results obtained with other mapping techniques. The approach chosen here was to immobilise the MAb on microtiter wells of an ELISA plate and then screen the pool of phage contained in the phage-display library by multiple rounds of selection and amplification (biopanning).

Table II.3 shows the yields (ratio of input to output phage) obtained from the four rounds of biopanning. From round 1 to 2 there was a 5-fold increase in yield followed by a 16-fold increase in round 3. Round 4 showed the smallest increase in yield. It has to be noted however that the amounts of input phage were not identical due to variations in titre for the amplified eluates. In this case the yields from each round do not truly reflect the increase in phage recovered from the wells. If the absolute number of phage recovered from each well is considered, it becomes evident that, with the coating conditions being identical for each round, the number of phage recovered rose from $2.5 \times 10^4$ t.u. in round 1 to $1.1 \times 10^8$ t.u. in round 4, equal to a 4,000-fold increase.

<table>
<thead>
<tr>
<th>Round</th>
<th>Input phage</th>
<th>Output phage</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1.9 \times 10^{10}$</td>
<td>$2.4 \times 10^4$</td>
<td>$1.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>2</td>
<td>$3.2 \times 10^{11}$</td>
<td>$2.0 \times 10^6$</td>
<td>$6.3 \times 10^{-4}$</td>
</tr>
<tr>
<td>3</td>
<td>$1.2 \times 10^{10}$</td>
<td>$2.4 \times 10^7$</td>
<td>$1.0 \times 10^{-2}$</td>
</tr>
<tr>
<td>4</td>
<td>$4.0 \times 10^{11}$</td>
<td>$1.1 \times 10^8$</td>
<td>$3.0 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

From rounds 1, 2 and 4 about a dozen phage were chosen at random and analysed by sequencing the gIII-insert of the phage genome. The nucleotide sequences were translated into the peptide sequences displayed on the individual phage. The colonies for template preparation were derived from the titration of the output titres. Fig.II.1 shows the alignment of all hexapeptides displayed on the analysed phage. After only one round of biopanning almost 50% of the phage analysed displayed peptides corresponding with three or four residues with the amino acid sequence of PreS1.
11. Results and Conclusions

Degree of conservation in % (29 matches total):

| 7 | 48 | 100 | 100 | 45 | 93 | 3 | 0 |

**PreS1 residues:**

<table>
<thead>
<tr>
<th>Q</th>
<th>L</th>
<th>D</th>
<th>P</th>
<th>A</th>
<th>F</th>
<th>R</th>
<th>A</th>
<th>Epitope as identified by Coursaget et al. (1991)</th>
</tr>
</thead>
</table>

**Round 1**

| 3/6 | T | V | D | P | S | F | E1-6 |
| 3/6 | I | D | P | A | Y | K | E1-10 |
| 4/6 | L | D | P | G | F | L | E1-4 |
| 4/6 | T | F | D | P | A | F | E1-3 |
| 4/6 | H | D | P | A | F | Y | E1-2 |

n. m.

| H | L | R | Y | F | R | E1-1 |
| P | V | L | V | L | T | E1-8 |
| L | S | S | W | P | N | E1-5 |
| H | Y | S | R | S | S | E1-7 |
| G | W | F | F | H | T | E1-11 |
| Y | P | K | R | G | V | E1-12 |

**Round 2**

| 4/6 | A | E | D | P | A | Y | E2-12 |
| 4/6 | L | L | D | P | G | F | E2-13, E2-19, E2-22 |
| 4/6 | S | L | D | P | G | F | E2-14 |
| 4/6 | T | L | D | P | G | F | E2-15 |
| 4/6 | R | Y | D | P | A | F | E2-20 |
| 4/6 | D | P | A | F | N | D | E2-17 |
| 4/6 | A | D | P | A | F | Q | E2-11 |
| 4/6 | V | D | P | A | F | K | E2-18 |
| 5/6 | L | D | P | A | F | V | E2-16 |
| 5/6 | V | L | D | P | A | F | E2-21 |

**Round 4**

| 3/6 | P | D | P | G | F | N | E4-2 |
| 3/6 | R | D | P | G | F | N | E4-8, E4-9 |
| 3/6 | F | D | P | G | F | L | E4-12 |
| 4/6 | K | D | P | A | F | E | E4-10 |
| 4/6 | R | Y | D | P | A | F | E4-6 |
| 4/6 | L | L | D | P | G | F | E4-1, E4-11 |
| 4/6 | L | I | D | P | A | F | E4-5 |
| 5/6 | Q | L | D | P | G | F | E4-4, E4-7 |
| 5/6 | S | L | D | P | A | F | E4-3 |

Figure II.1: Amino acid sequences of pIII fusion phage isolated from selection rounds 1, 2 and 4; residues 18-25 of PreS1, defined as the binding site for MA18/7 by Coursaget et al. (1991), are shown in bold; percentages shown at the top indicate the degree to which these residues were conserved among the phage matching in this area (29 in total); matching amino acids are shown in bold in grey boxes; figures to the left indicate how many amino acids of the hexapeptide were conserved for each individual phage clone (clone numbers to the right); n. m. = "no matches" and refers to hexapeptides from round 1 which display no obvious consensus motif with the PreS1 region shown here.
between position 19 and 23. All matching phage displayed the dipeptide DP (20, 21). Other phage from round 1 showed no obvious consensus with this or any other region of PreS1. From round 2 onwards all phage analysed matched in the region 18-24 of PreS1 and all phage displayed the dipeptide DP. No non-matching sequences were found. The phenylalanine (F) at position 23 was also conserved to a high degree and replaced by tyrosine (Y) in only two of the phage analysed. Position 22 which is sandwiched between the highly conserved residues P_{21} (100%) and F_{23} (93%) allowed substitutions from its native residue alanine to a high degree (55%) and indeed from round 2 onwards the core region DPGF seemed the most common one isolated. However, the amino acids substituting alanine at this position all had small side chains such as glycine (G), serine (S) and valine (V).

Fig. II.1 also shows the alignment of all pIII-fusion hexapeptides with the residues defined as the binding site for MA18/7 by Coursaget et al. (1991; bold box). Residues outside the core region DPAF, namely Q_{18} (7%), R_{24} (3%) and A_{25} (0%) were only very poorly or not at all conserved in the phage analysed. Further, the frequency of phage showing five out of six amino acids corresponding to PreS1 did not greatly increase after round 2 for the numbers analysed here, and no phage were found which had all six amino acids corresponding to PreS1.

The biopanning with MA18/7 clearly led to the identification of a consensus motif within the PreS1 sequence that consisted of three amino acids that were highly conserved. For residues D_{20} and P_{21} no substitutions were found whereas two of the phage had residue F_{23} substituted by another aromatic amino acid. Neighbouring residues were conserved to variable degrees but allowed considerable replacement, which did apparently not reduce binding affinity significantly. Interestingly, A_{22} which is located between the highly conserved residues P_{21} and F_{23} also allowed substitutions and therefore interrupts the strict linear sequence of residues that seem to be essential for binding.

Although the range of exchanges observed at this position is somewhat limited to amino acids with small side chains the properties of the actual side chain itself did not seem to matter. This is an example of a linear epitope where at least five linear residues do probably contribute to various degrees to the binding but where the three crucial residues sufficient for effective binding do not show a similar linear array.
II. Results and Conclusions

II.A.3) Three non-linear core residues, D, P and F, are sufficient for specific interaction of phage and antibody MA18/7

The biopanning experiment with MA18/7 defined three highly conserved core residues (D₂₀, P₂¹ and F₂₃) that are crucial for efficient binding of the antibody. Although this core motif was almost unchanged, neighbouring residues were exchanged to various degrees and also internal residues like A₂₂ allowed substitutions. However, the results of the biopanning alone did not allow precise conclusions about the affinities of these different phage clones for the antibody. It was of interest therefore to quantitatively compare the binding affinities of phage displaying different hexapeptides, for this could reveal the effects of specific amino acid substitutions on binding affinity. Preferably, the interaction should be in solution since conformational changes can occur when proteins are bound to solid supports (Friguet et al., 1984).

The assay employed here was developed from one used for studying antibody-antigen interactions by ELISA (Friguet et al., 1985). In contrast the modified assay used here relied on phage titre as the signal readout of the experiment. However, to allow dissociation constants (Kᵋ) to be determined from the data obtained, some theoretical conditions had to be fulfilled. As shown below phage E4-4 was chosen as a representative to prove that these criteria were met. Table II.4 is a summary of the phage chosen for the Kᵋ determination and the hexapeptides they display. Phage E2-12 represented the two phage where the highly conserved F₂₃ was replaced by another aromatic residue, tyrosine (Y). E4-2 has only the three highly conserved residues D₂₀, P₂¹ and F₂₃ corresponding with the primary sequence of PreS1 and therefore represents the essential core motif identified by biopanning. E2-14, E2-16 and E4-4 show an additional one or two residues corresponding with PreS1 and were chosen to identify the effects of the flanking residues on binding. E2-17 displays the core motif but additionally contains the frequently replaced internal A₂₂ residue.
II. Results and Conclusions

### Table II.4: Phage chosen for binding studies

<table>
<thead>
<tr>
<th>Phage No.</th>
<th>pIII fusion peptide(^a)</th>
<th>aa conserved</th>
<th>reason for selection(^b):</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2-12</td>
<td>ALDPAY</td>
<td>4 of 6</td>
<td>(F_{23}) replaced by tyrosine</td>
</tr>
<tr>
<td>E2-14</td>
<td>SLDPGF</td>
<td>4 of 6</td>
<td>core motif + (L_{19})</td>
</tr>
<tr>
<td>E2-16</td>
<td>LDPVFR</td>
<td>5 of 6</td>
<td>core motif + (L_{19}) and (R_{24})</td>
</tr>
<tr>
<td>E2-17</td>
<td>DPAFND</td>
<td>4 of 6</td>
<td>core motif + internal (A_{22})</td>
</tr>
<tr>
<td>E4-2</td>
<td>PDPGFN</td>
<td>3 of 6</td>
<td>core motif only</td>
</tr>
<tr>
<td>E4-4</td>
<td>QLDPGF</td>
<td>5 of 6</td>
<td>core motif + (L_{19}) and (Q_{18})</td>
</tr>
</tbody>
</table>

\(^a\) bold writing indicates residues matching with PreS1 amino acid sequence  
\(^b\) "core motif" = \(D_{20}\), \(P_{21}\) and \(F_{23}\)

The method developed for the determination of \(K_D\) for the interaction of phage with substrate molecules in solution can be divided into two steps:

**Step I** involves the incubation of a constant concentration of phage with variable amounts of antibody molecules in solution until equilibrium is reached. The purpose of **step II** is to determine the concentration of unbound phage present in this mixture by transferring the phage and antibody mixture to wells coated with the antibody for a short time period, after which all unbound phage are washed away and the bound phage eluted and titred as p.f.u. The concentration of phage eluted from these wells is proportional to the amount of unbound phage present in the mixture for the particular antibody concentration used.

That this proportional relation is valid was demonstrated by incubating various concentrations of phage (input) with wells coated with the antibody for 1 h. The wells were washed, bound phage eluted (output) and the titres of the eluates determined. Fig.II.2 shows that the output p.f.u. was directly proportional to the input p.f.u. over two log ranges. With this relationship established the concentration of free phage remaining in solution \(p\) could be determined since the total concentration of phage \(p_t\) incubated with the antibody was known:

\[
p / p_t = p.f.u. / p.f.u. \_O
\]

with

\[
p = p_t \times p.f.u./p.f.u. \_O
\]
II. Results and Conclusions

Figure II.2: Linear relationship between input and output p.f.u. from wells coated with MA18/7; varying concentrations of phage E4-4 (input) were incubated for 1 h on polystyrene wells coated with MA18/7 and the number of bound phage (output) after extensive washing determined by titrating eluted phage; assays were performed in triplicate.

Figure II.3: Time course experiment of phage binding to wells coated with MA18/7; phage E4-4 (10^8 p.f.u.) were incubated on wells coated with MA18/7 for different periods of time and the amount of bound phage determined by titrating the eluate; points represent the average of three assays.
**II. Results and Conclusions**

p.f.u.₀ is the titre obtained in the absence of antibody in step I (i.e. 100% of phage remain unbound in solution). The incubation time of 1 h was determined by a time course experiment in which a constant concentration of phage was incubated on antibody-coated wells for various time periods (Fig.II.3). The amount of phage bound to the wells coated with MA18/7 reached a plateau after 5 h with about 14% of the input phage bound. However, it was important that the percentage of free phage binding to the wells during step II was low since any large scale binding to the wells would have shifted the equilibrium towards the release of free phage from the phage-antibody complexes in solution:

\[
\text{[phage] + [antibody] } \Leftrightarrow \text{[phage/antibody]} \quad \text{(equilibrium in solution)} \quad (3)
\]

Fig.II.3 shows that it was unlikely that the equilibrium was disturbed since after 1 h the fraction of phage bound to the wells was under 5%. A further experiment was set up shown in Fig.II.4 to demonstrate this. Various concentrations of phage were incubated for 1 h on wells coated with MA18/7 and the unbound phage transferred to a second set of identically coated wells and again incubated for 1 h. The titres of phage eluted from the two sets of wells were identical within 95% of standard error limits indicating that the total number of phage in the solution was basically unchanged.

For determination of the $K_D^{rel}$ values of the phage shown in Table II.4, MA18/7 concentrations in step I were varied from $10^{-9.5}$ to $10^8$ M and measured at 4°C. All concentrations including one sample where no antibody was added (0M value) were assayed in triplicate and the mean calculated. For most phage more than one experiment was performed in which case the overall mean was taken. Table II.5 representatively shows the free phage remaining in solution for phage E4-2 expressed as percentage of the 0M value.
Figure II.4: Consistency of output titre for phage E4-4 binding to wells coated with MA18/7 over 1 h incubation time; plate 1 was incubated with various concentrations of phage E4-4 for 1 h (input plate 1), unbound phage then transferred to plate 2 and again incubated for 1 h (input plate 2); bound phage eluted from both plates are shown as p.f.u. (output plate 1 • and plate 2 ○); titres were determined in triplicate (plate 1) or duplicate (plate 2).
II. Results and Conclusions

Table II.5: Data obtained from binding experiments of phage E4-2 with MA18/7

<table>
<thead>
<tr>
<th>Concentration of MA18/7 in step I (M):</th>
<th>percentage of phage remaining in solution during step I ([\text{f.u.}/\text{p.f.u.o} \times 100]:)</th>
<th>fraction of phage engaged in phage-antibody complexes ((= x/pt):)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100 (set value)</td>
<td>0</td>
</tr>
<tr>
<td>(10^{-9.5})</td>
<td>35.6 ± 11.3</td>
<td>0.64</td>
</tr>
<tr>
<td>(10^{-9})</td>
<td>15.8 ± 1.9</td>
<td>0.84</td>
</tr>
<tr>
<td>(10^{-8.5})</td>
<td>5.8 ± 3.3</td>
<td>0.94</td>
</tr>
<tr>
<td>(10^{-8})</td>
<td>2.9 ± 0.6</td>
<td>0.97</td>
</tr>
</tbody>
</table>

From these data the \(K_{D_{\text{rel}}}\) can be determined in two different ways:

a) by plotting the percentage of free phage remaining in solution (i.e. the proportional percentage of phage bound to the solid phase during step II) at equilibrium versus the various antibody concentrations used during step I. As will be explained below the antibody concentration at which 50% of the phage remain unbound in solution during step I equals \(K_{D_{\text{rel}}}.\)

b) by applying the Scatchard equation which provides Scatchard plots. The actual values for \(K_{D_{\text{rel}}}\) can then be derived from the slope of these plots.

**a) Plotting the percentage of free phage against antibody concentration**

The titre obtained for the sample in the absence of MAb is set at 100% \((\text{p.f.u}_0)\). The concentrations of the reactants at equilibrium, with \(x\) being the concentration of the phage-antibody complex and \(a\) the concentration of free antibody can be calculated with the known total concentrations of \(p_t\) and \(a_t\), the total antibody concentration:

\[
x = p_t - p \quad (4)
\]

\[
a = a_t - x \quad (5)
\]

The association constant \(K_A\) is defined as:
II. Results and Conclusions

\[ K_A = \frac{x}{p \cdot a} \quad (6) \]

The reciprocal of \( K_A \) is \( K_D \), the dissociation constant, and one can replace \( p \) after (4) to give:

\[ \frac{1}{K_A} = K_D = a \cdot \frac{(p_t - x)}{x} \quad (7) \]

which can be transformed into:

\[ \frac{x}{p_t} = a / a + K_D \quad (8) \]

For the plots the fraction of free phage remaining in solution can be expressed as

\[ (1 - \frac{x}{p_t}) \quad (9) \]

which is plotted versus \( a \) so that:

\[ (1 - \frac{x}{p_t}) = y = a / a + K_D \quad (10) \]

\( y \) equals 1 (or 100%) when \( x = 0 \), i.e. when all phage remain free which is the case in the absence of antibodies (\( a = 0 \)). At the concentration at which 50% of the total phage are engaged in a phage-antibody complex (i.e. \( y = \frac{1}{2} \) because \( x = \frac{1}{2} \cdot p_t \)) \( a \) equals \( K_D \).

Fig.II.5 shows these plots for all six phage of Table II.4. The curves were fitted to the points using the least square curve fitting option to a hyperbolic function of the graphic package used for plotting the graphs (SigmaPlot, Jandel). From this operation the program delivers the value for the \( a_t \) at which 50% of the phage would remain in solution (i.e. the \( K_D \)).
II. Results and Conclusions

Figure 11.5: Percentage of free phage remaining in solution after incubation of phage ($10^8$ p.f.u.) with various concentrations of MA18/7 in solution at 6°C; percentage of free phage at equilibrium is plotted against the concentration of MAb in solution (free phage in the absence of MAb = 100%); curves were fitted to the points using the hyperbolic least square curve fitting option (Marquardt-Levenberg algorithm, Marquardt (1963)) of SigmaPlot (Jandel Scientific GmbH) and the $K_D^{rd}$ values were derived from this process; plots shown represent data of at least two experiments performed for each phage.
II. Results and Conclusions

b) Applying the Scatchard equation:

\[
\frac{x}{a} = \frac{(p_t - x)}{K_D} \quad (11)
\]

As shown above the concentrations of the reactants needed for the Scatchard equation can be calculated after (4) and (5). For Scatchard-plots the equation is rearranged to give:

\[
\frac{x}{p_t} \cdot a = 1/K_D - \frac{x}{K_D} \cdot p_t \quad (12)
\]

\(x/p_t \cdot a\) is then plotted against \(x/p_t\) with the slope equal to \(-1/K_D\). Fig.II.6 shows Scatchard plots for four phage from Table II.4. The values for \(-1/K_D\) were derived from a regression line put through the points by using the regression curve option of SigmaPlot.

Table II.6 shows a summary of the \(K_D^{rel}\) values obtained for the various phage by the two methods described above. All constants were in the nM range and showed only small differences for the various phage studied. The differences between the two methods of calculating the \(K_D^{rel}\) values a) or b), were neglectable. The minor differences observed can probably be attributed to the fact that in the case of the curve fitting determination an additional point at 100% (for \(a = 0\)) was created which might have constrained the curve fitting operation.

<table>
<thead>
<tr>
<th>Phage No.</th>
<th>pIII fusion peptide(^a)</th>
<th>(K_D^{rel}) values(^b) (nM)</th>
<th>(K_D^{rel}) values(^c) (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2-12</td>
<td>ALDPAY</td>
<td>0.79*</td>
<td>n.d.</td>
</tr>
<tr>
<td>E2-14</td>
<td>SLDPGF</td>
<td>0.26*</td>
<td>n.d.</td>
</tr>
<tr>
<td>E2-16</td>
<td>LDPVFR</td>
<td>0.11±0.08</td>
<td>0.10±0.01</td>
</tr>
<tr>
<td>E2-17</td>
<td>DPAFND</td>
<td>0.20±0.07</td>
<td>0.22±0.01</td>
</tr>
<tr>
<td>E4-2</td>
<td>PDPGFn</td>
<td>0.18±0.06</td>
<td>0.17±0.02</td>
</tr>
<tr>
<td>E4-4</td>
<td>QLDLPGF</td>
<td>0.18*</td>
<td>0.25±0.02</td>
</tr>
</tbody>
</table>

\(^a\) bold indicates residues matching with the PreS1 sequence  
\(^b\) determined by curve-fitting to a hyperbolic function
Figure II.6: Scatchard plots for the interaction of four phage (Table II.4) with MA18/7 at equilibrium at 6°C; $x$ is the concentration of phage-antibody complexes, $p_t$ the total phage concentration (for further details see text); points represent the average of two experiments.
II. Results and Conclusions

c determined by Scatchard plots
* no standard deviation given since curve fitting operation could not fit the curve to all points (see Figure II.5)
n.d. = not determined, note: these data were published together with other results by Dyson et al., (1995).

