Epitope Identification and Discovery Using Phage Display Libraries: Applications in Vaccine Development and Diagnostics

Lin-Fa Wang* and Meng Yu

CSIRO Livestock Industries, Australian Animal Health Laboratory, Geelong 3220, Australia

Abstract: Antigenic epitopes are the part (contact points) of an antigen involved in specific interaction with the antigen-binding site (the paratope) of an antibody or a T-cell receptor. Detailed analysis of epitopes is important both for the understanding of immunological events and for the