It is important to note that the values determined here should be regarded as relative values (\(K_D^{rel}\)) since several factors have to be considered. First, the presence of several copies of pIII-fusion hexapeptides per phage particle could mean that one phage is able to interact with one antibody molecule in solution and one antibody molecule on the solid support at the same time. The binding to solid phase-bound or solution-born antibodies was therefore probably not completely excluding each other. Second, the presence of multiple copies of pIII-fusion proteins per phage could also lead to one phage particle binding via multiple interactions to the antibodies on the solid support thus enhancing the apparent affinity of this phage. The affinities determined here are hence the result of a complex interplay of such effects and are therefore not to be confused with the \(K_D\) for an "ideal" interaction. The aim of these studies was to compare the affinities of phage displaying different hexapeptides and thus allowing conclusions about the contribution of individual amino acids to the interaction. Since the effects discussed above should affect all phage to a similar degree a relative comparison should therefore be valid.

The most dramatic effect on the affinity was observed when F\(_{23}\) was replaced by tyrosine (E2-12) with a three-fold reduced \(K_D^{rel}\) in comparison to the weakest interaction for a phage displaying phenylalanine in this position (E2-14). Phage E2-14 and E4-4 displayed additional PreS1 residues (L\(_{19}\) or Q\(_{18}\) and L\(_{19}\), respectively) to the core motif, but did not have significantly enhanced binding affinities no matter whether one or two additional residues were involved. Interestingly, phage E2-16 had a lower \(K_D^{rel}\) although A\(_{22}\) was replaced by an amino acid with a slightly bulkier side chain (valine).

It is possible that the valine residue had positive effects on the binding affinity for MA18/7 but it cannot be excluded that residues L\(_{19}\) and R\(_{24}\) also contributed to the binding. However, since R occurred only twice at this position among the 35 phage analysed this effect is probably rather small.
II. Results and Conclusions

If the two phage representing the core region are compared (E2-17 and E4-2) it is interesting to note that the replacement of A22 by a glycine is effectively neutral in respect of binding to MA18/7 which supports the hypothesis of A22 not being crucial for binding the antibody and therefore also would explain the high rate of replacement observed for A22.

The comparison of E2-14 with E4-2 shows that the contribution of L19 to binding of the antibody is at best very small since the phage with this residue replaced by a rather different proline residue had an even lower $K_{d,rel}$ value.

II.A.4) Synthetic peptides compete with selected phage for the interaction with MA18/7

The comparison of the dissociation constants of individual phage isolated from the biopanning experiment revealed that the three residues D20, P21 and F23 are crucial and also sufficient for effective binding of phage to the monoclonal antibody. Based on these results, it was of interest to see whether free synthetic peptides, representing to various degrees residues of the epitope were able to compete for binding of a phage carrying five out of six residues (E4-4, Table II.4) to the monoclonal antibody. If peptides competed with the phage it would be additional proof for the fact that the nature and the spacing of the residues were crucial for efficient interaction and not necessarily the way they were presented in the context of the phage particle. In particular, a peptide only containing the three core residues could demonstrate that these amino acids are sufficient for a specific interaction. Three peptides of variable lengths were synthesised:

\[Table\ II.7:\ \textit{Synthetic\ peptides\ used\ for\ competition\ assays}\]

<table>
<thead>
<tr>
<th>Peptide</th>
<th>amino acid sequence$^a$</th>
<th>comment:</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>LDPAFR</td>
<td>All six aa correspond to the PreSi sequence (18-24)</td>
</tr>
<tr>
<td>P2</td>
<td>DPGF</td>
<td>Core residues only; A22 replaced by glycine</td>
</tr>
<tr>
<td>P7</td>
<td>ALLTRILG</td>
<td>Negative control peptide with no relation$^b$</td>
</tr>
</tbody>
</table>

$^a$ bold residues indicate residues that match the PreSi sequence, $^b$ Note: Flanking A and G residues are present in all phage displayed sequences

59
II. Results and Conclusions

Figure II.7: Competition between phage E4-4 (10^8 p.f.u.) and various concentrations of peptides P1 (○), P2 (●) and P7 (▲) for binding to wells coated with MA18/7; percentage of phage eluted from the wells is plotted against the concentration of peptide added as competitor (p.f.u. eluted in the absence of peptide = 100%); points represent the average of two experiments; note: concentration of P2 was slightly lower than other peptides since aa analysis of the hydrolised peptide revealed a slightly lower concentration.
P1 had six out of six residues matching with the PreS1 sequence, a situation not encountered amongst the 35 phage analysed in II.A.2; Fig.II.1, p.46). P2 could be referred to as the minimal possible epitope containing the three crucial residues as discussed above. It should be noted that A$_{22}$ was replaced by glycine to restrict the matching residues to D, P and F only. It was shown in II.A.3) that the effect on binding of replacing A$_{22}$ was relatively small.

For the competition assays constant concentrations of phage E4-4 were mixed with serial dilutions of peptides ranging from $10^{-6}$ to $10^{-11}$ M for P1 and $10^{-4}$ to $10^{-9}$ M for P2 and P7. The mixtures were then incubated on wells coated with the antibody. After washing the amount of phage eluted from each well was then plotted against the concentration of peptide added to that well (Fig.II.7). P1 shows strong inhibition effects on binding of phage E4-4 to MA18/7 at concentrations above 10 nM and completely prevents binding of the phage at 1 pM. P2 required concentrations above 1 µM to show an effect on phage binding, however, at 0.1 mM phage binding was also significantly reduced to 8% of the value with no peptide present. P2 is therefore about 300-fold less effective as a competitor for the binding of E4-4 to MA18/7 than P1. Control peptide P7 showed no effects on binding of E4-4 even at the highest concentration. The results shown are the summary of two experiments. It should be mentioned that P2 was very hygroscopic which made it difficult to achieve precise input concentrations. The concentration was therefore determined by amino acid analysis of the hydrolysed peptide which resulted in a slightly lower than expected concentration. Fig.II.7 shows the corrected values for P2.

Although peptide P1 had only one amino acid more corresponding with the PreS1 primary sequence than the peptide sequence of E4-4 it required concentrations 10-fold higher than expected from the $K_D$ value determined for E4-4 to achieve 50% inhibition. This might be related to stereo chemical factors of unknown character and has been observed by other workers in similar experiments (Smith et al., 1993; Yayon et al., 1993). The fact that phage display multiple copies of the hexapeptides might explain the apparently higher affinity constant determined for the phage whereas the synthetic peptides are more likely to interact in a monovalent fashion and therefore have a higher $K_D$. Further, the conformational freedom of a free peptide in comparison to the same peptide displayed in a phage context is most likely to be considerably greater thus lowering the effective concentration of a particular functional conformation (Sachs et al., 1972). These effects might well explain the
II. Results and Conclusions

reduced efficiency of the even shorter peptide P2 to compete for binding, but could also reflect weak contributions of the additional residues present in P1. Additionally, the close proximity of charged residues at the two termini to the binding site might have had destabilising effects (Smith et al., 1993) for P2. Contribution of neighbouring residues present in the phage context have also been reported (Felici et al., 1993; Yayon et al., 1993) and cannot be excluded either.

Nevertheless, the three residues highlighted as being crucial for binding antibody MA18/7 in II.A.2) and II.A.3) were effective as a free peptide to act as specific competitor for the interaction of phage and antibody and therefore emphasise the essential character of these residues within the epitope recognised by MA18/7.
II.B) Investigating Antibody Specificities in the Serum of a Chimpanzee Vaccinated with Recombinant Surface Antigen

II.B.1) Biopanning polyclonal IgG antibodies with the random phage-display library

As described in the introduction (I.A.5, p.10) many different systems can be used to provide a matrix for immobilising antibodies for the biopanning procedure. Two approaches were chosen here, 1) streptavidin-coated petri dishes (3.6 cm diameter; BPIII), and 2) streptavidin-coated magnetic beads (2.8 μm diameter; DBXII). The amount of beads coated with the biotinylated antibody preparation was chosen on the basis to provide a similar surface area to the petri dish. The effects of the two different systems on the selection procedure or the nature of the selected phage could hardly be predicted before the biopanning experiments were carried out and therefore the two parallel routes were envisaged. It could be that the magnetic beads allowed a better presentation of the antibodies to the surrounding solution due to less steric hindrance on the surface of one bead. Consequently, the accessibility of the antibody molecules could be very different in the two systems possibly eliminating a number of F\textsubscript{ab} domains from being properly exposed to the phage in the case of the petri dishes.

a) Affinity selection (biopanning)

Table II.8 provides an overview of the two experiments performed to select phage from the phage-display library on the basis of their affinity for polyclonal IgG preparations from lanthe. The experiments differed not only in the solid support system used to immobilise the antibodies but also in the way the final round of biopanning was performed. In the case of BPIII, round 3 favoured the selection of higher affinity binders since amplified phage were in competition for binding sites (see Materials and Methods, p.113). There was no such pressure on the selection of phage using the magnetic beads which might reflect the fact that the yield actually decreased in round 3 for the immune serum IgG. Both outputs increased similarly in round 2 but only the biopanning with neutral serum also increased in round 3.
Generally however, yields are not a reliable indicator for a successful selection and even rounds with identical yields or yields close to background level (usually around $10^{-5}\%$) can contain an enriched population of phage (Smith and Scott, 1993).

**Table II.8: Biopanning experiments involving polyclonal serum IgG from lanthe**

1) Experiment using petri dishes coated with immune serum IgG (BPIII):

<table>
<thead>
<tr>
<th>Round</th>
<th>input ($\times 10^{11}$)</th>
<th>output ($\times 10^{5}$)</th>
<th>yield ($% \times 10^{-4}$)</th>
<th>coating (µg)</th>
<th>wash.</th>
<th>amplif. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4</td>
<td>0.77</td>
<td>1.8</td>
<td>10</td>
<td>8 x</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>1.5</td>
<td>5.7</td>
<td>0.4</td>
<td>10</td>
<td>8 x</td>
<td>21</td>
</tr>
<tr>
<td>3a</td>
<td>2.5</td>
<td>73</td>
<td>29</td>
<td>1.8</td>
<td>8 x</td>
<td>21</td>
</tr>
</tbody>
</table>

2) Experiment using Dynabeads® coated with immune and [pre-immune] serum IgG (DBXII)

<table>
<thead>
<tr>
<th>Round</th>
<th>input ($\times 10^{11}$)</th>
<th>output ($\times 10^{5}$)</th>
<th>yield ($% \times 10^{-4}$)</th>
<th>coating (µg)</th>
<th>wash.</th>
<th>amplif. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.36</td>
<td>0.04 [0.02]</td>
<td>0.1 [0.06]</td>
<td>5 [5]</td>
<td>6 x</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>0.26 [0.29]</td>
<td>1.8 [2.7]</td>
<td>6.9 [9.3]</td>
<td>5 [5]</td>
<td>6 x</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>1.6 [6.8]</td>
<td>0.3 [3.3]</td>
<td>1.2 [11.0]</td>
<td>5 [5]</td>
<td>3 x</td>
<td>100</td>
</tr>
</tbody>
</table>

**Table legend:**

1 = Input p.f.u.; amplified phage-display library for round 1  
2 = Output p.f.u. (total number of recovered phage prior to amplification)  
3 = Output expressed in percent of input phage  
4 = Concentration of IgG used to coat the dishes or beads  
5 = Washing conditions before elution of bound phage (TBS/Tween 20 0.05%)  
6 = Percentage of eluate taken for amplification; note that for round 1 of BPIII the eluate was reduced to 100 µl by spinning on microsept filters, for following rounds only 100 µl equal to 21% were amplified  
6a (round 3 for BPIII was performed after the "PL+S method", i.e. the amount of IgG indicated refers to the amount incubated in solution with the amplified eluate from round 2)  
[ ] = values for experiment DBXII with pre-immune serum IgG
Table II.9: pIII-fusion peptides of 75 phage analysed from round 3 of BPIII (immune serum IgG, A) and DBXII (immune (B) and pre-immune (C) serum IgG); a (b) sequence occurred twice (four times)

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>no group</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>TRVPRR(^a)</td>
<td>MSYLEG</td>
<td>RWFHRH(^a)</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>SRLPLR(^b)</td>
<td>LIYLVN</td>
<td>YLRWMY</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRLPKR</td>
<td>GYLVN</td>
<td>PMWMIG</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RWSDMA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>SRVPLT</td>
<td>ARYLMY</td>
<td>ARWPML(^a)</td>
<td>ARLNY</td>
</tr>
<tr>
<td></td>
<td>SRFPWS</td>
<td>SYLFSN</td>
<td></td>
<td>ARFVNS</td>
</tr>
<tr>
<td></td>
<td>TRMPFF(^a)</td>
<td></td>
<td></td>
<td>ARLKF</td>
</tr>
<tr>
<td></td>
<td>TRMFER</td>
<td></td>
<td></td>
<td>ARYLMY</td>
</tr>
<tr>
<td></td>
<td>TILPKR</td>
<td></td>
<td></td>
<td>ARWPML(^a)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>none</td>
<td>TRYVYL</td>
<td>none</td>
<td>ARLLLS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
II. Results and Conclusions

b) sequence analysis of pIII-fusions of isolated phage from round 3
The most obvious way of analysing phage recovered from the biopanning procedure was to sequence the phage genome of various isolated phage in the region of the random DNA insert and translate the nucleotide sequence into the pIII-fusion peptide displayed on the phage. The colonies chosen for template preparation were picked at random thereby giving an unbiased overview over the population of phage present after three rounds of selection. Table II.9 shows 75 phage analysed for BPIII and DBXII.

Some of the peptide sequences obtained from BPIII share sequence motifs which could be divided into three groups with:

**Group 1)** Spanning 6 aa featuring serine or threonine at position 1, arginine at position 2, proline at position 4 and arginine at position 6 [S/TR-P-R].

**Group 2)** Featuring tyrosine followed by a leucine [YL].

**Group 3)** A motif with RW, WM or RWM pattern [RWM].

Four phage occurred more than once with two of them being from group 1.

Representatives of groups 1)-3) were also present in DBXII:

**Group 1)** Four out of six phage (67%) had threonine instead of serine at position 1 (compared with only 29% for BPIII) and position 6 (R) was only conserved in 33% of these phage (compared to 100% in BPIII).

**Groups 2)** and 3) were also not as well represented as in BPIII with only one sequence fitting the pattern involving tryptophane (group 3) and two fitting the YL motif (group 2).

**Group 4)** A new rather abundant group could be created with the general pattern A followed by R and in four out of six cases (67%) followed by an aromatic residue.

Two of the three sequences that appeared several times could be arranged in group 1) and 4) indicating an enrichment of certain sequence motifs.

If the short linear motifs of groups 1-4) represented crucial residues for recognition by antibodies directed against HBsAg some of them could probably be aligned with...
the primary sequence of this polypeptide. The results of this alignment were as follows:

Table II.10: Alignment of identified group motifs with amino acid sequence of HBsAg

<table>
<thead>
<tr>
<th>Group of pIII-fusions</th>
<th>motif can be aligned with:</th>
<th>non-identical matchesa</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>SR-P-R, no alignment</td>
<td>7R-P-R</td>
</tr>
<tr>
<td></td>
<td>S--P-R, position 117</td>
<td>T--P-R</td>
</tr>
<tr>
<td>2)</td>
<td>YL, no alignment</td>
<td>-</td>
</tr>
<tr>
<td>3)</td>
<td>RWM, position 73</td>
<td>-</td>
</tr>
<tr>
<td>4)</td>
<td>ARF, position 168</td>
<td>ARY, ARW</td>
</tr>
</tbody>
</table>

a sequences from the same group with related but non-identical amino acid substitutions (italic)

Interestingly, the two most abundant groups 1) and 4) matched well with short linear regions of HBsAg. The 3 amino acid match of the RWM motif of group 3) was also interesting since both experiments had overlapping pIII-fusions with either RW or WM in this area (see also Fig.II.8, p. 71).

Did sequences obtained from DBXII using pre-immune serum IgG also fall into groups 1-4)?

No group 1) or group 3) motifs were found and AR of group 4) and YL of group 2) were each found once among the 13 phage analysed. The probability of a given dipeptide occurring by chance is 1 in 400. Since each hexapeptide can represent five adjacent dipeptides the 13 analysed phage clones represent 65 dipeptides. This corresponds to about a one in six chance to find one of the dipeptides amongst even this low number of analysed phage.

Some sequence motifs were shared by a number of phage in both experiments and could be aligned with identical or similar amino acids which could indicate a directed selection for phage that required a certain motif for binding. Three of the four group motifs could be aligned with the primary sequence of HBsAg and may
therefore represent short linear parts of epitopes on HBsAg which were recognised by antibodies specific for surface antigen.

Since there are potentially many different antibodies in a polyclonal serum it seemed surprising that there should be only a few of such group motifs i.e. putative epitopes, but this could mean that many of the antibodies in the serum were "ignored" by the selection procedure maybe for conformational reasons. It is possible that the displayed hexapeptides were too short to span all residues crucial for binding or reproduce the conformational situation encountered in the native environment for many of the antibodies present in the mixture. In other words the selection procedure could have been biased towards phage that bound antibodies that recognised epitopes with short linear core sequences and possibly, but not necessarily, were abundant in the serum. The multiple occurrence of some sequences was interesting in this respect although it could not be excluded that at this stage of the biopanning procedure they were clonal siblings and not necessarily the result of specific selection.

The nature of polyclonal IgG molecules used for selection of phage here means that not only specificities against HBsAg but also against other immunogens, present in the chimpanzee before vaccination, were exposed to the phage pool. However, the experiment using pre-immune serum demonstrated that there was no significant selection for phage displaying distinct motifs as in the case of immune serum. Nevertheless, it should be noted that when immune serum was being used for biopanning the selection for phage recognising HBsAg-specific antibodies would always occur alongside the selection for phage recognising non-HBsAg-specific antibodies. It was therefore impossible to distinguish by sequence analysis and alignment between phage recognising anti-HBsAg antibodies, other antibodies or in fact sites on other proteins or reagents present in the IgG preparation.

In this respect it is interesting to note that in other experiments performed on petri dishes (data not shown) the selection procedures led to highly related sequences of the type xRHRHIH, always consisting of several histidines (H) and arginines (R) and often lysines (K). It was not clear to what these sequences bound and why they were so effectively enriched, but they were typified as background since they also occurred in experiments where different chimpanzee sera were used (F. Gray, personal communication) and never occurred in experiments using magnetic beads, implying a specific problem occurring in connection with the petri dishes. Isolation
of sequences like these when using streptavidin-coated petri dishes was also observed by others who argued the phage recognised streptavidin sites different from the biotin binding site or unspecific sites on the plastic support (G.P. Smith, personal communication). Although this could be true the fact that this problem did not occur with the streptavidin-coated beads almost eliminates streptavidin as the source of the problem. In most cases the occurrence of these sequences completely obscured the results of the biopanning experiment so that they had to be abandoned. For reasons unknown, the problem was not observed in the case of experiment BPIII although some sequences of the type discussed were still isolated (Table II.9, p.65).

II. Results and Conclusions

II.13.2) Sequence comparison and alignment of pIII-fusion peptides with the primary sequence of HBsAg

Instead of aligning the hexapeptides first with each other and then search for matches of the conserved group motifs with the primary sequence of the antigen, each hexapeptide was also aligned separately with the HBsAg sequence. This search has two main consequences: i) it includes all residues in the search and not just the ones that matched with a group motif and ii) it also included hexapeptides that did not fit in any group (Table II.9, p.65; labelled 'no group'). The analysis was also based on the assumption that phage binding an antibody directed against HBsAg could resemble parts of the epitope recognised by the same antibody on the native antigen. It was further assumed that even in the case of assembled epitopes, where parts of the epitope originate from different regions of the primary sequence, short stretches of the epitope, like two or three amino acids, may be linear. Unless a phage wholly mimicked an epitope and therefore was completely unrelated to the sequence of HBsAg, i.e. a mimotope, short linear stretches of epitopes may be identified. There also was the possibility of functionally pre-screening the phage for their ability to preferentially bind immune serum in comparison to pre-immune serum which could distinguish specific from non-specific phage. This approach will be described in section II.B.3).

Sequence comparisons were performed using the "Findpatterns" application of the GCG7 package of the University of Wisconsin. This program aligns sequences, here the pIII-fusion hexapeptides derived from analysing phage genomes, with other sequences, here HBsAg from residue 1-226 of the subtype adyw (Pugh et al., 1986).
II. Results and Conclusions

The programme thereby looks for identity of amino acids and by default only looks for absolute matches i.e. in this case six out of six residues matching with the primary sequence of HBsAg. For the reasons mentioned above it was of interest to allow a certain number of mismatches in this case three, for three amino acid matches and four, for two amino acid matches. Obviously, the total number of matches greatly increased with the number of allowed mismatches and the program did not distinguish whether matching amino acids were adjacent or interrupted by non-matching residues. Table II.11 shows the results of the computer alignments with all different pIII-fusion sequences from BPIII and DBXII.

Table II.11: Computer alignment of pIII-fusion peptides with the aa sequence of HBsAg using "Findpatterns" of GCG7

<table>
<thead>
<tr>
<th>Experiment</th>
<th>pIII-fusions (^a)</th>
<th>4aa matches</th>
<th>3aa matches</th>
<th>2aa matches</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPIII</td>
<td>27</td>
<td>1 (0.3%)</td>
<td>35 (10.5%)</td>
<td>297 (89.2%)</td>
<td>333 (100%)</td>
</tr>
<tr>
<td>DBXII (immune)</td>
<td>35</td>
<td>5 (1.2%)</td>
<td>33 (7.8%)</td>
<td>385 (91.0%)</td>
<td>423 (100%)</td>
</tr>
<tr>
<td>DBXII (pre-immune)</td>
<td>13</td>
<td>-</td>
<td>12 (6.6%)</td>
<td>169 (93.4%)</td>
<td>181 (100%)</td>
</tr>
</tbody>
</table>

\(^a\) number of total pIII-fusion peptides including multiple aa sequences

As expected, the large majority of matches were 2 aa matches, around 10% 3 aa matches and only up to ~1% 4 aa matches. The results of this computer search can be divided into three main categories:

1. Linear matches involving 3 and 4 amino acids:
These matches were either interrupted by non-matching amino acids or uninterrupted. Fig.II.8 shows in detail the alignment of all 3 and 4 amino acid matches with the primary sequence of HBsAg. At the termini of the polypeptide and in an area around residues 80-100 both experiments matched with higher frequencies, but other areas where matches accumulated were different. BPIII showed a concentration of matches around residues 110-120 an area not well represented in DBXII. DBXII sequences on the other hand accumulated mainly around residues 150-170 not represented in BPIII at all (for comparison also see Fig.II.9, p.75).
Figure 11.8
(for legend please see next page)
II. Results and Conclusions

Figure II.8:
Alignment of pIII-fusion hexapeptides from BPIII (blue) and DBXII (dark green) (Table II.9, p.65) with the primary sequence of HBsAg (adyw, Pugh et al., 1986) which match with at least 3 aa; group motifs shown in Table II.9 which could be aligned with the HBsAg sequence are shown in red (see also Table II.10, p.67); bold residues in the HBsAg sequence (purple) indicate residues which were part of a motif found in a pIII-fusion peptide; underlined HBsAg residues indicate point mutations investigated by various workers (see also I.C.6, pp.35-38); bold upper case residues in the hexapeptides indicate residues that could be matched with residues in the HBsAg primary sequence, lower case are non-matching residues; asterisks mark 4 aa matches; black lines represent peptides that have been used by various workers to investigate the immunogenicity and antigenicity of HBsAg (see also I.C.6, pp.35-38), number references according to code shown in legend table (p.73).
II. Results and Conclusions

Table legend for Figure II.8, p. 71:

<table>
<thead>
<tr>
<th>Peptide</th>
<th>aa residue&lt;sup&gt;a&lt;/sup&gt;</th>
<th>X-reactivity with anti-S serum</th>
<th>Induction of anti-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-10</td>
<td>yes (e)</td>
<td>n.d.</td>
</tr>
<tr>
<td>2</td>
<td>1-121</td>
<td>n.d.</td>
<td>yes (g)</td>
</tr>
<tr>
<td>3</td>
<td>2-16</td>
<td>n.d.</td>
<td>yes (d, f)</td>
</tr>
<tr>
<td>4</td>
<td>22-35</td>
<td>n.d.</td>
<td>yes (f), no (d)</td>
</tr>
<tr>
<td>5</td>
<td>48-81</td>
<td>n.d.</td>
<td>yes (f), no (d)</td>
</tr>
<tr>
<td>6</td>
<td>95-109</td>
<td>n.d.</td>
<td>yes (d, f)</td>
</tr>
<tr>
<td>7</td>
<td>110-137</td>
<td>n.d.</td>
<td>yes (d)</td>
</tr>
<tr>
<td>n.s.</td>
<td>110-139</td>
<td>yes (d)</td>
<td>yes (d)</td>
</tr>
<tr>
<td>8</td>
<td>117-137(c)</td>
<td>n.d.</td>
<td>yes (c)</td>
</tr>
<tr>
<td>9</td>
<td>121-226</td>
<td>n.d.</td>
<td>yes (g)</td>
</tr>
<tr>
<td>10</td>
<td>122-137(c)</td>
<td>n.d.</td>
<td>yes (c)</td>
</tr>
<tr>
<td>11</td>
<td>124-137</td>
<td>yes (e)</td>
<td>n.d.</td>
</tr>
<tr>
<td></td>
<td>124-137(c)</td>
<td>yes (e)</td>
<td>n.d.</td>
</tr>
<tr>
<td>12</td>
<td>125-139</td>
<td>n.d.</td>
<td>yes (d)</td>
</tr>
<tr>
<td>13</td>
<td>135-155</td>
<td>yes (h)</td>
<td>yes (h)</td>
</tr>
<tr>
<td>14</td>
<td>138-149</td>
<td>yes weak (i)</td>
<td>yes (i)</td>
</tr>
<tr>
<td>15</td>
<td>139-147</td>
<td>yes (a, b, e)</td>
<td>yes (a), n.d.(e)</td>
</tr>
<tr>
<td></td>
<td>139-147(c)</td>
<td>yes (b, e)</td>
<td>n.d.</td>
</tr>
<tr>
<td>16</td>
<td>139-158</td>
<td>yes (a)</td>
<td>yes (a)</td>
</tr>
<tr>
<td>17</td>
<td>140-158</td>
<td>yes (a)</td>
<td>yes (a)</td>
</tr>
</tbody>
</table>

<sup>a</sup> (c) stands for "cyclic" and indicates that the peptide was used in cyclic form; n.s. = "not shown" and refers to Fig.1; n.d. = "not determined"; reference letter codes: (a) Bhatnager et al., 1982; (b) Brown et al., 1984; (c) Dreesman et al., 1982; (d) Gerin et al., 1983; (e) Howard et al., 1988; (f) Lerner et al., 1981; (g) MacKay et al., 1981; (h) Prince et al., 1982; (i) Neurath et al., 1982.
II. Results and Conclusions

Interestingly, two uninterrupted 4 aa matches occurred, one at position 11 (DBXII) and one at position 112 (BPIII).

2. Short linear matches of two adjacent amino acids:

2 aa matches also accumulated within certain areas of the surface antigen (Fig.II.10). However, some of the dipeptides like FL and LL occur several times in the primary sequence of HBs. They not only account for a larger proportion of the total matches (e.g. each FL occurring in a pIII-fusion scores 7 matches) but areas of accumulated 2 aa matches coincided with the areas where these dipeptides occurred in HBs (not shown). Dipeptides of interest were therefore the ones that occurred in several phage but only occurred once in the sequence of HBs. Of these, TR (position 23) occurred in five pIII-fusions, LR (position 77) in four and RR (position 78) in three. However, these dipeptide-motifs not only occurred in BPIII and DBXII but also in the fewer pIII-fusions analysed from the biopanning experiment with pre-immune serum IgG (Table II.9, p.65 and Fig.II.10).

3. Non-adjacent 2 amino acid matches:

Matches of this category were the majority of 2 aa matches. There was a good correlation between the two experiments but also with the experiment with pre-immune serum IgG (Fig.II.10). However, in many cases the alignment of a 2 aa match allowed a further third residue to be identified which was maybe not identical but related with amino acids involved in a previously identified three amino acid match covering the same area. For example, the group of 3 aa matches at position 117-122 of the S--P--R type (Fig.II.8, p.71) were related to a number of sequences in BPIII and DBXII which occurred as 2 aa matches since they only matched with P--R in this region but they possessed a threonine instead of a serine at position 117 which can be considered a conservative substitution (see also phage 6, 9 and 14 Fig.II.11, p.81).

These results generated with the "Findpatterns" application can be divided in two respects:

1.)

The identification of high quality matches which in this case meant that at least three out of six linear amino acids were identical, could indicate the representation of linear stretches of HBsAg sequence on phage isolated by affinity selection with polyclonal serum IgG containing antibodies directed against HBsAg. In other words,
II. Results and Conclusions

Figure II.9: Percent of 3 and 4 amino acid matches of pIII-fusion peptides isolated from experiment BPIII and DBXII with the primary sequence of HBsAg at 10 aa resolution; figures shown on x-axis mark the last residue of the relevant decapeptide, i.e. "10" means residues 1-10.

Figure II.10: Percent of 2 aa matches of pIII-fusion peptides isolated from BPIII and DBXII (including pre-immune serum) with the primary sequence of HBsAg at 10 aa resolution; figures shown on x-axis mark the last residue of the relevant decapeptide, i.e. "10" means residues 1-10.
II. Results and Conclusions

The pIII-fusion sequences of this category could have functioned as target sequences for antibodies that *in vivo* would recognise an similar linear region of the surface antigen by imitating or mimicking parts of this native epitope. In the case of mimickry the sequence could be entirely different from the native epitope recognised on HBsAg and a linear alignment could not identify such mimotopes. It cannot be excluded that some hexapeptides with different amino acid sequences that could not be aligned at all or only in different sequence areas, in fact mimicked the same epitope. Binding of phage like this would occur on the basis of structural similarity rather than sequence identity.

Interestingly, both experiments selected for phage that matched with three or four amino acids and sometimes overlapped with the same matching amino acids, as for example at position 73-75 where two hexapeptides from DBXII matched with xWMCxR and xRWxxL and one from BPIII matched with xxRWMx, overlapping with both other matches (Fig.II.8, p.71). These residues were not always in a continuous arrangement but residues that are crucial for binding antibodies do not necessarily occur in a continuous array as the findings for the linear epitope recognised by MAb 18/7 clearly showed (chapter II.A).

There were differences between the areas of higher frequency matches between the two experimental approaches as already mentioned. BPIII showed concentration in the area 110-120 whereas the most striking concentration in the case of DBXII was in the area 160-170 with six different pIII-fusions matching within three amino acids (168-170; see Fig.II.8, p.71).

2.)

The object of analysing 2 aa matches was that they might provide a way to identify areas within the antigen which contribute short linear parts to conformational epitopes. Apart from the assumptions outlined in the introduction of this chapter it was also believed that despite these matches being extremely short, the phage could still efficiently bind the antibody during the biopanning experiment. The latter could perhaps be justified in the light of the results obtained for MAb 18/7 (chapter II.A) where as few as three conserved residues were sufficient for specific and high-affinity binding of the MAb, but binding efficiency also depends on the environment of the conserved dipeptide in the displayed hexapeptide since the neighbouring amino acids could to some extent support the binding without being necessarily
II. Results and Conclusions

identical to the native environment in the epitope. Further, the stringency conditions during the biopanning experiment would also influence this class of interaction and in this respect DBXII could have been favouring the identification of weaker ligands since there was less competition for binding sites, although there was little evidence for that.

The experiment with pre-immune serum IgG (DBXII) showed a similar pattern thus making conclusions about the distribution itself and the meaning of 2 aa matches very difficult. As already mentioned, the 2 aa matches were valuable in areas where there were already higher quality matches and where additional pIII-fusion could be aligned which included the two matching amino acids and another related residue thereby supporting the conserved motif in this area. This was mainly the case in area 117-122 (STGPCR) which was identified previously by aligning the consensus motif of group 1) pIII-fusions with the HBsAg sequence (Table II.10, p.67).

II.3) Phage that preferentially recognise immune serum IgG share common sequence motifs and match in the same area of HBsAg

Analysis of the pIII-fusion peptides of phage picked at random from round 3 of the biopanning experiments gave a representation of the population of phage present after affinity selection against polyclonal serum IgG (II.B.1). The peptides could be aligned with each other or with the sequence of the antigen the antibodies were raised against (HBsAg) and matches with the primary sequence could indicate potential binding sites for antibodies (II.B.2). However, the situation encountered when using polyclonal serum IgG as substrate was very complex and analysis and alignment of sequences alone was not sufficient to predict areas of potential epitopes with great confidence. Although the sequence comparisons in II.B.2) gave an indication of where some of these areas might be located within HBsAg it did not prove that the phage identified were mimicks of these particular areas or indeed any other region of HBsAg. It should be stressed that the complete standard repertoire of the chimpanzee was present throughout the whole selection procedure and phage could therefore also mimic epitopes recognised by non-HBsAg specific antibodies.

In this respect it was important to identify phage in the pool of affinity selected phage that preferentially bound specificities contained in the IgG repertoire of the
II. Results and Conclusions

immunised animal used for selection and not with the repertoire of the same animal before immunisation. If most of the phage would have been selected against specificities also present in the serum before immunisation, one would not expect such differential binding behaviour since both sera would probably contain a similar representation of non-HBsAg specific antibodies.

**Table II.12: Selection procedure for 40 individual phage based on preferential binding to wells coated with immune serum IgG**

<table>
<thead>
<tr>
<th>Experiment:</th>
<th>BPIII</th>
<th>DBXII</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Micropanning:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indiv. phage (input)</td>
<td>1-20 (20)</td>
<td>21-40 (20)</td>
</tr>
<tr>
<td>positive</td>
<td>1, 2, 4, 5, 6, 7, 9, 14, 19 (9)</td>
<td>21, 26, 27, 31, 36, 37, 38 (7)</td>
</tr>
<tr>
<td><strong>ELISA:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>input</td>
<td>1, 2, 4, 5, 6, 7, 9, 14, 19 (9)</td>
<td>21, 26, 27, 31, 36, 37, 38 (7)</td>
</tr>
<tr>
<td>positive</td>
<td>2, 4, 5, 6, 7, 9, 14 (7)</td>
<td>26, 31, 37, 38 (4)</td>
</tr>
<tr>
<td>sequence analysis on:</td>
<td>2, 4, 5, 6, 7, 9, 14 (7)</td>
<td>26, 31, 37, 38 (4)</td>
</tr>
<tr>
<td>number eliminated:</td>
<td>(13)</td>
<td>(16)</td>
</tr>
</tbody>
</table>

*figures shown in bold indicate phage finally selected for sequence analysis, figures in parenthesis in second columns indicate the relevant total number of phage

Twenty new colonies, not previously analysed, were taken at random from round three eluates of biopanning experiments BPIII and DBXII and amplified in overnight cultures. Two consecutive elimination procedures were employed using different methods for detecting bound phage. The first assay was adapted from the micropanning procedure of Smith and Scott (1993). Phage supernatants of the overnight cultures were diluted 25-fold serving directly as input. After incubation on wells coated with IgG molecules of preparations from immune or pre-immune serum, wells were washed and bound phage eluted and spot titred on agar plates.
II. Results and Conclusions

Phage were considered positive if they showed fewer than half the number of colonies in the eluate from the pre-immune IgG-coated well in comparison to the eluate from the immune IgG-coated well.

Phage clones selected by this method were subjected to a second selection procedure featuring a different detection system. Bound phage were not eluted after washing but detected by ELISA using anti-fd phage rabbit serum ("Indirect Phage Sandwich ELISA"). This second selection step led to exclusion of further phage from the initially 20 clones. Selection criterion for the ELISA was the same as for micropanning i.e. an at least twice as strong signal for phage binding to wells coated with immune IgG in comparison to wells coated with pre-immune IgG. Table II.12 illustrates the entire selection procedure for the 40 phage clones. Phage regarded as positive after the micropanning and ELISA selection procedure were analysed by sequencing and the pIII-fusions they display compared:

Table II.13: pIII-fusion peptides of phage identified by selective binding procedure from clones derived from round 3 of BPIII and DBXII

<table>
<thead>
<tr>
<th>Experiment: BPIII</th>
<th>phage no.</th>
<th>pIII-fusion</th>
<th>found previously?*</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>SRLPKR</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>SRRPSV</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>SRLPKR</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>TRVPRR</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>SRLPKR</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>TRVPRR</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>TRVPRR</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experiment: DBXII</th>
<th>phage no.</th>
<th>pIII-fusion</th>
<th>found previously?*</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>VLLRST</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>SKPPTI</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>PYNLDV</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>GRMPTR</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

* the symbols indicate whether the same sequence had been found before when phage clones from round 3 were chosen at random for sequence analysis (Table II.9, p.65)

The seven phage analysed from BPIII consisted of only three different phage species with six of the phage displaying either SRLPKR or TRVPRR peptide fusions (Table II.13). All seven had a common motif with serine (S) or related threonine (T)
residues at position 1, proline (P) at position 4 and, in six out of seven phage, arginine (R) at position 6 (S--P-R motif).

In contrast the phage isolated by the same procedure from DBXII were completely new peptide sequences and did not show a comparable conservation of residues or positions, although phage 31 and 38 showed some similarity with S--P or P-R motifs, respectively.

How do these sequences isolated by selection procedure align with the amino acid sequence of HBsAg? Equivalents of the dominating phage species from BPIII (phage clones 2, 5, 6, 7, 9 and 14) were also found previously when clones were selected at random from round 3 eluate (Table II.9, p.65). Accordingly, SRLPKR could be aligned with three amino acids involving S₁₁₇, P₁₂₀ and R₁₂₂ (Fig.II.11). Although TRVPRR did not match with three amino acids anywhere over the entire HBsAg sequence it could also be aligned with the same region as SRLPKR involving P₁₂₀ and R₁₂₂ but a related threonine instead of serine at position 117 (see also "non-adjacent 2 aa matches" II.B.2, p.74). The third phage, SRRPSV, shared the S--P motif and could therefore also be aligned in this area matching with S₁₁₇ and P₁₂₀. In this case however, the arginine was not conserved and replaced by a valine (Fig.II.11).

Of the more diverse sequences obtained from DBXII, SKPPTI (31) and GRMPTR (38) could also be aligned in the area 117-122 since these hexapeptides showed S--P or P-R motifs, respectively (Fig.II.11). Additionally, SKPPTI matched with four amino acids at residues 64-68, and VLLRST and PYNLDV matched with three amino acids in other areas of HBsAg (positions 96 and 29, respectively; not shown).

The pool of affinity selected phage was screened functionally by a combination of different binding assays to identify phage that preferentially bound to immune serum IgG. It was assumed that phage that recognised both repertoires equally well were probably not specific for anti-HBs antibodies. However, it cannot be excluded that these phage were mimicking extremely rare epitopes and that the frequency of these epitopes, even in a serum with high anti-HBs titres, was so low that the signals obtained with the binding assays employed were close to background. Nevertheless, the selection procedure led to the exclusion of the majority of phage and the remaining phage were analysed by sequencing. Strikingly, in the case of BPIII, the
selected phage showed predominantly two hexapeptide sequences with a common consensus motif (Table II.13, p.79). When aligned with the HBsAg primary sequence the hexapeptides matched in the same region and, more importantly, with the same amino acids (Fig.II.11). Both sequences were also detected when the eluate was examined at random (II.B.1) but did not occur at the same frequency (Fig.II.9, p.65).

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Phage</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td></td>
<td>PGSSTTSTGPRCRTCCTTPAQG</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SrlPkr phage 2/5/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SrrPsv phage 4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TrvPrr phage 9/6/14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SkpPti phage 31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gmpPrt phage 38</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SGSTTP phage 10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>STGPRCRTCmT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CrTCahPgeha</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TsntahCrTCsnPsr</td>
<td></td>
</tr>
</tbody>
</table>

Figure II.11: Amino acid sequence of HBsAg (bold; adyw subtype) within the borders indicated aligned with pIII-fusion peptides of phage isolated by selective binding experiments (Table II.13, p.79); upper case bold indicates residues identical to HBsAg (adyw) sequence, upper case related aa and lower case different aa; underlined residue C124 was shown to be crucial for recognition by subtype-specific polyclonal serum (Ashton-Rickard and Murray, 1989a); phage numbers relate to Table II.14; a linear epitope for HBs-specific MAb A15A7 identified using overlapping hexapeptides (PepScan; Coursaget et al., 1991); b phage clone isolated from phage display-library recognised by specificities contained in sera of hepatitis B patients (Folgori et al., 1994). c phage clone isolated from phage-display library which mimicks the epitope recognised by MAb RFHBs6 (Motti et al., 1994); note: R122 of b and c is shown as a non-matching residue here since it was not matching with the antigen sequence of the subtype used in the quoted references.

The selected phage resembled a linear stretch of amino acids of the HBsAg sequence and could therefore represent a linear part of an epitope recognised by antibodies present in the serum of the immunised chimpanzee. The fact that the same residues were conserved in both phage families could indicate their crucial function in binding these antibodies.

Among the four phage analysed from DBXII two phage shared parts of the S-P-R motif (Table II.13, p.79). This could indicate that phage representing the same
II. Results and Conclusions

epitope were also selected in DBXII but were perhaps more "diluted" by the occurrence of other, possibly lower affinity ligands.

In addition to the synthetic peptide studies mentioned in the Introduction (p.34) and shown in Fig.II.8 (p.71), the area around residues 110-130 of HBsAg has recently been in the focus of other workers in connection with possible linear parts of immunogenic determinants. Coursaget et al. (1991) and Motti et al. (1994) identified regions recognised by monoclonal antibodies specific for the short (S) surface antigen, using overlapping hexapeptides (PepScan) or phage-display techniques, respectively. The peptides binding to MAb A15A7 (Couroucé et al., 1983) encompassed residues 117-126 whereas the phage that mimicked the epitope for MAb RFHBs6 (Waters et al., 1992) could be aligned with residues 115-127 (Fig.II.11). In the latter study the matching residues could indicate amino acids crucial for binding the antibody. In an approach similar to the one chosen for this project, Folgori et al. (1994) screened a nonapeptide phage-display library with human sera from hepatitis B patients and found four phage clones that were only recognised by sera from patients but not from healthy individuals. One of the pVIII-fusion phage isolated matched with four amino acids within a linear stretch of HBsAg in the region 121-127 (Fig.II.11). The three other phage clones had no significant matches with HBsAg and were probably mimicking a different discontinuous epitope (Meola et al., 1995).

The idea of a linear epitope in this region is also interesting in the light of a number of models proposed for the structure of this region of HBsAg in which the cysteine immediately following the identified motif (C<sub>124</sub>) forms a disulphide bridge with one of the two possible cysteines further downstream (137 or 139) generating a loop exposing conformationally sensitive epitopes (Carman, 1994; see Fig.I.6, p.37). P<sub>120</sub>, the central amino acid in the S--P-R motif, could be the site of a proline-induced turn of the peptide chain and would therefore make the area an accessible epitope with linear features (Ogata, 1994). This could explain why this epitope was preferentially isolated from the diverse mixture of specificities present in the polyclonal serum. Proline restricts the range of possible conformations of a peptide due to the rigid cyclic structure of its side chain. The limitation to fewer possible conformations could have biased the selection not only for linear rather than discontinuous epitopes, but also for phage (i.e. epitopes) mimicking proline-containing epitopes.
II. Results and Conclusions

Interestingly, an amino acid change from $P_{120}$ to $Q$ was reported in a subpopulation of viruses in a chimpanzee inoculated with serum from a mutant virus strain infection (Ogata, 1994). It was thought that the mutation had arisen anew in the chimpanzee, but the biological effects of this mutation were not further investigated.

II.B.4) Binding studies with individual phage that preferentially recognise immune serum IgG

Phage identified by micropanning and Indirect Phage Sandwich ELISA from experiments BPIII (Table II.13, p.79) were very similar and matched with positions 117-122 of the amino acid sequence of HBsAg. This region thus represented a potential candidate for a linear part of possibly larger conformational epitope recognised by antibodies present in the polyclonal immune serum of lanthe and it was the aim of these binding experiments to focus in more detail on this particular region.

Table II.14: Phage chosen for Indirect Phage Sandwich ELISA binding experiments with immune and pre-immune lanthe serum IgG

<table>
<thead>
<tr>
<th>phage number</th>
<th>pIII-fusion peptide</th>
<th>pre-selected by</th>
<th>matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>TRVPFR</td>
<td>ELISA</td>
<td>3aa, 117-122</td>
</tr>
<tr>
<td>7</td>
<td>SRLPFR</td>
<td>ELISA</td>
<td>3aa, 117-122</td>
</tr>
<tr>
<td>10</td>
<td>SGSSTP</td>
<td>not pre-selected</td>
<td>4aa, 112-115</td>
</tr>
<tr>
<td>E1-8</td>
<td>PVLVLT</td>
<td>-</td>
<td>negative control</td>
</tr>
</tbody>
</table>

note: bold marks residues that match with the HBsAg primary sequence with identical or related amino acids

The phage chosen for the binding experiments are shown in Table II.14. Phage 6 and 7 were representatives of the two groups of phage identified after the selection procedure described in II.B.3. Phage 10 was also derived from experiment BPIII but not identified by functional analysis i.e. not on the basis of its selective binding ability for immune serum IgG, but was chosen at random from the colonies derived from round 3 of biopanning and turned out to match with four adjacent amino acids in a region just upstream of the potential epitope area (112-115; Fig.II.11, p.81). The
II. Results and Conclusions

object of including this phage was to examine whether the mere occurrence of such a high degree of identity with a linear stretch of HBsAg sequence was reflected in an increased affinity for immune serum IgG. If this was not the case it would further support the argument that functional analysis is preferable for selecting specific phage and that sequence homology alone does not identify specific phage.

The experiment was set up as an Indirect Phage Sandwich ELISA (IPS-ELISA) as in II.B.3 with either purified (by CsCl-gradient centrifugation) or partly purified (by PEG precipitation) phage particles. Experiments with a reversed set-up i.e. the wells of the ELISA plates coated with phage particles and the IgG molecules added and subsequently detected by anti-human alkaline phosphatase-conjugated antibodies (Motti et al., 1994) were not successful, giving extremely low signals and not showing differences between various phage and various sera (data not shown). This was probably due to the fact that the ligands used here were polyclonal antibodies and the concentration of available specific binding sites possibly quite low. Since the pIII-fusions, in contrast to the pVIII-fusions used by Motti et al. (1994), only occur on one tip of the phage particle, orientation is also an important aspect and given that the binding to the wells was through unspecific adsorption to plastic the number of accessible copies of pIII-fusions could have been even lower.

Phage ($10^{10}$ t.u.) were incubated overnight at 4°C on wells coated with purified immune or pre-immune serum IgG, the wells washed and bound phage detected by ELISA as described. Four independent experiments were performed with each sample in duplicate. Fig.II.12A shows the signals derived from wells coated with immune serum IgG in comparison to the signals of the same phage derived from wells coated with pre-immune serum IgG. The average signal obtained for the control phage on wells coated with immune serum IgG was just under 2-fold higher than the signal derived for the wells coated with pre-immune serum IgG. Phage 10 showed even smaller differences than E1-8 and was hardly elevated over the signal from wells coated with pre-immune serum. The two phage isolated previously on the basis of having elevated signals on wells coated with immune serum IgG (II.B.3)
II. Results and Conclusions

Figure II.12 A: Binding experiment involving phage 6, 7, 10 and control phage E1-8 and wells coated with immune serum and pre-immune serum IgG; columns indicate how many fold the signal on the immune IgG-coated wells was elevated over the signal obtained for the same phage on wells coated with pre-immune serum IgG; x-axis crosses at the value obtained for the control phage (E1-8 = 1.92-fold increase) to show the true variations obtained for the other phage; values shown are the mean of three experiments.

Figure II.12 B: Diagram shows the same binding experiment as II.12 A, however, columns show the percentage of the ELISA signal obtained on immune and pre-immune serum-coated wells in relation to the signal obtained for phage 6 on wells coated with immune serum IgG (value set at 100%).
II. Results and Conclusions

were 3.5 and 4.7-fold increased over the signals obtained from wells coated with pre-immune serum IgG. If the signal obtained for E1-8 was taken as a reference, as the base line in Fig.II.12A indicates, phage 6 and 7 had 2.7 and 1.5-fold increased signals on the wells coated with immune serum IgG, respectively. In Fig.II.12B all ELISA signals were related to that obtained for phage 6 on wells coated with immune serum IgG. This not only shows the relative differences, but also the absolute signals obtained for the phage. All signals from wells coated with pre-immune serum IgG were around 25-30% of the signal obtained for phage 6 on immune serum IgG-coated wells but signals for the wells coated with immune serum IgG differed considerably. Phage 6 and 7 both bound immune serum IgG with similar efficiency but the control phage and also phage 10 only reached about half of this level (61% and 45%, respectively). The difference of the ELISA signals between pre-immune and immune serum coated wells was therefore smaller for the control and phage 10 (46% and 31% difference, respectively).

Given that phage 6 and 7 were identified according to their ability to preferably recognise immune serum IgG the results obtained here were reassuring, for they confirmed that the phage that had been the major species isolated by the two step selection procedure (II.B.3) were indeed recognising the immune serum significantly more efficiently than the same phage recognised pre-immune serum of the same animal or an unrelated control phage. This demonstrated that these phage were specific for antibodies contained in the serum only after vaccination and challenge. Further, this experiment emphasised that a functional selection of phage on the basis of selective binding is important when complex populations are obtained from biopanning which may contain many ligands to unrelated specificities. The actual concentration of certain specificities within the polyclonal serum was of course unknown but since both sera were used as purified IgG preparations at least the amount of IgG molecules used for coating the wells was similar since otherwise there would be a difference in the concentration of IgG molecules between a neutral and a boosted serum.

It is possible that the control phage coincidentally mimicked epitopes recognised by antibodies present in the immune sera and therefore bound more tightly than expected, which might explain why E1-8 recognised immune serum IgG almost twice as efficiently as pre-immune serum. Phage 10 matched the primary sequence of HBsAg with high homology and thus was a good candidate for a linear epitope, but
did not preferentially recognise components in the immune serum and therefore did probably not represent a putative epitope or even parts of it. It could not be excluded however that phage 10 did recognise antibodies specific for HBsAg but then this reactivity could not be demonstrated with this assay.

Interestingly, although only phage 7 features the native serine residue (S_{117}) in comparison to a related threonine for phage 6 there was not much difference between the efficiencies of the interaction with immune serum IgG. This might support the argument already advanced that this exchange of amino acids does not effect binding and is a conservative substitution.

II.B.5) Competing for binding of phage 6 and 7 to immune serum IgG with synthetic peptides

Although the two dominating phage sequence motifs isolated from BPIII (II.B.3) could be aligned with area 117-122 of HBsAg they still showed considerable amino acid exchanges with the native sequence in the non-matching positions (Fig.II.11, p.81). To test whether it was this particular linear region of HBsAg which was mimicked by the phage, the effects of a free hexapeptide representing this region on the interaction of phage 6 and 7 with immune serum IgG was examined. This would demonstrate that, if competition occurred, the interaction was specific for this particular target sequence and not unspecific, in which case adding the peptide would not necessarily affect binding. A peptide with an unrelated sequence served as a negative control and comparison with a peptide representing one of the phage displayed motifs was helpful for assessing effects observed with the "native" hexapeptide (P3).

The assay was performed as a competition ELISA and for each phage examined two sets of wells were coated with immune serum IgG and incubated with fixed amounts of purified fusion phage in the presence or absence of synthetic peptides. Experiments with peptide concentrations lower than 1 mM proved unsuccessful and had no effects on binding of the phage (data not shown). Wells were washed and bound phage detected by anti-fd serum and secondary alkaline phosphatase
Figure II.13: Competition of the interaction of phage 6, 7 and E1-8 with wells coated with immune serum IgG by adding synthetic peptides P1, P3 or P5; results shown are the average of three experiments with the exception of P5 competition of phage 7 and E1-8; the ELISA signal obtained in the absence of peptides is set at 100% in each case and represented by the thick line.
conjugated anti-rabbit antibodies as before. Fig.II.13 displays the average of results obtained from three experiments.

Control peptide P1 slightly reduced binding for all three phage but to very similar degrees (91-95% compared to the absence of peptide). P3, the linear HBsAg sequence 117-122, had inhibitory effects on phage 6 and 7 reducing binding by 13% and 14%, respectively. Interaction of the control phage was not reduced by adding P3. P5, representing the hexapeptide displayed by phage 6 had a weak effect on phage 6 but not on phage 7 or E1-8. Competition experiments with a peptide (P6) representing another frequently isolated phage sequence motif which was almost identical to the pIII-fusion of phage 7 except for one amino acid exchange (SRLPLR instead of SRLPKR), showed inversed effects on binding i.e. interfered only with phage 7 but not with phage 6 or E1-8 (data not shown).

The effects of various peptides on phage binding were quite weak and required high peptide concentrations, but since this assay was complex and required manifold amplification of the signal possibly due to very low concentrations of available binding sites it probably did not allow for larger effects. A further reason could be that the free peptides bound less strongly than the same peptide in a pIII-fusion environment, possibly reflecting contributing effects from neighbouring residues (Felici et al., 1993; Yayon et al., 1993). In this respect recent work by Meola et al. (1995) emphasises the importance of the phage context for correct presentation of mimotopes and raises doubts about looking at mimotopes as isolated peptide sequences outside their, in this case, phage context.

Several experimental aspects remain speculative, for example, how many antibodies that recognised the epitope postulated here were actually available for binding on the solid matrix?

It could not be excluded that adding synthetic peptides at such high concentrations interfered with the interactions in a sequence-independent way but since control phage and control peptide were included findings could be related to each other and led to the following conclusions:

- The control peptide did have a slight competitive effect but it affected all phage including the control phage and was comparatively small (5-9%) despite the high concentrations.
II. Results and Conclusions

- Inhibitory effects of the peptide representing the native HBsAg sequence (P3, STGPCR) were only observed for phage 6 and 7 but not for the control phage.
- The peptide representing the peptide motif of phage 6 (P5, TRVPRR) had a weak effect on phage 6 but non on phage 7 or the control phage.

Although the results obtained here varied considerably between different experiments and maybe regarded as trends rather than definite results, they support phage 6 and 7 representing linear target sequences for HBsAg-specific antibodies that cross-react with the phage and the related native sequence as a free peptide. In comparison with the corresponding area of HBsAg the phage displayed sequence motifs had several amino acid exchanges (Fig.II.11, p.81) which could mean that the interaction was relatively weak.

Free peptides representing the phage-displayed sequences competed only with their pIII-fusion counterpart (e.g. P5 with phage 6) but not with phage of the other dominant sequence family (e.g. P5 with phage 7). Perhaps peptides P5 (TRVPRR) and P6 (SRLPLR, not shown) were both equally well mimicking the native epitope but not very efficiently each other, particularly when one partner was a free peptide and the other a phage fusion as it was the case in this assay and could explain why there was no competition between the two sequence families (SRLPKR or TRVPRR) in this assay.
CHAPTER III: DISCUSSION
Phage-display libraries have become a widely used tool and found their way into many areas of investigation of macromolecular interactions. Besides the advantage of being easy to use and requiring minimal expenditure, the key factor for the success of the phage-display technology is the possibility that it offers to probe for peptide ligands to almost any molecule, be it a whole protein, a protein domain or even nucleic acids, without requiring any information about the molecule used for screening. More complex targets such as organelles, viruses or nucleoprotein complexes may also be probed in a similar way. Many ligands for a whole range of molecules have been successfully isolated, many of them being specific for antibodies (see Introduction I.A.6, p.12).

As mentioned in the Introduction, screening with antibodies allowed in many cases identification of the epitopes recognised by these particular antibodies since the isolated peptide ligands showed amino acid sequences similar to the primary sequence of the antigen against which the antibodies were raised.

**Epitope mapping studies on MA18/7**

In this project a relatively well characterised monoclonal antibody (MA18/7) which specifically recognises the PreS1 domain of the HBV surface antigen was used to screen a random hexapeptide phage-display library. The epitope recognised by MA18/7 had been identified previously by means of other epitope mapping techniques which relied on overlapping synthetic peptides (Coursaget et al., 1991; Howard et al., 1991) or analysis of deletion mutants expressed on bacteriophage (Sominskaya et al., 1992). By mapping the epitope recognised by MA18/7 with a random phage-display library one could directly compare the results achieved with these different strategies. Based on the results of the previously applied mapping techniques it was expected that MA18/7 would recognise a linear sequence of amino acids between 17-25 of PreS1 and probably would not require complex conformational structures for efficient binding (Fig.II.1, p.46).

**Three non-contiguous residues were highly conserved amongst isolated phage**

Phage isolated from biopanning rounds 1, 2 and 4 carried peptide sequences that matched within area 18-24 of PreS1, but only a few residues within this area were conserved, in some cases as few as three. In particular, the motif D20, P21 and F23 was conserved in almost all of the phage analysed whereas neighbouring residues of the
PreS1 sequence were found only between 3% and 48% of all phage analysed (Fig.II.1, p.46).

Although the epitope would have been located clearly after round 1 of biopanning since about 50% of the phage analysed from this round showed a striking similarity to the PreS1 sequence, it was of interest to proceed with further rounds of selection in order to ascertain whether the pattern of conserved residues displayed by the isolated phage would change. For example, would further neighbouring residues be found increasingly frequently amongst the phage population derived from later selection rounds?

After round 2 the population of phage did change in the respect that all phage analysed now matched with this motif from PreS1 indicating a strong selection for these residues. Although the pattern of the conserved residues did not change significantly, residue L9, neighbouring the core motif DP-F, was found more frequently than after round 1 indicating some selection for this residue. Other neighbouring residues of the core motif were found in only two cases in total (Fig.II.1, p.46).

Interestingly, the three most conserved core residues D20, P21 and F23 were not contiguous and amino acid exchanges of the intervening residue were tolerated although only amino acids with small side chains were found. Position 22 may therefore itself not be involved in the interaction with the paratope of MA18/7, but the steric situation may require a small amino acid at this position for the interaction to be strong enough to persist under the experimental conditions used here.

• Relative dissociation constants of the interaction between phage and MA18/7 were in the nM range

The above findings were supported by quantitative examination of the interaction of individual phage with MA18/7 using an assay that was developed on the basis of an immunoassay procedure described by Friguet et al. (1985) and which yielded relative dissociation constants for this interaction. The assay developed here was based on phage titre and allowed the individual phage that bound to MA18/7 to be ranked according to their relative dissociation constant $K_D^\text{rel}$. Comparison of the values obtained provided evidence that, i) the three core residues D20, P21 and F23 were sufficient for an efficient interaction between phage and MA18/7 with a $K_D^\text{rel}$ in the
III. Discussion

nM range, ii) the exchange of the conserved F23 residue by another aromatic residue (Y) had only a small effect on affinity and iii) exchange of the intervening A22 and neighbouring residues had practically no effect on the strength of the interaction (Table II.6, p.56).

- **Synthetic peptides competed effectively with phage ligands specific for MA18/7**
  To prove the specificity of the isolated phage for MA18/7, synthetic peptides were used for competition studies which were also based on phage titre as the signal readout. The result demonstrated that the peptide sequences identified did not require the context of the phage particle to interact efficiently with MA18/7 and that even an extremely short peptide (P2, 4 aa) representing only the three non-contiguous core residues with the correct spacing was recognised by MA18/7 (Fig.II.7, p.60). Further, complex conformational structures were apparently not required for interaction with MA18/7 since short peptides were efficiently bound.

- **Characterising the epitope recognised by MA18/7**
  Collectively, the results demonstrated that the core of the epitope recognised by MA18/7 consisted of only three amino acids which were sufficient for a strong and specific interaction and which were arranged in a linear, but non-contiguous array. They probably make direct contact with the paratope of the antibody and therefore contribute significantly to the binding energy of this particular interaction. In this respect it is interesting to speculate that the reported cross-reactivity of MA18/7 with peptides spanning 131-137 of PreS2 (Howard et al., 1991) may be due solely to the presence of the DP dipeptide which contributes most efficiently towards the binding of this antibody. In addition to DP, the multiple sequence alignment of PreS2 subtypes presented by Howard et al. (1991) also contains a leucine (L132) with the same spacing as L19 in PreS1. Based on the results described above, a peptide carrying this leucine residue in association with the DP motif would be efficiently recognised and hence explain the observed cross-reactivity. In fact it could be argued that MA18/7 may recognise many other antigens and peptides that feature a DP dipeptide and one of the other neighbouring residues within a linear stretch of amino acids.

- **How do the mapping strategies compare?**
  By using a random phage-display library a high-resolution epitope mapping could be achieved since internal residues were readily accessible for investigation and
revealed the discontinuous pattern of the crucial residues. If similar resolution was to be achieved with for instance the PepScan method, complex substitution analysis and synthesis of many peptides would be required. Further, the ease with which individual fusion phage carrying different peptide sequences could be amplified and examined in a non-radioactive quantitative binding assay highlights the value of this biological method in comparison to approaches with synthetic peptides.

**Screening polyclonal anti-HBsAg IgG revealed consensus motifs amongst the phage ligands isolated**

The second major objective in this project was to apply the phage-display approach in a polyclonal context to provide an overview of the target sequences recognised by a mixture of antibodies directed against a complex antigen, in this case the HBV surface antigen. As explained earlier (Introduction I.A.5) there are many experimental variations for applying phage-display libraries and here two approaches, using different solid support matrices and different final rounds of selection, were pursued in parallel to examine possible influences on selection. The two methods led to two quite diverse populations of phage, but with some highly related pIII-fusion peptides from both experiments (Table II.9, p.65).

Closer analysis of the fusion peptide sequences of the phage isolated showed a few consensus motifs that were of four basic types and, with the exception of one, occurred in both experiments, although at different frequencies. The group motifs were represented by two amino acids or by three amino acids and could represent short linear elements, possibly within larger conformational structures, which were sufficient for recognition by antibodies directed against HBsAg under the experimental conditions used for the selection (Table II.9, p.65). It could not be excluded at this stage though, that the antibodies to which these motifs bound were not necessarily HBsAg-specific, but part of the standard repertoire.

Considering the potential complexity of a polyclonal serum, it seemed surprising that only four sequence consensus motifs were found but this could be explained either if the majority of antibodies in the serum were directed against similar target sequences or, alternatively, if there was a bias during the selection towards certain types of sequences that were selected preferentially. One reason for such a preference could be that, although it is considered as one of the assets of the phage-display technique that conformational epitopes can to some extent be reproduced, in practise, the short
III. Discussion

hexapeptides only poorly represent secondary conformations, and that therefore the selection was biased towards phage that were recognised by antibodies that did not require a specific conformation for efficient interaction (as for example linear type epitopes). Difficulties connected with conformational epitopes have also been experienced by others and the few successful examples of conformational determinants being revealed by screening of phage libraries underline this problem (see also Introduction I.A.6, p.12).

- Evidence for the consensus motifs being parts of epitopes recognised on native HBsAg

Three of the group motifs were represented within the sequence of HBsAg and may represent cases in which short linear elements contributed efficiently towards the binding of an antibody and were therefore selected on this basis. Group motif 1) (S--P--R) was particularly interesting since it involved three amino acids covering six residues and matched within an immunological important region, the a determinant.

The relevance of this motif was further emphasised by several other results and also by independent work by other research groups:

1. Sequence comparison of pIII-fusion peptides isolated from round 3 with HBsAg revealed several 2 aa matches within the sequence 117-122 which could be aligned with an additional non-identical, but related third residue, and because of this degeneracy these sequences were not placed previously in group 1) (Table II.10, p.67).

2. The motif S--P--R or T--P--R was present in six out of seven phage selected from round 3 of experiment BPIII (using streptavidin-coated petri dishes) on the basis of preferential recognition of immune serum IgG in comparison to pre-immune serum IgG in a two step selection procedure (Table II.12, p.78), and the motifs selected were almost identical to group motif 1) (Table II.9, p.65).

3. Phage representatives displaying these two motifs recognised immune serum 3.5-fold (phage 7, SRLPKR) or almost 5-fold (phage 6, TRVPRR) better than pre-immune serum IgG (Fig. II.12A, p.85).
4. A synthetic hexapeptide representing the sequence 117-122 of HBsAg was able to compete weakly for binding with phage 6 and 7 to immune serum IgG although high peptide concentrations were required; effects varied considerably between experiments, but control phage and control peptide were included in the experiment and allowed a trend to be established (Fig.II.13, p.88).

5. The epitopes recognised by two monoclonal antibodies specific for HBsAg were mapped to an area overlapping substantially with region 117-122 (Coursaget et al., 1991; Motti et al., 1994; see Fig.II.11, p.81).

6. Screening of several sera of HBV infected patients with a pVIII phage-display library produced one phage clone that matched with four residues between 121 and 127 of HBsAg (Folgori et al., 1993; see Fig.II.11, p.81).

- **The putative linear epitope did not induce antibodies cross-reactive with native HBsAg when injected into rabbits**

In order to provide a more extensive picture of the investigations into this putative epitope the immunogenicity of the isolated phage particles was also tested by injecting rabbits with phages 6 and 7 (data not shown). It was hoped that this would generate antibodies that would recognise native HBsAg, but the immunisation did not generate detectable levels of such antibodies, although phage particles *per se* are very immunogenic carriers (Meola et al., 1995). Similar immunisation experiments recently carried out with the pVIII phagotopes isolated by Folgori et al. (1994) showed however, that the phage clone matching with 121-127 of HBsAg, and those phage carrying mimotopes with no sequence similarity were able to induce antibodies that recognised HBsAg (Meola et al., 1995). It was also argued that although pVIII fusions probably display 10-100 times more fusion peptides per phage particle than do pIII fusions, the doses of the administered peptide did not account for the observed differences in titre between the immunisation with pVIII- or pIII-fused mimotopes.

Interestingly, the two best conserved residues, $P_{120}$ and $R_{122}$, among the phage isolated in this work were not conserved in the phagotopes (i.e. "disease related mimotopes displayed on fusion phage"; Folgori et al., 1994) isolated by Folgori et al. (1994), and $C_{121}$ was not found amongst any of the phage analysed from BPIII (see Fig.II.11, p.81). Whether these differences account for the crucial difference
observed in terms of immunogenicity is difficult to assess and the immunisation with mimotopes by Folgori et al., (1994) remains the only successful case described to date. However, the presence of the two cysteine residues in the phagotope isolated by Folgori et al. (1994) could restrict the conformation of the displayed peptide in a way that may more closely resemble this area of the antigen, and therefore be more efficient in inducing cross-reactive antibodies. Clearly, detection of residues spanning more than the six residues carried by the phage was not possible with the hexapeptide library used here.

*Distribution patterns of short amino acid matches with HBsAg; what do they mean?*

Many of the isolated and analysed sequences could not be aligned with any of the groups of sequence motifs but nevertheless matched with sequences of three or even four amino acids within the sequence of HBsAg (Fig.II.8, p.71). This individual comparison of all the hexapeptides with the sequence of HBsAg generated a distribution pattern of matches which accumulated in some areas. These accumulations could point towards the involvement of these areas in immunogenic regions including conformational determinants.

Both experiments, using either streptavidin-coated petri dishes (BPIII) or streptavidin-coated magnetic beads (DBXII), showed accumulation of three amino acid matches in the N- and C-terminal sequences, but otherwise accumulation was found in different areas for BPIII and DBXII (Fig.II.8, p.71). The differences observed between the two experiments are difficult to explain, but it cannot be excluded that individual hexapeptides, or indeed whole sequence groups, mimicked the same binding site although they were not related at the amino acid sequence level. The alignment with HBsAg alone could not answer this question and only provide a basis for speculations on the immunogenicity of certain areas and their involvement in larger epitopes. Analysis of pre-immune serum (as in DBXII) could in principle offer a way of resolving this problem, but would appear to require the analysis of many more phage clones to provide statistical relevant data.

Conclusions drawn from the analysis of two amino acid matches have to be considered even more carefully in this respect, particularly since the experiment involving pre-immune serum provided a similar pattern of 2 amino acid matches to that of the other two experiments using immune serum (Fig.II.10, p.75). Further, the
odds for a certain dipeptide to occur by chance are relatively high (1 in 80 within a hexapeptide, approximately).

The functional assay for selection of phage that preferentially interacted with specific components of the immune serum IgG pool in comparison to the pre-immune pool excluded phage that were recognised equally well by both sera and therefore were probably recognised by antibodies of the standard repertoire. Problems were encountered in setting up this assay, probably for the following reasons: i) the number, or concentration, of available binding sites for one particular phage in the polyclonal serum was potentially relatively low, thus requiring high sensitivity, which in turn resulted in high background levels; ii) the pIII fusion-phage system displays at the most five copies of the fused peptide on one tip of the phage, which is disadvantageous in two respects; first, the absolute concentration of the displayed peptides is low and, second, the correct orientation is essential to enable efficient interaction. Fusion systems using pVIII do not have these disadvantages and therefore detection assays are more sensitive and appear to deliver clearer results (Motti et al., 1994; Folgori et al., 1994; Meola et al., 1995).

Detection of bound phage required signal amplification

For these reasons a selection assay was developed based upon the detection of phage particles by anti-phage antibodies, thereby offering the whole surface of the phage as a target for interaction independent of orientation and with a high degree of amplification. These antibodies could then be detected by a second antibody conjugated with the enzyme used in the subsequent ELISA procedure to provide another level of amplification. In order to provide an additional level of selection a related selection procedure was carried out prior to the ELISA which relied on phage titre and not on enzyme-linked colorimetric detection (Chapter II.B.3, p.77; Table II.12., p.78).

The problems just outlined can be encountered equally in the ELISA-based competition assay with synthetic peptides and although alternative experimental procedures were also investigated, the assay presented here was the only successful one in distinguishing between different phage and different peptides. High peptide concentrations were required to achieve inhibitory effects; this could reflect the fact that the free peptide may not be recognised as efficiently by the antibodies as the same sequence in the context of HBsAg or even the phage. However, in a study that
also used polyclonal serum and phage-display technology to detect antibodies specifically recognising antigens connected with rheumatoid arthritis, it was also found that peptide concentrations in the mM range were required to compete for binding of serum IgG to the phage, although the arguably more favourable reversed ELISA assay could be employed here because a pVIII fusion phage library had been used (Dybwad et al., 1993).

**Overall conclusions**

A random hexapeptide phage-display library was used to study epitopes recognised by monoclonal and polyclonal antibodies directed against HBV surface antigen. Various strategies were employed for presenting the antibodies to the library including unspecific binding to plastic microwells and streptavidin-coated plastic dishes and magnetic beads. In the case of the monoclonal antibody simple unspecific coating of microwells was successful for presenting the antibodies efficiently. Presentation of the polyclonal antibodies on plastic dishes or magnetic beads in two independent experiments selected for similar phage pools but with different frequencies of the individual phage clones.

Screening with the monoclonal antibody MA18/7 demonstrated the power of the biopanning procedure and that phage specific for the paratope of an antibody can be isolated from a random pool of hexapeptides. Even by just comparing the amino acid sequences of the hexapeptides isolated it was possible to evaluate residues which were preferentially selected and therefore highly conserved. The importance of these residues which matched with a linear stretch of PreS1, was supported by quantitative binding experiments which allowed the comparison of individual phage clones in terms of binding avidity expressed by their relative dissociation constants. These values were determined by a simple assay based on phage titres which does not involve radioactive labelling, can be performed without the purchase of any special equipment and is comparatively quick.

The phage-display library was also applied to polyclonal serum which revealed some problems connected with the complexity of a polyclonal situation and the potentially low concentration of specific binding sites in such a mixture. However, the results demonstrated that sera can be successfully analysed with phage-display libraries and that interesting consensus motifs can be identified which may represent immunologically relevant areas of the antigen. Although the identified areas were of
a linear nature in this case, due in part to possible restrictions of the phage library to mimic conformational binding sites, the results generated could provide insight into the immunological properties of any antigen. This approach could be valuable in cases where the immunological areas of an antigen are of interest but where only the amino acid sequence of the antigen is known. In a complex case like the surface antigen of HBV where some areas are buried in the envelope and others are exposed to the surrounding environment, the results may contribute towards the understanding of the topology of HBsAg in the membrane. Despite failing to induce antibodies that recognised HBsAg the results obtained imply that the sequence between residues 117-122 was recognised by antibodies generated on vaccination of a chimpanzee with HBsAg and this sequence is therefore likely to be exposed on the surface. Further, since the putative binding site was identified as a short linear sequence it is possible that parts of this area are not involved in predicted complex three-dimensional structures that are part of the \(a\) determinant and which start immediately downstream of this region (see Fig.I.6, p.37).
CHAPTER IV: MATERIAL AND METHODS
IV. Materials and Methods

IV.A) Materials and Equipment

IV.A.1) Suppliers of reagents and laboratory equipment (in alphabetical order)

acrylamide gel stock solutions
bacterial culture plates
BCIP
biotin
biotinylation Kit
BSA
Coomassie blue G-250
CsCl
DE52 column matrix
Dynabeads® (M-280, streptavidin coated)
ELISA 96-well microplates
HBsAg
kanamycin
Klenow DNA polymerase fragment
Microsept spin filters
milk powder
Mini-Protean II gel apparatus
NBT
PEG
petri dishes
polyacrylamide stock solutions
pNPP
SDS
Sephatax G25 gel matrix columns
sequencing apparatuses
sequencing Primer
Sigma Plot graph plotting program
streptavidin
strip plates (polystyrene, high capacity)

Scotlab, UK
Philip Harris, UK
Boehringer, Mannheim
Boehringer, Mannheim
Amersham, UK
Sigma, USA
BioRad, USA
Boehringer, Mannheim
Pharmacia, Uppsala
Dynal, Oslo
Corning, USA
Green Cross, Japan
Gibco, UK
NBL Gene Sciences, UK
Filtron, USA
Marvel
BioRad, USA
Sigma, USA
BDH, UK
Falcon, USA
NBL Gene Sciences, UK
Boehringer, Mannheim
Sigma, USA
BioRad, USA
BRL, USA; Hybaid, UK
Oswel, Dept. of
Chemistry, University of
Edinburgh
Jandel GmbH, Germany
Gibco, UK
Costar, USA
**IV. Materials and Methods**

<table>
<thead>
<tr>
<th>Item</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEMED</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>tetracycline</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Tween 20</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>X-ray films</td>
<td>Cronex</td>
</tr>
</tbody>
</table>

**IV.A.2) Bacteriological material and solutions**

**L(Luria)-broth:**
Bacto-tryptone (10g), bacto-yeast extract (5g), NaCl (5g) were dissolved in water (1 l), adjusted to pH 7.2 and autoclaved.

**TB (Terrific Broth):**
Bacto-tryptone (12g), bacto-yeast extract (24g), glycerol (4 ml) were dissolved in water (900 ml) and autoclaved in aliquots (90 ml). Before use, autoclaved potassium phosphate buffer (2.31g KH₂PO₄ (anhydrous), 12.54g K₂HPO₄ (anhydrous) per 100 ml water; 10 ml) was added.

**L (Luria)-agar:**
As LB but with containing agar (10g /l).

**Top layer agar:**
Trypticase (10g), NaCl (5g), bacto-agar (10g) were dissolved in water (1 l) and autoclaved.

**Antibiotics:**
Tetracycline was used as a 20 mg/ml stock solution (dissolved in water and filter sterilised), kanamycin was used as a 100 mg/ml stock solution (dissolved in water and filter sterilised)

**Bacteria strains (all from G. Smith, University of Missouri, USA):**
K91: (derived from N.D. Zinder, Lyons and Zinder, 1972); male *Escherichia coli* strain for propagation of filamentous phage and therefore has an elevated number (~5) of F pili.
K91Kan: same as K91 except that a kanamycin mini element was inserted into the lacZ gene by using a transposon vector (Way et al., 1984). K91Kan cells are therefore kanamycin resistant.

Bacteriophage library:
The library was given to us by G. Smith, University of Missouri, USA and is based on the fuse5 fd phage vector (for details see I.A.4, p.8).

IV.A.3) Standard solutions and buffers

PBS (phosphate buffered saline):
137 mM NaCl, 2.7 mM KCl, 4.3 mM Na$_2$HPO$_4$, 1.4 mM KH$_2$PO$_4$ adjusted to pH 7.4.

TBS (Tris buffered saline):
50 mM Tris-HCl, pH 7.5 (25 ml of a 1M stock solution), 150 mM NaCl (15 ml of a 5M stock solution) ad 500 ml and autoclaved.

TBE (Tris buffered EDTA solution), 10x stock solution:
90 mM Tris (109g), 90 mM boric acid (55.6g), 2 mM EDTA (9.3g of disodium salt) ad 11 and autoclaved.

PEG/NaCl solution:
16.7% w/v PEG 6000 (100g), 3.3 M NaCl (117g) ad 475 ml water (~600 ml final volume), stirred until PEG has dissolved completely (heated if necessary) and stored at 4°C.
IV. Materials and Methods

IV.A.4) Antibodies and anti-sera

<table>
<thead>
<tr>
<th>Antibody/serum</th>
<th>type</th>
<th>specificity</th>
<th>provided by</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA18/7</td>
<td>monoclonal</td>
<td>HBsAg (PreS1)</td>
<td>W.H. Gerlich, University of Göttingen</td>
</tr>
<tr>
<td>Alkaline phosphatase - conjugated goat anti-rabbit antibodies</td>
<td>polyclonal</td>
<td>rabbit IgG</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Chimpanzee anti-HBsAg serum (lanthe)</td>
<td>polyclonal</td>
<td>recombinant HBsAg</td>
<td>Murray et al., (1984)</td>
</tr>
</tbody>
</table>

IV.A.5) Synthetic peptides

All synthetic peptides were synthesised by the Department of Chemistry, University of Edinburgh:

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Amino acid sequence (single letter code):</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>LDPAFR</td>
</tr>
<tr>
<td>P2</td>
<td>DPGF</td>
</tr>
<tr>
<td>P3</td>
<td>STGPCR</td>
</tr>
<tr>
<td>P4</td>
<td>STGSCR</td>
</tr>
<tr>
<td>P5</td>
<td>TRVPRR</td>
</tr>
<tr>
<td>P6</td>
<td>SRLPLR</td>
</tr>
<tr>
<td>P7</td>
<td>ALLTRILG</td>
</tr>
</tbody>
</table>
IV. Materials and Methods

IV.B) Methods

IV.B.1) Protein purification and analysis

_Purification of antibodies (IgG) from sera_
Serum samples were thawed and pasteurised at 56°C for 30 min. A 2 ml aliquot was removed from the stock and cleared by spinning briefly in a benchtop centrifuge. The total volume was adjusted to 4 ml with PBS and the IgG molecules precipitated by two consecutive ammonium sulphate precipitations as follows: The solution was constantly stirred in the cold and 1.04g (NH₄)₂SO₄ added in small portions over 30 min. The then slightly turbid solution was stirred for another 30 min and centrifuged (Sorvall SS34; 10000 rpm; 10 min) and the pellet dissolved in 2 ml PBS. After clearing the solution by spinning briefly, the second precipitation was performed as described above by adding 0.54g (NH₄)₂SO₄ over 30 min. The then heavy precipitate was centrifuged as above and resuspended in 2 ml PBS. The solution was dialysed overnight against 2 l PBS at 4°C with gentle stirring, changing buffer twice. The dialysed solution was passed over an ion-exchange column (DE52, 10 ml volume) using (Na)₂HPO₄ buffer (20mM, pH 8.0) and 0.5 ml fractions collected. Aliquots (10 µl) of the fractions were analysed by SDS-polyacrylamide gel electrophoresis, PAGE) and fractions containing the IgG pooled. Overall protein concentration was determined by Bradford assays.

_Bradford assay for the determination of protein concentrations_
Bradford dye was prepared as a 5x solution (400 ml) by dissolving 200 mg Coomassie blue G-250 in 10 ml 95% ethanol and adding 200ml 85% phosphoric acid. The volume was adjusted to 400 ml with water and the solution filtered to remove undissolved dye.

For the standard curve 5, 10, 15 and 20 µl of a BSA solution (1mg/ml) were transferred to a plastic cuvette (total volume 100 µl adjusted with PBS). 1 ml of the dye solution was added and the OD₅₉₅ measured instantly. Samples were measured the same way using various amounts of protein solution (5-10 µl) instead of BSA solution. Concentrations were determined graphically from the standard curve. IgG solutions purified by ammonium sulphate precipitation from crude sera as described above usually had concentrations of around 1mg/ml.
Biotinylation of antibody molecules
Antibodies were biotinylated using an Amersham Biotinylation Kit according to the manufacturers instructions. The protein was diluted with borate buffer (50 mM; pH 8.6; final concentration 1 mg/ml) and 40 μl biotinylation reagent (biotin-spacer arm N-hydroxysuccinimide ester in methyl formamide) per mg of protein added. The mixture was incubated 1 h at room temperature on a rotary wheel and then passed over an equilibrated (PBS, pH 7.5) G25-Sephatex column (10 ml bed volume). Fractions (0.5 ml) were collected and analysed by SDS-PAGE (10 μl/fraction). Fractions containing biotinylated IgG (two bands of 55 and 25 kD) were pooled and the concentration estimated from SDS-PAGE gels by comparing with standard amounts of BSA electrophoresed alongside the fractions. Biotinylated antibodies were usually stored at 4°C.

SDS-polyacrylamide electrophoresis (PAGE)
A Mini-Protean II gel apparatus (BioRad) was used to analyse proteins under denaturing conditions. Separating and stacking gel solutions were prepared as follows:

<table>
<thead>
<tr>
<th></th>
<th>Separating gel (12% w/v)</th>
<th>Stacking gel (4% w/v)</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>3.35 ml</td>
<td>6.1 ml</td>
</tr>
<tr>
<td>1.5 M Tris-HCl, pH 8.8</td>
<td>2.5 ml</td>
<td>2.5 ml</td>
</tr>
<tr>
<td>0.5 M Tris-HCl, pH 6.8</td>
<td>100 μl</td>
<td>100 μl</td>
</tr>
<tr>
<td>acrylamide/bisacrylamide a</td>
<td>4.0 ml</td>
<td>1.3 ml</td>
</tr>
<tr>
<td>10 % w/v ammonium</td>
<td>50 μl</td>
<td>50 μl</td>
</tr>
<tr>
<td>persulphate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total volume:</td>
<td>10 ml (= 2 gels)</td>
<td></td>
</tr>
</tbody>
</table>

a of a 30% w/v acrylamide, 0.8% bisacrylamide solution

Prior to pouring the gels TEMED (10 μl) was added in each case. Before electrophoresis of the samples, 4x sample buffer (0.15 M Tris pH 6.8, 1.4 M β-mercaptoethanol, 4% SDS, 20% glycerol, 5mM EDTA, 0.01% bromophenol blue) was added (final concentration 1x; total volume 10-15 μl) and the mixture boiled for 5 min. After loading the samples electrophoresis was performed in running buffer (25
mM Tris, 192 mM glycine, 0.1 % SDS; pH 8.8) at 0.4 mA and 250-300 V for ~25 min.

IV.B.2) Phage amplification, purification and analysis

*Large-scale preparation of phage particles*

Overnight cultures (LB-medium containing tetracycline, 20 mg/ml; 1 litre) that had been inoculated with single colonies of phage-infected K91Kan *E.coli* cells were cleared by two centrifugation steps at 5000 rpm (Sorvall GS3 rotor) and 8000 rpm (Sorvall GSA rotor) for 10 min. The phage particles were precipitated by adding PEG/NaCl solution (0.15 volumes) to the cleared supernatant, mixing thoroughly by inversion and incubating for 16 h at 4°C. Phage were collected by centrifugation for 30 min at 8000 rpm (Sorvall GSA rotor) and the solution decanted carefully. Pellets were resuspended in TBS (total of 30 ml per l original culture) and cleared by centrifugation at 15 000 rpm (Sorvall SS34 rotor). The dissolved phage were precipitated a second time by again adding NaCl/PEG solution (0.15 volume) and incubation for at least 1 h at 4°C. Phage were collected by centrifugation at 10 000 rpm (Sorvall SS34), pellets again dissolved in TBS (10 ml) and the solution cleared by centrifugation as above. 5.64 g of CsCl were weight in a new tube, the phage solution added and mixed. The total net weight was then adjusted to 18.2 g with TBS which provided about 14 ml of phage/CsCl solution with a density of 1.30 g/ml. The solution was centrifuged in 14 ml tubes at 37 000 rpm (Sorvall, TST 41.14 rotor), 10°C for 60 h. After carefully removing the opaque band in the middle of the gradient (usually around 1 cm wide) and avoiding taking flocculent material from beneath the phage band, the phage solution was transferred to tubes for the TST60.4 rotor (Sorvall; 4.4 ml), topped up with TBS, mixed, and the phage pelleted by centrifugation at 45 000 rpm for 4 h at 10°C. This centrifugation step removes most of the remaining CsCl and can be repeated if very clean phage are required as, for example, for immunisation of animals. After the last centrifugation step the pellet was dissolved in TBS (3 ml per l original culture) containing 0.02% NaN₃. The virion concentration of the solution was measured at 320 and 269 nm at a 1/100 dilution in water and the concentration calculated using the formula $A_{269} - A_{320} \times 6 \times 10^{16}/9.2$ kb. A typical concentration obtained by the purification method described here was about $10^{14}$ virions/ml.
IV. Materials and Methods

Titration of phage as plaque forming units (p.f.u.)
For titering of plaque forming units, phage were diluted serially (usually to $10^{-8}$ for an amplified and purified phage solution) in TBS containing gelatine (1 g/l) and aliquots of this dilution (100 µl) added to a suspension of K91 Kan cells (400 µl) grown to OD$_{600}$ 0.2 (1/10 dilution) in TB-medium from a 1/100 inoculation with an overnight culture. Melted soft agar (3 ml) was then poured into the tubes containing the phage dilution/host cell suspension, gently mixed and poured on L-agar plates containing kanamycin (100 mg/ml). The plates were incubated overnight at 37°C and plaques counted the next day.

Titration of phage as transforming units (t.u.)
For titering of transforming units, phage were diluted serially as described above, aliquots (10 µl) gently mixed with K91 Kan cells (10 µl) from a TB-culture grown as described above and then left for 10 min at room temperature to allow the phage to infect the E.coli cells. LB-medium (1 ml) containing tetracycline (0.2 µg/ml) was added to the cell suspension and incubated 30-40 min at 37°C on a rotary wheel to allow expression of antibiotic resistance genes. Aliquots (200 µl) were plated on L-agar plates containing tetracycline (40 µg/ml) and kanamycin (100 µg/ml) and incubated overnight at 37°C. Colonies were counted the next day.

Preparation of single-stranded phage DNA for sequence analysis
Single-stranded phage DNA was prepared from 1.5 ml overnight cultures inoculated with single colonies of K91 Kan cells infected with phage derived from titering t.u. Cells were removed by centrifugation at 11000 rpm for 5 min (Heraeus Biofuge 15R, also all following centrifugations) and the clear supernatants poured into fresh tubes containing PEG/NaCl solution (0.15 volumes). Tubes were inverted 100 times and put on ice for 4h for precipitation of phage particles. Phage were collected by centrifugation at 11000 rpm, 10°C for 15 min and the PEG/NaCl solution carefully removed by aspiration. Pellets were redissolved in TBS (250 µl) and extracted once with phenol (200 µl) and once with chloroform/iso-amylalcohol (24:1; 200 µl). DNA was precipitated by adding 1/10 volume sodium acetate (3 M, pH 5.2) and 2.5 volumes ethanol to the aqueous phase. After vortexing the tubes were stored overnight at -20°C. DNA was collected by centrifugation at 11000 rpm, 4°C for 30 min and the pellets washed with cold 70% ethanol. Pellets were dried in a desiccator and dissolved in water (20 µl). Aliquots (8 µl) were used directly for sequencing reactions.
Sequence analysis of phage DNA

Sequencing reactions were performed essentially after the dideoxy-mediated chain termination method (Sanger et al., 1977).

Annealing mix:

8 μl  single-stranded template DNA in water
1 μl  1 M Tris/50 mM MgCl₂
1 μl  primer (OD 0.23, 5'AGT TTT GTC TTT CC 3', complementary to fd-phage genome downstream of the multiple cloning site)

The annealing mixture was heated to 80°C for 5 min, briefly centrifuged and left to cool at room temperature. Aliquots (2 μl) were then distributed to four wells of a microtiter plate for each sample analysed. To each well containing nucleotide mixes (order G, A, T, C; 2 μl) reaction mix (2 μl) was added:

Nucleotide mixes (volumes in μl):

<table>
<thead>
<tr>
<th></th>
<th>G</th>
<th>A</th>
<th>T</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5mM dGTP</td>
<td>12.5</td>
<td>250</td>
<td>250</td>
<td>250</td>
</tr>
<tr>
<td>0.5mM dTTP</td>
<td>250</td>
<td>250</td>
<td>12.5</td>
<td>250</td>
</tr>
<tr>
<td>0.5mM dCTP</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>12.5</td>
</tr>
<tr>
<td>10mM ddGTP</td>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10mM ddATP</td>
<td>-</td>
<td>0.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10mM ddTTP</td>
<td>-</td>
<td>-</td>
<td>25</td>
<td>-</td>
</tr>
<tr>
<td>10mM ddCTP</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>TE</td>
<td>500</td>
<td>250</td>
<td>500</td>
<td>500</td>
</tr>
</tbody>
</table>

Reaction mix (volumes are per template analysed):

8 μl  10mM Tris pH 8.0
0.1 μl 1 M DTT
0.5 μl [α-³⁵S]dATP (10 μCi/μl)
1.5 U  E.coli DNA polymerase I Klenow fragment (5 U/μl)

The reagents were gently mixed by tapping the plate and incubated at room temperature for 25 min. Chase mix was added (0.25 mM dGTP, dATP, dTTP and dCTP in TE; 2 μl) and incubation continued for a further 20 min. The reaction was stopped by adding gel loading buffer [(0.1% (w/v) xylene cyanol, 0.1% (w/v)
bromophenol blue, 10 mM EDTA in formamide; 2 μl). For electrophoresis aliquots (2 μl) of each reaction were heated to 80°C for 5 min and immediately loaded onto a 6% denaturing polyacrylamid gel and the DNA fragments separated at 60 W in 1x TBE on a S2 gel apparatus (BRL) for about 3-4 h. Gels were dried for 2 h on a gel dryer and autoradiographed overnight at -70°C.

**Urea-polyacrylamide gel solution (for 50 ml):**
7.5 ml acrylamide (40% w/v acrylamide/2.105 % w/v bisacrylamide)
5.0 ml 10x TBE (90 mM Tris, 90 mM boric acid, 2 mM di-sodium EDTA)
23 g urea
250 μl ammonium persulfate (10% w/v)
10 μl TEMED

**Immuno dot blots with anti-fd phage rabbit serum**
Various dilutions of fusion phage particles (10^5 to 10^10 particles in 50 μl PBS) were blotted in four sets on nitrocellulose membrane (Costar) by using a Schleicher & Schuell dot blot apparatus. The membrane was blocked in blocking buffer (PBS containing 5 % w/v milk powder) for 90 min, the membrane cut into strips each carrying one full set of phage dilutions and each strip incubated with different dilutions of anti-fd phage rabbit serum (1/100, 1/200, 1/500 and 1/1000 in PBS containing 1% w/v milk powder) at room temperature for 2h. The membrane strips were washed twice for 10 min with PBS and incubated with alkaline phosphatase conjugated anti-rabbit IgG (diluted 1/5000 in PBS containing 1% w/v milk powder) for 2h at room temperature. After washing again as above the blot was developed using BCIP (50 mg/ml in water) and NBT (50 mg/ml in 70% v/v DMF).

The detection threshold of phage particles with anti-phage serum was 10^8 particles. For use in indirect phage-sandwich ELISAs IgG molecules were purified from the anti-fd phage rabbit serum as described and used at 0.5 μg/ml.

**IV.B.3) Biopanning procedures**

*Amplification of the random hexapeptide phage-display library*
The phage library was received from Prof. G. Smith as an aliquot of an amplified library (Scott and Smith, 1990). To obtain sufficient material for biopanning
experiments the library was amplified further as follows: Two TB-cultures (50 ml) were inoculated with K91Kan cells of an overnight culture (0.5 ml) and grown to a density of OD sub 600 0.2 (1/10 dilution). Shaking was slowed down for 5 min, the culture infected with the library (10^{12} physical particles; 4 μl) and slow shaking continued for 15 min. Each pre-culture was poured into a fresh flask containing LB-medium (500 ml) supplemented with tetracycline (0.2 μg/ml) and incubated on the shaker for 35 min. The tetracycline concentration was adjusted by adding 0.5 ml (20 mg/ml stock solution) to each flask and incubation continued overnight. The following day aliquots of the culture were titred as t.u. (10^{-5} dilutions) and the total number of infected cells for each culture determined. Phage particles contained in the overnight culture representing the amplified library were prepared as described for the large-scale preparation of phage particles.

Biopanning with polyclonal serum IgG and petri dishes as solid support
3.5 cm petri dishes (Falcon) were coated overnight at 4°C with streptavidin (10 μg in 1 ml 0.1 M NaHCO₃, pH 8.6) and then blocked with blocking solution (0.1 M NaHCO₃; 5 mg/ml BSA; 0.1 μg/ml streptavidin) for 1 h at 4°C. Dishes were washed 6 times with washing buffer (TBS/0.05% Tween 20) and then filled with washing buffer (400 μl) containing BSA (1 mg/ml). For rounds 1 and 2 the streptavidin-coated dishes were saturated with substrate by adding biotinylated, purified IgG molecules (30 μl containing around 10 μg) to the solution and incubating 2 h at 4°C ("Phage+LigateStreptavidin method"). After coating the plate remaining streptavidin binding sites were blocked by adding biotin (10 mM; 4 μl) and further incubation for 1 h. After washing the dishes as above, plates were filled with washing buffer (400 μl) containing biotin (10 mM; 4 μl) and the amplified phage library (5 μl, 4 x 10^{10} TU; round 1 only) or amplified eluate (100 μl round 2 only; 1.5 x 10^{10} TU), respectively. Phage were incubated 4 h at 4°C with gentle agitation on a Bellydancer®. After discarding the solution, dishes were washed 8 times with washing buffer at room temperature and bound phage eluted by adding elution buffer (400 μl; 0.1 N HCl adjusted to pH 2.2 with glycine) and gently rocking for 10 min at room temperature.

For round 3 ("PL+S method") 100 μl of amplified eluate from round 2 was incubated with 6-12 μl biotinylated substrate (around 2-4 μg) in an Eppendorf tube overnight at 4°C on a rotary wheel and then together with washing buffer (400 μl total volume) transferred to freshly streptavidin-coated and blocked dishes. After incubation for 10
min at room temperature dishes were washed and phage eluted as described above for rounds 1 and 2.

Phage eluates were neutralised with Tris-HCl (1 M, pH 9.5; 75 μl). For amplification of phage contained in the first eluate (475 μl) the whole eluate was used and therefore was reduced to 100 μl by centrifugation in a Microsept spin filter with 160 kD cut-off size (Sorvall SS34 rotor; 5000 rpm). For subsequent rounds only 100 μl (21%) of the eluate was taken as input for the amplification.

**Biopanning with polyclonal serum IgG using paramagnetic Dynabeads® as solid support**

Streptavidin-coated magnetic beads (Dynal, M-280) were used as solid support to affinity select phage specific for polyclonal antibodies. The quantity of beads used here, had a surface area approximately that of 3.5 cm petri dishe, with a maximal binding capacity of 3 μg biotinylated IgG molecules (according to the manufacturer's instructions).

After washing the beads twice with PBS containing BSA (0.1% w/v), using a magnetic concentrator (Dynal, MPC-E), the equivalent of 30 μl original bead suspension was coated with biotinylated substrate (~5 μg purified IgG) according to the manufacturer's instructions for 30 min at room temperature. The substrate-coated beads were blocked by adding biotin (10 mM; 0.5 μl) to the coating solution and continuing incubation at room temperature for 1 h. Beads were washed 4-5 times as above. Input phage for the biopanning were either the phage library (5 μl round 1; 3.6 x 10¹⁰ TU) or amplified eluates (100 μl round 2 and following). The total volume of the beads and phage suspension was adjusted with PBS containing BSA (0.1% w/v; final volume 250 μl) and incubated for 4 h at 4°C. Beads were then washed 6 times with PBS containing Tween 20 (0.05% v/v) and the bound phage eluted with elution buffer (50 μl) for 10-15 min at room temperature. Eluates were neutralised with Tris-HCl (1 M, adjusted to pH 2.2 with glycine; 10 μl) and phage amplified and titered. Due to the smaller volumes of the eluates using this method there was no need to concentrate the first eluate prior to amplification.

**Biopanning with monoclonal antibodies using microwells as solid support**

Wells of a 96-well ELISA plate were coated with purified monoclonal antibody (1 μg/ml in 0.2 M ammonium bicarbonate buffer, pH 9.6; 200 μl) overnight at 4°C and
then blocked for 2 h with TBS containing Tween 20 (0.05% v/v) and BSA (1%). After washing with TBS containing Tween 20 (0.05% v/v) and BSA (0.1%), the phage (5 µl containing $10^{10}$ t.u. phage library for round 1; 50 µl amplified eluate for following rounds) in TBS containing BSA (0.1% w/v; 50 µl total volume) were incubated in the wells for 4 h at 4°C. Wells were washed 10 times with the same buffer as before and the bound phage eluted with elution buffer (50 µl) for 10 min at room temperature. The eluates were neutralised with 1 M Tris-HCl, pH 9.5 (10 µl) and phage amplified and titered.

**Amplification and titration of phage contained in eluates of biopanning experiments (output phage)**

Whole or concentrated eluates (100 µl) were gently mixed with K91 Kan cells (100 µl) from a TB-culture grown to OD$_{600}$ 0.2 (1/10 dilution) and incubated for 10 min at room temperature. The suspension of infected *E.coli* cells was then added to LB-medium (20 ml) containing tetracycline (0.2 µg/ml) in an Erlenmeyer flask and incubated on a shaker at 37°C for 30 min. Afterwards, an aliquot (10 µl) was removed from the culture for titration of output phage and tetracycline added (20 µl of a 20 mg/ml stock solution). Incubation was continued overnight. Phage titres of the aliquot taken from the culture were determined as t.u. as described above and the total number of phage contained in the 20 ml culture before amplification calculated.

Amplified phage were prepared from the overnight culture by two precipitations with NaCl/PEG solution as described above for the large-scale purification and dissolved in TBS (200 µl) containing NaN$_3$ (0.02% v/v).

**IV.B.4) Analysis of individual phage**

**Phage binding assay on solid phase (input-output relationship)**

U-shaped wells of polystyrene strip plates were coated with MAb (1 µg/ml in 0.2 M ammonium bicarbonate buffer, pH 9.6; 200 µl) overnight at room temperature, washed three times with TBS and blocked for 2 h at room temperature by adding TBS containing BSA (10 mg/ml; 200 µl). After three further washes the wells were filled with various concentrations of purified phage particles ($10^7$ to $10^{10}$ p.f.u.; 200 µl) diluted in TBS containing BSA (0.2 mg/ml) and NaN$_3$ (0.02% v/v) and incubated for 1 h at 4°C. Each phage concentration was assayed in triplicate. After one hour the
phage solutions were removed from the wells and plates were then washed 10 times with the buffer used above and bound phage eluted with elution buffer (200 μl). The amount of phage recovered from each well was determined by titration as p.f.u. and plotted against the amount of input phage per well.

**Phage binding assay over time (time course experiments)**

For these experiments five identical sets of plates were prepared with coating conditions as above and the input phage concentrations kept constant (10⁸ p.f.u./well; 200 μl). Each plate represented a point on a time course with incubation times of 10, 31, 100, 316 and 1000 min (corresponding to half log intervals). After each incubation period the relevant plate was washed as described above, the bound phage eluted and the titer of the eluate determined as p.f.u. The amount of phage recovered from each well was plotted against the time of incubation.

**Determination of relative dissociation constants (K_D^rel) for the interaction of phage with monoclonal antibodies in solution**

This procedure was adapted from Friguet et al. (1985). **Step I** involved incubation of a constant concentration of purified phage particles with variable amounts of MAb 18/7 until equilibrium was reached: MAb 18/7 at various concentrations (0.32 to 30 nM in TBS and one blank sample containing buffer only; measurements for each MAb concentration were made in triplicates) was mixed with a constant amount of phage [10⁸ p.f.u./ml, in TBS containing BSA (0.2 mg/ml) and NaN₃ (0.02% v/v)] and incubated for 18 hr at 4°C in a water bath. **Step II** was to determine the number of unbound phage remaining in the solution after 18 h. For this, an aliquot (100μl) of each mixture was transferred to the wells of a U-shaped well polystyrene strip plate that had been coated with MAb as described above. After 1 hr at 4°C the wells were washed 10 times with TBS, supplemented with BSA (0.2 mg/ml). Bound phage were eluted with elution buffer (200μl) and quantitated by titration of p.f.u.

For the determination of the K_D^rel values two methods were used:

1.) The mean of the values of one set of triplicates from one or several experiments representing the number of phage remaining in solution at equilibrium for a certain MAb concentration was determined and the percentage of free phage remaining in solution calculated with the number obtained for the well in the absence of MAb (set at 100%). Percentages were then plotted against the concentration of MAb (added in **step I**) and a curve fitted to these points using the least square curve fitting option.
contained in the Sigma Plot Graph plotting program (see Fig. II.5). The concentration of MAb at which 50% of the total amount of phage remain unbound at equilibrium equals the $K_D^\text{rel}$ value (the theoretical background is described in Results section II.A.3)).

2.) The number of free phage remaining in solution at various MAb concentrations allows calculation of the concentration of all other reactants in the equilibrium given that the input titers and the input concentrations were known. With these figures a plot can be generated according to the Scatchard equation with the slope of the plot yielding the $K_D^\text{rel}$ value (the theoretical background is described in Results section II.A.3)).

**Indirect phage-sandwich enzyme-linked immunosorbent assay (ELISA) using anti-fd phage antibodies**

Wells of ELISA microplates were coated overnight at room temperature with purified polyclonal serum IgG [100 ng in ammonium bicarbonate buffer (0.2 M, pH 9.6; 100 μl)]. The coating solution was removed, the wells washed once with washing buffer (TBS, 0.05% v/v Tween 20) and blocked for 90 min at room temperature with blocking buffer (TBS, 5% w/v dry milk powder; 200 μl). The wells were washed three times with washing buffer and the phage solution added to the wells (10$^{11}$ p.f.u. in TBS, 1% w/v milk powder; 100 μl). The phage were either purified by centrifugation in CsCl gradient or where large numbers of phage were screened, crudely purified by one PEG precipitation and dissolved in 1/10 of the original culture volume. The phage solutions were incubated in the wells at 4°C overnight. Next day the wells were washed three times as above at 4°C and then incubated with rabbit anti-fd phage antibodies (purified IgG, 0.5 μg/ml in blocking buffer; 100 μl) for 2 h at 4°C. Wells were washed again as above and incubated with alkaline phosphatase-conjugated goat anti-rabbit IgG antibodies (diluted 1/5000 in blocking buffer; 100 μl). The wells were washed again as before, slammed on a pile of paper towels to remove traces of liquid and substrate added (pNPP 1 mg/ml in diethanolamine, 10 mM, pH 9.6; 100 μl). The assays were developed at room temperature for 1 h. The reaction was stopped by adding EDTA (0.1 M; 100 μl) and the optical densities determined by reading on a Dynatech® plate reader at 405 nm. Blank wells for calibration contained substrate and stop solution only. All ELISA assays were carried out in duplicate.
Competition experiments by indirect phage-sandwich ELISA using synthetic peptides
The binding of phage particles to wells coated with purified polyclonal serum IgG was competed by addition of synthetic peptides with various amino acid sequences. The assay was performed as described above except that the phage solution additionally contained synthetic peptides at variable (10^{-7} to 10^{-3} M) or constant concentrations (10^{-3} M). Peptides were freeze dried after analysis and then dissolved in water to make up defined stock solutions. Since P2 was extremely hygroscopic the precise peptide concentration was determined after dissolving the peptide in water by amino acid analysis of a hydrolysed peptide aliquot, kindly performed by M. Daniels and Prof. R.P. Ambler, ICMB, University of Edinburgh.

IV.B.5) Immunisation of rabbits with phage particles
Fuse5 based fusion phage purified by CsCl gradient centrifugation (Smith and Scott, 1993) and dissolved were used directly as immunogens. Two rabbits per immunogen were injected three times in four week intervals in the legs and ears with diluted phage solution (100 µl, 1.7 x 10^{12} particles in PBS). Serum from the third test bleeding (two weeks after third injection) was tested for reactivity with fd phage particles by performing immuno dot blots with various dilutions of phage particles (IV.B.2, p.112). IgG molecules were purified as described in IV.B.1. The immunisations were carried out by Glynis Leadbetter of the Department of Medical Microbiology, University of Edinburgh.
CHAPTER V: REFERENCES
V. References


three-dimensional structure of the Fab D1.3-lysozyme complex at 2.5-Å resolution. *J. Biol. Chem.* 266, 12915-12920.


V. References


unliganded Fab and a deoxynucleotide-Fab complex. *Proteins* 11, 159-175.


V. References

1590-1593.


V. References


V. References

to peptides detect new hepatitis B antigen: serological correlation with

Motti, C., Nuzzo, M., Meola, A., Galfré, G., Felici, F., Cortese, R., Nicosia, A. and
Monaci, P. (1994) Recognition by human sera and immunogenicity of
HBsAg mimotopes selected from an M13 phage display library. Gene 146,
191-198.

immunization with a synthetic conjugate. Proc. Natl. Acad. Sci. USA 79, 569-
573.

Murray, K., Bruce, S.A., Hinnen, A., Wingfield, P., Van Erd, P.M.C.A., De Reus, A.
immunise against infection. EMBO J. 3, 645-650.

and Cys-61 determines the structure of the secretory core gene product (e
antigen) of hepatitis B virus. J. Virol. 67, 4307-4315.

synthetic peptide having a sequence in common with a fragment of a virus
protein, the hepatitis B surface antigen. Proc. Natl. Acad. Sci. USA 79, 7871-
7875.

pre-S gene coded immunodominant epitope of hepatitis B virus. Science 224,
392-395.


chemical synthesis of a host cell receptor binding site on hepatitis B virus.
Cell 46, 429-436.
V. References


V. References


V. References


V. References


CHAPTER VI: APPENDIX
Identification of polyclonal serum specificities with phage-display libraries

Volker Germaschewski and Kenneth Murray

Institute of Cell and Molecular Biology, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, UK


Abstract:

A random hexapeptide fusion-phage library was screened to isolate phage that bound antibodies in a serum induced by hepatitis B virus surface antigen (HBsAg). Analysis of the isolated phage and comparison with their displayed peptide sequences with the primary sequence of HBsAg revealed areas where 3 and 4 amino acid matches accumulated. Differential binding studies of individual phage clones with immune and pre-immune sera identified phage carrying sequences that matched with region 117-122 of HBsAg which may represent a linear epitope or part of a larger antigenic determinant. Synthetic hexapeptides representing this region competed for binding with the matching phage clones.
Screening a Monoclonal Antibody With a Fusion-Phage Display Library Shows a Discontinuity in a Linear Epitope Within PreS1 of Hepatitis B Virus

Volker Germaschewski and Kenneth Murray

Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh, Scotland, UK

The epitope recognized by the monoclonal antibody MA18/7, specific for the PreS1-domain of the hepatitis B virus surface antigen, has been defined precisely by means of a library of fusion-phage carrying random hexapeptides on the tip of filamentous phage fd particles. Phage, isolated after only one round of affinity selection, displayed hexapeptides showing strong conservation of the PreS1 primary sequence in the region 19-23 with three noncontiguous residues, DP (20 and 21) and F (23) appearing in phage that bound the antibody. The importance of these core residues was supported by comparing the antibody binding of individual phage in solution, which provided relative dissociation constants for these interactions. Replacement of F (23) by Y was the only substitution observed in the three core residues, and resulted in somewhat weaker binding. Synthetic tetra- and hexapeptides containing these key residues inhibited the reaction between the phage and the antibody.

INTRODUCTION

The region of an antigen which binds to a specific antibody, usually termed an epitope, may be defined by a short sequence of amino acid residues (linear epitopes), or the site may be more complex and involve the juxtaposition of residues that are well separated from each other within the linear sequence (conformational epitopes). Definition of the amino acid residues that constitute an epitope is of fundamental importance to immunological studies and to the design and development of vaccines and the evaluation of their effectiveness. Where the amino acid sequence of an antigenic protein is known, synthetic oligopeptides representing parts of this sequence have frequently proved to be valuable reagents for the location of regions that constitute or contain epitopes for a particular antibody or antiserum. Such studies have helped to locate linear and conformational epitopes in the antigens of hepatitis B virus (HBV) [Gerin et al., 1983].

The surface antigen (HBsAg) of HBV may contain three polypeptides termed S (small), M (middle), and L (large), which result from the variable use of three start codons within the open reading frame of the surface antigen gene. The L polypeptide begins with the PreS1 region (108 amino acids), followed by the PreS2 domain (55 residues) and the amino acid sequence of S (226 residues). The M polypeptide comprises the PreS2 sequence as its N-terminal region, followed by the S sequence. S is the most abundant product and occurs in glycosylated and nonglycosylated forms. M and L also occur in differentially glycosylated forms. In addition to its presence in the viral envelope, HBsAg is found in large quantities in the serum of infected individuals as both spherical and filamentous particles, and proportions of the L, M, and S polypeptides in these three forms differ appreciably. Both polyclonal and monoclonal antibodies have been identified against immunological determinants in the domains PreS1, PreS2, and S of HBsAg, and studies with synthetic peptides have identified several distinct determinants including components of conformational epitopes in S (summarized by Ashton-Rickardt and Murray [1989]; see also Gerin et al. [1983]). Heermann et al. [1984] found that the monoclonal antibody (MAb) MA18/7 specifically recognizes the PreS1 domain, and therefore interacts only with the L polypeptide. They also showed that the only observed amino acid change between different HBV serotypes, glycine to arginine at position 24 of PreS1, was tolerated without reducing antibody binding. Coursaget et al. [1991] mapped the epitope recognized by this MAb to residues 18–25 (QLDPAFRA). Sominskaya et al. [1992] then used progressive deletions around this sequence to show that the central tetrapeptide DPAF constituted the epitope recognized by MA18/7.
In the work described here, fd-fusion phage that bind specifically to MAb MA18/7 were isolated from a random hexapeptide phage-display library [Scott and Smith, 1990]. This essentially unbiased approach has proved to be very powerful for determining native epitopes recognized by MAbs [Scott and Smith, 1990; Cwirla et al., 1990; Felici et al., 1991; Stephen and Lane, 1992], but may also identify new, apparently unrelated sequences, or mimotopes, which can open new vistas in terms of targets (such as components of conformational epitopes) for existing antibodies [Scott and Smith, 1990; Christian et al., 1992]. The results show that the phage-display method can provide a very simple route to high-resolution mapping, which yields information about the contribution of individual amino acids to the binding of the antibody which could not be obtained as readily with most methods based upon synthetic peptides or deletion analysis. In this specific case the phage isolated from the library carry sequences that lie within the region DPAP (PreSi aa 20–23), identified by Sominskaya et al. [1992], but not all of these amino acids are necessary for binding MA18/7. Furthermore, we have shown that the system can be used quite simply in a quantitative mode to measure relative affinities of a MAb for different peptide sequences. The identity of the epitope was confirmed by quantitative studies, showing that binding of the cloned fusion phage to the MAb was inhibited by synthetic peptides carrying the core epitope sequence.

MATERIALS AND METHODS

Monoclonal Antibody and Fusion Phage-Display Library

The MAb MA18/7 was a gift from Dr. K.H. Heermann and Prof. W. Gerlich, University of Göttingen, Germany.

The phage-display library was amplified from a primary library, consisting of approximately $2 \times 10^8$ independent transfectants, and was generously given to us by Prof. G.P. Smith, University of Missouri, Columbia, MO.

Biopanning Procedure

Wells of an ELISA plate (Corning) were coated with MAb MA18/7 (1 μg/ml in 0.2 M ammonium bicarbonate buffer, pH 9.6; 200 μl) overnight at 4°C, and then blocked for 2 hr with TBS (0.05 M tris-HCl, pH 7.5, 0.15 M NaCl) containing Tween-20 (0.05%) and bovine serum albumin (BSA; 1%). After washing with TBS containing Tween-20 (0.05%) and BSA (0.1%), the phage library (5 μl; $10^{10}$ transforming units (t.u.) or 50 μl of amplified eluate for rounds 2–4) in TBS containing BSA (0.1%; 50 μl total volume) was incubated in the wells for 4 hr at 4°C. The wells were washed 10 times with the same buffer as before, and the bound phage were eluted with elution buffer (0.1 N HCl, adjusted to pH 2.2 with glycine; 50 μl) for 10 min at room temperature. The eluates were neutralized with 1 M tris-HCl, pH 9.5 (10 μl), and amplified by gently mixing the eluate (60 μl) with a suspension of K91Kan cells (60 μl) at OD$_{600}$ (optical density) 0.2 (1/10 dilution), and leaving it for 10 min at room temperature. The cells were suspended in 20 ml LB (Luria-Bertoni) medium containing tetracycline (0.2 mg/ml) and incubated at 37°C for 45 min. Tetracycline (20 mg/ml; 20 μl) was then added, and an aliquot (10 μl) removed for titration as t.u. [Smith and Scott, 1993b]. Incubation was continued overnight. Phage particles were purified the following day as described by Smith and Scott [1993b].

Titration of Phage as Plaque-Forming Units (p.f.u.)

For the titration of plaque-forming units (p.f.u.), phage were diluted serially (usually to $10^{-8}$ for an amplified phage solution) in TBS containing gelatin (1 g/l) and aliquots of this dilution (100 μl) added to a suspension of K91Kan cells (400 μl) at OD$_{600}$ 0.2 (1/10 dilution). Melted soft agar (3 ml) was then poured into the tubes containing the phage dilution/host cell suspension, and poured onto agar plates containing kanamycin (100 mg/ml). The plates were incubated overnight at 37°C, and plaques were counted the next day.

Sequence Analysis of Phage DNA

fd-DNA (single-stranded) was prepared from 1.5 ml overnight cultures of K91Kan cells infected with individual colonies essentially as described by Sambrook et al. [1989].

Half of the DNA was used for sequencing with the dideoxy-mediated chain termination method [Sanger et al., 1977], using Klenow DNA-polymerase (NBL Gene Sciences Ltd., Cramlington, UK) and the primer 5′AGT TTT GTC GTC TTT CC 3′; synthesized by Ossel DNA Service, Dept. of Chemistry, University of Edinburgh).

Binding Assay in Solution Between Individual Phage and MAb MA18/7

This procedure was adapted from Friguet et al. [1985]. Each assay was carried out in triplicate. Each phage used in the assay was purified from the supernatant of a 500 ml overnight culture of infected K91Kan cells [Smith and Scott, 1993b]. Purified phage (2 \times 10^9 p.f.u./ml; 200 μl) were then incubated with various concentrations of MA18/7 (final concentrations ranging from $10^{-7.5}$ to $10^{-9.5}$ M) in TBS containing BSA (0.2 mg/ml) overnight at 6°C (total volume, 400 μl). The concentration of free phage remaining in solution at equilibrium was determined for each sample containing different concentrations of MA18/7, by transferring the phage/MAb mixture (100 μl containing approximately $10^8$ p.f.u.) to microtiter wells (Costar strip plates, high capacity) coated with MA18/7 (as described above). After incubation for 1 hr at 6°C, the wells were washed extensively and the phage particles eluted as described above, and titered as p.f.u.
Fig. 1. Hexapeptide sequences displayed on phage isolated from selection rounds 1, 2, and 4. The experiment with MA18/7 and the random phage library is described in Materials and Methods. The bold box represents the residues of PreSi defined as the binding site for MA18/7 by Coursaget et al. [1991]. The amino acids within the thin boxes are the conserved residues with respect to residues in the PreS1 sequence; the hexapeptides labelled as “no matches” are from round 1, and display no obvious consensus with the PreS1 region examined here.

**Competition Experiments With Synthetic Peptides**

Phage particle concentration was adjusted to a titer of $4 \times 10^9$ p.f.u./ml in TBS containing BSA (0.2 mg/ml). Aliquots of the phage solution (50 µl) were mixed with serial dilutions (50 µl) of the synthetic peptides (final concentration ranging from $10^{-6}$–$10^{-11}$ M for P1, and from $10^{-4}$–$10^{-9}$ M for P2 and P3, diluted in TBS containing BSA (0.2 mg/ml)) and transferred to microtiter wells coated with MA18/7 as described above. After incubation for 90 min at 6°C, the wells were emptied and washed and the phage eluted as described above. Phage titers were determined as p.f.u., as described above.

The peptides were synthesized in the Department of Chemistry of the University of Edinburgh (for which we thank Prof. R. Ramage), and had the following amino acid sequences: P1) LDPAFR, P2) DPGF, and P3) ALLTRILG. Due to its very hygroscopic nature, the concentration of P2 was determined after dissolving it in TBS by amino acid analysis of the hydrolyzed peptide, as kindly carried out by M. Daniel and Prof. R.P. Ambler, ICMB, University of Edinburgh.

**RESULTS**

**Biopanning the Fusion-Phage Library With the MAb MA18/7**

About a dozen phage-transformed E. coli colonies isolated from the eluates from each of four successive rounds of affinity selection and amplification (biopanning) were analyzed by sequencing the insert in gene III of the phage genome. Figure 1 shows that after only one round of biopanning, almost 50% of the phage analyzed carried peptides corresponding with three or four
residues of the sequence LDPAF of PreS1 (aa 19–23) with the dipeptide DP present in all of the matching sequences. In fact, DP was conserved throughout all the phage analyzed from rounds 2 and 4. The phenylalanine (F) at residue 23 was also strongly conserved with a conservative substitution by tyrosine (Y) in two cases. Alanine (A) at residue 22 was only present in 45% of the phage examined, but was otherwise replaced by amino acids with small side chains such as glycine (G), valine (V), and serine (S).

Figure 1 depicts the alignment of the hexapeptide sequences displayed by individual phage with the amino acid sequence of PreS1 in the region identified as the epitope for MA18/7 by Coursaget et al. [1991] (bold box). The matching residues in the various phage lie within this sequence, but there was no conservation of the alanine at residue 25, and residues 18 and 24 (Q and R, respectively) were present in only two or one of the phage, respectively.

Affinity Assays in Solution

The relative dissociation constants (K\textsubscript{Drel}) for four of the isolated phage (Table I) were measured to compare the binding affinity of phage with different residues in the hexapeptide from those in the native PreS1 sequence.

The K\textsubscript{Drel} values were determined from curve-fitting plots of the proportion of unbound phage at various antibody concentrations (Fig. 2) which gave the values shown in Table I, column 5. All were in the nM range for the temperature chosen for the experiment (6°C). The differences between the four phage analyzed are small. E4-4 had the largest number of amino acids (five), corresponding to the PreS1 primary sequence, and had one of the lowest K\textsubscript{Drel} values. Phage E2-14, E4-2, and E4-4 all had similar K\textsubscript{Drel} values, but phage E2-12, which represents the two phage where residue 23 (F) is not conserved, had the highest K\textsubscript{Drel}, i.e., the lowest affinity. These values emphasize the relative unimportance of residues 18 (Q) and 19 (L) for the binding of the antibody.

Competition Experiments With Synthetic Peptides and Phage Specific for MA18/7

The hexapeptide P1 covered six conserved amino acids at the center of the epitope (LDPAFR). Parts of this sequence occurred among all the phage analyzed (Fig. 1). P2 was a tetrapeptide that had only the three conserved amino acids D, P, and F from the PreS1 sequence, with the A at position 22 replaced by G. P3 was an octapeptide of unrelated sequence.

These peptides were used as competitors for binding to MAb MA18/7, with cloned phage isolated after the biopanning procedure. Phage E4-4 was incubated with various amounts of the synthetic peptides in microtiter wells coated with the MAb (see Materials and Methods). Figure 3 shows the effects of increasing concentrations of competing peptide on the binding of this phage to the antibody.

**TABLE I. Characteristics of the Four Phage Chosen for the Affinity Assay in Solution**

<table>
<thead>
<tr>
<th>Phage no.</th>
<th>Sequence of hexapeptide insert*</th>
<th>Conserved residues out of 6</th>
<th>K\textsubscript{Drel} at 6°C (nM)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2-12</td>
<td>ALDPAY</td>
<td>4</td>
<td>F (23) subst. by Y</td>
<td>0.8</td>
</tr>
<tr>
<td>E2-14</td>
<td>SLDPGF</td>
<td>4</td>
<td>Core motif + L (19)</td>
<td>0.3</td>
</tr>
<tr>
<td>E4-2</td>
<td>PDPGFN</td>
<td>3</td>
<td>Core motif only</td>
<td>0.2</td>
</tr>
<tr>
<td>E4-4</td>
<td>QLDPGFF</td>
<td>5</td>
<td>Frequent substitution of G for A (22)</td>
<td>0.2</td>
</tr>
</tbody>
</table>

\*The bold type identifies residues that occur in the PreS1 sequence.
At concentrations above $10^{-8}$ M, peptide P1 strongly inhibited the binding of the phage to the antibody and completely abolished it at 1 μM (Fig. 3). Similar results were obtained with other phage displaying fewer amino acids of the PreS1 sequence (data not shown). The tetrapeptide P2, representing the core motif DPGF of the PreS1 epitope, was less effective in competing for binding, and concentrations above 1 μM were needed to reduce phage-binding substantially, but at 0.1 mM, phage-binding was reduced to less than 10%. The octapeptide P3, used as a negative control, showed no such inhibitory effects even at the highest concentration.

**DISCUSSION**

The purpose of this study was to define precisely the amino acids within the PreS1 sequence of HBV which are essential for binding the monoclonal antibody MA18/7. A fusion-phage library consisting of hexapeptides displayed on the tip of filamentous phage particles was used to provide an unbiased and unconstrained survey of the residues involved in binding the MAb. The results were compared with those from the study by Coursaget et al. [1991], with overlapping hexapeptides spanning amino acids 12–53 (i.e., 1–42 of PreS1 as used here), which located the epitope within residues 18–25 (QLDPAFRA) of PreS1, and with those from Sominskaya et al. [1992], who used progressive deletions in the coding sequence for PreS1 to express a series of PreS1 peptides with deletions flanking the epitope, which was thus located within residues 20–23 (DPAPF).

In the experiments described here, phage were isolated after only one round of affinity selection, which showed three highly conserved residues corresponding with the PreS1 sequence which are critical for efficient binding to the antibody. This minimal core motif of residues 20 (D), 21 (P), and 23 (F or Y) was conserved in all phage examined, and these residues are therefore believed to make direct contact with the antibody. Neighboring amino acids showed little or no conservation, and therefore probably have little influence on antibody binding. While 20 (D) and 21 (P) were invariant, a conservative substitution was observed for the other key residue 23 (Y for F). Residue 22 (A), between the core residues 21 and 23, allowed a wider range of substitution by other amino acids with small neutral side chains, and may therefore serve to maintain the correct spacing for the key residues in the presented peptide, while not necessarily interacting directly with the MAb.

This high-resolution mapping of residues directly involved in binding could not be achieved by using overlapping peptides of the same length [Coursaget et al., 1991], or by more complex studies using deletions in the coding sequence cloned into the RNA bacteriophage fr expression vectors [Sominskaya et al., 1992]. While complex omission- and alanine-substitution analysis of the critical residues recognized by another MAb gave comparable resolution [Appel et al., 1990], the result described now was obtained readily from the conservation of amino acids observed within a relatively small number of phage analyzed.

The relative dissociation constants for four of the cloned phage that bound the antibody (Fig. 2 and Table I, column 5) gave quantitative support to the definition of the amino acids that constitute the epitope. Substitution of the phenylalanine residue (23) by tyrosine lowered the strength of binding slightly, and phage E4-2, which carries only the core motif D, P, and F at residues 20, 21, and 23, respectively, had essentially the same $K_D$ as phage E4-4, which has two additional residues in common with the sequence that occurs in PreS1. The contribution of residues Q (18) and L (19) therefore can be, at best, only very small. The dissociation constants in Table I were obtained from measurements of the titer of phage eluted from antibody bound to a solid phase after equilibration with various concentrations of the antibody in solution, as described in Materials and Methods.

The competition studies with synthetic peptides (Fig. 3) showed that a hexapeptide (P1) with all six amino acids corresponding to the PreS1 epitope is an effective competitor against phage E4-4, which has the most conserved amino acids (five, QLDP and F). The $IC_{50}$ of the interaction of this peptide with the MAb is about an order of magnitude less than the $K_D$ for phage E4-4. The shorter peptide carrying the three highly conserved residues (P2) also competed effectively with phage E4-4 for the MAb, but at a substantially higher concentration ($IC_{50}$, about 10 μM), whereas an octapeptide of unrelated sequence showed no inhibition, even at 0.1 mM. Stereochemical factors of one form or another may well contribute to the relatively high concentrations of the peptides required to compete effectively with the fusion phage for the antibody binding sites; this has also been observed by others [Smith et al., 1993a; Yayon et al., 1993]. The phage filaments have a number (probably three or five) copies of the hexapep-
tide at their tip which may all bind with antibody molecules in close proximity to each other on the solid phase, and so give an enhanced binding affinity for the phage in relation to that of the free peptide. The free peptides may also occupy a larger population of conformations than their counterparts in the fusion phage, giving a lower effective concentration of a favorable conformation for interaction with the antibody. The need for a higher concentration was enhanced further when the tetrapeptide, P2, was used in the experiments described here, and might be a further reflection of the role of undefined stereochemical factors. But it could also reflect, in part, weak contributions of the additional residues of the larger peptide, or perhaps, more probably, destabilizing effects of charged residues at the two termini of the smaller peptide in close proximity to the binding site [Smith et al., 1993a]. Despite these effects it could be demonstrated that a peptide containing as few as three critical amino acids is sufficient for specific interaction with the MAb.

The results described here define the residues within the PreS1 sequence which are conserved within all the fusion phage analyzed which bind to the monoclonal antibody MA18/7, and in contrast to other work [Coursesag et al., 1991; Sominskaya et al., 1992] are able to distinguish between residues within a core motif essential for binding. The conclusion is that even in this case where the residues previously identified would probably be considered as a “linear epitope,” the critical residues for binding are not contiguous in the primary sequence.

ACKNOWLEDGMENTS

We thank Prof. G.P. Smith for generously providing the fusion-phage library, and Prof. W. Gerlich and Dr. K.H. Heermann for the MAb MA18/7. We also thank Dr. M. Dyson for valuable discussion and advice on the binding assays. V.G. was supported by the Darwin Trust of Edinburgh, and the work was supported in part by Biogen, Inc. (Cambridge, MA).

REFERENCES


Direct measurement via phage titre of the dissociation constants in solution of fusion phage–substrate complexes

Michael R. Dyson, Volker Germaschewski and Kenneth Murray*

Institute of Cell and Molecular Biology, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, UK

Received January 5, 1995; Revised and Accepted March 13, 1995

ABSTRACT

Studies of interactions between filamentous fusion phage particles and protein or nucleic acid molecules have gained increasing importance with recent successes of screening techniques based upon random phage display libraries (biopanning). Since a number of different phage are usually obtained by biopanning, it is useful to compare quantitatively the binding affinities of individual phage for the substrate used for selection. A procedure is described for determination of relative dissociation constants ($K_d$) between filamentous phage carrying peptide fusions to the coat protein gplll and substrates in solution. This novel method is based on the measurement of phage titres. Phage selected from a random fusion phage library for binding to a monoclonal antibody or a viral structural protein exhibited $K_d$ values in the nanomolar and micromolar ranges for their respective substrates, thus validating the method over a wide range of binding affinities.

INTRODUCTION

Libraries of phage displaying random peptide sequences are being used increasingly to identify sequences that interact specifically with a particular protein or nucleic acid. In some instances the substrate used for screening the library has been a monoclonal antibody and the system has been used for epitope mapping (1). The method can give very high resolution, defining the specific amino acids involved in binding the antibody (2). However, the power of the random approach lies in the discovery of amino acid sequences that bind to proteins or specific nucleic acid sequences about which little or nothing is known of potential binding sites (3–8). In such cases, ligands isolated from the library frequently show little or no primary sequence motifs of any known proteins. These so-called mimotope sequences presumably resemble the structure of a potential binding site and therefore bind the protein (9).

To identify mimotopes as specific ligands and subsequently compare them with other sequences in terms of affinity for the substrate used in screening it is necessary to examine the interaction quantitatively as well as qualitatively. The method described here provides direct comparison of the binding affinity of different phage for a substrate in solution without the need to synthesize peptides. That the assay is performed in solution is crucial, since the conformation of a protein can change on binding to a solid phase, with a resulting change of binding affinity (10).

We have used two groups of phage that were isolated from a random phage display library by screening with: (i) a monoclonal antibody specific for the PreS1 domain of the surface antigen of hepatitis B virus (HBV); (ii) a structural protein from the nucleocapsid of HBV, to determine the relative dissociation constants ($K_d$) for the interactions between the phage particles and their target proteins in solution. All phage were previously identified as specific ligands for the proteins used for screening the library and the selection procedures have been described elsewhere (2,8). The essential features of the method are that the substrate is incubated with the phage in solution until equilibrium is reached and the amount of unbound phage remaining in solution is then determined by adsorption on a solid phase coated with the substrate. The adsorbed phage are then eluted from the solid phase and titred, representing a value proportional to the free phage in the solution of the reactants at equilibrium. The titration of infective units (plaque forming units, p.f.u.) thereby offers a sensitivity that could not be achieved with most other techniques, as, in principle, one free phage particle can be detected. Scatchard plots of such results from a series of reactions at a range of protein concentrations deliver $K_d$ values.

MATERIALS AND METHODS

Phage particles and proteins

Phage particles B1–B4 are specific for the HBV core antigen particle (HBcAg) and were isolated from a random hexapeptide phage display library (11) by three rounds of biopanning against HBcAg adsorbed on nitrocellulose discs (8). HBcAg was expressed in Escherichia coli (12). Phage particles E2-16, E2-17, E4-2 and E4-4 are specific for the monoclonal antibody MA18/7, directed against the PreS1 protein domain of HBV surface antigen (13), and originated from rounds two and four of a
biopanning procedure using the same random hexapeptide phage display library and MA187-coated wells of an ELISA plate (2).

**Large scale preparation of phage particles**

Phage particles were purified from the supernatant of 500 ml overnight cultures of infected E.coli K91 Kan cells by polyethylene glycol precipitation and CsCl gradient centrifugation, as described by Smith and Scott (14).

**Determination of phage titre as plaque forming units (p.f.u.)**

Phage solutions were diluted serially in TBS (0.05 M Tris–HCl, pH 7.5, 0.15 M NaCl) containing gelatin (1 g/l) and aliquots of these dilutions (100 μl) were added to a suspension of K91 Kan cells (400 μl) at OD 600 0.2 (1/10 dilution). Melted soft agar (3 ml) was then poured into the tubes containing the suspension of phage and host cells and the mixture poured on agar plates containing kanamycin (100 μl). The plates were incubated overnight at 37°C and plaques counted the following day.

**Phage binding assay on solid phase**

U-shaped wells of ‘high capacity’ polystyrene strip plates (#2585; Costar) were coated with protein [20 μg/ml HbcAg in 25 mM potassium phosphate, pH 7.5, 150 mM NaCl (PBS) or 1 μg/ml MA18/7 in 0.2 M ammonium bicarbonate buffer, pH 9.6; volume 200 μl] overnight at room temperature, washed three times with TBS and blocked for 2 h at room temperature by adding TBS containing bovine serum albumin (BSA) (200 μl, 10 mg/ml). After three further washes with TBS, wells were filled with various concentrations of purified phage particles diluted in TBS buffer containing 0.2 mg/ml BSA, 0.02% NaN₃ (10⁻⁷–10⁻¹⁰ p.f.u./200 μl) and incubated for 1 h at 6°C. After 1 h the phage solutions were removed from the wells and the plates were then washed 10 times with TBS containing 0.2 mg/ml BSA and bound phage eluted by incubation with elution buffer (200 μl, 0.1 M HCl, titrated to pH 2.2 by addition of solid glycine) for 10 min. Eluates were neutralized with 40 μl 1 M Tris–HCl, pH 9.5, and the phage titres determined. Triplicate measurements were made at each phage concentration.

For time course experiments five identical sets of plates were prepared with coating conditions as above and the input phage concentrations were kept constant (200 μl, 10⁸ p.f.u./well). Each plate represented a point on a time course with incubation times of 10, 31, 100, 316 and 1000 min (corresponding to half-log intervals). After each incubation period the relevant plate was treated as described above and the bound phage eluted and titred.

**Phage binding assay in solution**

This procedure was adapted from Friguet et al. (15). First, the proteins (HbcAg or MAb) were incubated in solution with the purified phage particles until equilibrium was reached (18 h). In the case of binding of phage B1 to HbcAg this was attained as follows. The antigen at various concentrations (0.32–10 μM) was mixed with a constant amount of phage (10⁹ p.f.u./ml) in TBS containing 0.2 mg/ml BSA, 0.02% NaN₃. After 18 h at 6°C, an aliquot (100 μl) of each mixture was transferred to the wells of a U-shaped, polystyrene, radioimmunoassay strip plate that had been coated with HbcAg as described above. After 1 h at 6°C the wells were washed 10 times with TBS supplemented with 0.2 mg/ml BSA. Bound phage were then eluted, neutralized and titred as described above; all assays were performed in triplicate. In the case of binding phage B2 and B3 to HbcAg, the antigen concentration was varied from 1.58 to 50 μM and for phage B4 the antigen concentrations were varied from 0.63 to 20 μM. The binding assay between monoclonal antibody MA18/7 and phage was performed in the same way, except that the antibody concentration was varied from 0.32 to 30 nM and the solid phase was coated with MA18/7 as described.

**RESULTS**

The method for determination of relative dissociation constants between fusion phage and substrate molecules in solution is based upon that used for the measurement of affinity constants in solution between antigens and antibodies (15). A constant concentration of fusion phage was incubated with varying concentrations of substrate until equilibrium was reached. The number of free phage can be quantitated by incubating the mixture in microtitre plates coated with antigen or antibody, washing the wells, eluting the phage bound to the wells with glycine–HCl buffer, pH 2.2, and quantitating the phage in the eluate by titration of p.f.u.. To show that the phage concentration in the eluate reflected the free phage concentration in the reaction mixture, varying concentrations of phage (input) were incubated in microtitre wells coated with antigen or antibody for 1 h, the wells were washed and bound phage eluted (output) and titred. The output p.f.u. was directly proportional to input p.f.u. over a two log range for fusion phage B1 and E4-4 (Fig. 1), which differ greatly in their binding affinities (Table 1). The standard error involved in measuring phage titre varied between 5 and 10%.
Figure 2. Time course of phage binding to substrate-coated solid phase. Fusion phage (10^8 p.f.u.) E4-4 (●) or B1 (○) were incubated on polystyrene wells coated with either MA18/7 or HBcAg respectively for different periods of time and bound phage determined by subtracting the p.f.u. Points represent the average of three assays and error bars the standard deviation.

Table 1. Relative dissociation constants and characteristics of phage analysed

<table>
<thead>
<tr>
<th>Phage</th>
<th>gpIII fusion</th>
<th>Specificity</th>
<th>K_dRel</th>
<th>K_dRel</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>LLGRMK</td>
<td>HBcAg</td>
<td>0.15 ± 0.01 μM</td>
<td>0.17 ± 0.01 μM</td>
</tr>
<tr>
<td>B2</td>
<td>YLLRRF</td>
<td>HBcAg</td>
<td>1.4 ± 0.1 μM</td>
<td>1.5 ± 0.1 μM</td>
</tr>
<tr>
<td>B3</td>
<td>LLGRKK</td>
<td>HBcAg</td>
<td>1.1 ± 0.1 μM</td>
<td>1.1 ± 0.1 μM</td>
</tr>
<tr>
<td>B4</td>
<td>LLGRFK</td>
<td>HBcAg</td>
<td>0.56 ± 0.02 μM</td>
<td>0.64 ± 0.03 μM</td>
</tr>
<tr>
<td>E4-2</td>
<td>PDPPFN</td>
<td>MA18/7</td>
<td>0.17 ± 0.02 μM</td>
<td>0.2 μM</td>
</tr>
<tr>
<td>E4-4</td>
<td>QLDPPF</td>
<td>MA18/7</td>
<td>0.25 ± 0.02 nM</td>
<td>0.2 nM</td>
</tr>
<tr>
<td>E2-16</td>
<td>LDPPFR</td>
<td>MA18/7</td>
<td>0.10 ± 0.01 nM</td>
<td>nd</td>
</tr>
<tr>
<td>E2-17</td>
<td>DPAPND</td>
<td>MA18/7</td>
<td>0.22 ± 0.01 nM</td>
<td>nd</td>
</tr>
</tbody>
</table>

aItalic indicates matching amino acids with the sequence of the PreS1 antigen.
bRelative dissociation constants calculated by linear regression of Scatchard plots (Fig. 3).
cRelative dissociation constants calculated by curve fitting to a hyperbolic function (2,8).
dNot determined.

determined.

Figure 2 shows a time course experiment in which a constant concentration of free phage (10^8 p.f.u.) was incubated on substrate-coated wells for different time periods. A plateau was reached after 5 h, with ~14% of phage E4-4 bound to MA18/7-coated wells and ~4% of phage B1 captured by HBcAg-coated wells. These results presumably reflect the differing strengths of binding in the two cases.

The linear relationship between input phage titre and output of phage eluted from ligand-coated wells allows the determination of free phage concentration at equilibrium, provided that the total phage concentration is known. Indeed, if the total phage concentration (p_t) is incubated with the binding substrate at any given concentration, the free phage concentration (p_f) is related to the p.f.u. measured in the eluate by the following equation:

\[ p/p_f = \text{p.f.u.}/\text{p.f.u.}_o \]  

where \( p/p_f = \text{p.f.u.}/\text{p.f.u.}_o \) is the titre measured for the (eluted) phage in the absence of antigen. This will only be true if the equilibrium in the liquid phase is undisturbed during exposure of the mixture to the coated wells. It is unlikely that this will result in significant perturbation of the equilibrium in solution, since the number of phage binding to the solid phase under the experimental conditions used (i.e. quantity and accessibility of the protein substrate coated on the wells and time of incubation of the mixtures in the wells) represents <5% of the total free phage concentration (Figs 1 and 2). To verify this, fusion phage B1 and E4-4 were incubated at various concentrations (in triplicate) for 1 h with substrate-coated wells (HBcAg and MA18/7 respectively) and the contents of each well were then transferred to another coated well and again incubated for the same time. The bound phage in the two sets of wells were quantitated as described previously and found to be identical, within 95% standard error limits. This linear relationship between input and output p.f.u. from HBcAg-coated wells was also demonstrated for fusion phage B2, B3 and B4 (data not shown).

Since Equation 1 was shown to be valid and if one phage particle binds one antigen particle or antibody molecule, the concentrations of bound phage (x) and of free antigen or antibody (a) at equilibrium can be calculated from the mass conservation equations:

\[ x = p_t - p \]  

\[ a = a_0 - x \]  

where \( a_0 \) is the total concentration of antigen and \( p_t \) and \( p \) represent the total and free phage concentrations respectively. At equilibrium, x, a and \( p_t \) are related to the dissociation constant \( K_d \) by the Scatchard equation (16):

\[ x/a = [p_t - x]/K_d \]  

Equation 4 can be rearranged to give:

\[ x/p_t \ a = (1/K_d) - (x/K_d \ p_t) \]  

The gradient of a plot of \( x/p_t \ a \) against \( x/p_t \) is therefore equal to \(-1/K_d\). Figure 3 shows Scatchard plots for phage B1→4 binding to truncated HBcAg (Fig. 3A) and phage E2-16, E2-17, E4-2 and E4-4 binding to MA18/7 (Fig. 3B). \( K_d \) values were calculated from linear regressions of the plots and show the order of affinity of the hexapeptides displayed on the phage for the ligands against which they were selected (Table 1).

Once the linearity of the input to output ratio (Fig. 1) has been confirmed, up to five dissociation constants may be conveniently determined over a 2 day period, including data analysis. For ease of handling of larger numbers of plates, an image analyser may be used for counting plaques.

**DISCUSSION**

The method described here for measuring relative dissociation constants between fusion phage and ligands in solution should be generally applicable and enables direct comparison of the binding abilities of individual phage clones. However, dissociation constants obtained in this way should be regarded as relative values, since the presence of up to five copies of the gpIII fusion protein of the filamentous phage complicates matters. This could enable a single phage particle to simultaneously bind one molecule or particle on the solid phase and another molecule in solution or...
a single phage particle could bind to the ligand (particularly in the case of the HBCAg particle) via multiple interactions, thus enhancing the affinity. Further, since HBCAg and MA18/7 were present in large excess over the fusion phage, a single phage could bind more than one protein molecule or particle unless steric hindrance played a role.

However, these factors should remain constant when comparing individual fusion phage clones selected for a given substrate or a given phage clone for different ligands and the method described for measuring $K_d^{Rel}$ is valid because two essential criteria have been fulfilled. First, and most importantly, input and output phage titres are linearly related (Fig. 1), although the high affinity of phage E4-4 for the MAb apparently leads to saturation with inputs higher than $10^{8.5}$ p.f.u. This linear relationship between input and output phage concentrations has been observed in all cases that have been studied, which include examples that differ by four orders of magnitude in their apparent $K_d$ values. Secondly, in both cases <5% of free phage were bound to the wells during the 1 h incubation period used in the experiment (Fig. 2). Thus with both high and moderate affinity interactions it is unlikely that the equilibrium between the reactants in solution is significantly disturbed during this phase of the assay.

Competed with the high affinity interactions of phage E2-16, E2-17, E4-2 and E4-4 with the monoclonal antibody, the relative dissociation constants measured for the interactions of phage B1-4 with HBCAg represent moderate binding affinities (Table 1). This validates the method over a wide range (four orders of magnitude) of binding constants for reactions between phage and proteins. The $K_d^{Rel}$ values (Table 1), calculated by linear regression of the Scatchard plots (Fig. 3), are very similar to those obtained by a curve fitting calculation using the same data (2,8). The slight differences are probably attributable to the additional constraint introduced into the curve fitting that $p = p_t$ at zero antigen or antibody concentration.

Surface plasmon resonance has been used to evaluate the binding of fusion phage clones to a monoclonal antibody (17), but it did not appear possible to determine the rates of association or dissociation. Measurement of phage titre offers greater sensitivity compared with refractive index changes at the surface of the sensor chip. For example, phage concentrations greater than $10^{11}$ p.f.u./ml were required to observe a signal above background in the surface plasmon resonance experiments, whereas here concentrations of $10^9$ p.f.u./ml were employed and, in principle, this could be much lower. Additionally the determination of phage titres involves minimum expense.

The procedure described here for calculating relative dissociation constants between fusion phage and substrates in solution should find wide application in the growing field of selection of peptide ligands to complex macromolecules and assemblies by the use of phage display libraries (3,8,18). After the selection process has been completed and amplified phage clones sequenced, it is necessary to confirm that the selected phage bind to their substrate in solution and to compare relative binding affinities between the various phage clones. The procedure described here achieves both these objectives. Further, the choice of synthetic peptides for competition studies is facilitated by comparing binding affinities for the same ligand of fusion phage with varying amino acids in their displayed peptide sequences. Although examples do exist where the gpIII protein context contribute significantly to the binding affinity of the fusion phage for the substrate (19), it is expected that the relative order of binding of a set of fusion phage will be similar to that of the free peptides. In addition, for the examples discussed here the free peptides were synthesized and found to bind the selected substrates (2,8). The method may also find use in studying structure-function relationships where whole protein domains, such as human growth hormone (20), zinc finger domains (7,21) or antibodies (22), have been expressed on filamentous phage. Therefore, it should be possible to rank affinity variants in the phage-associated form without the need to purify free domains.

**ACKNOWLEDGEMENTS**

We thank Prof. G. Smith for the fusion phage library and Prof. W. Gerlich for the gift of Mab MA18/7. VG is currently supported by the Roseanne Campbell Trust for Hepatitis Research, Edinburgh. The work was also supported in part by Biogen Inc., Cambridge, MA, USA.

**REFERENCES**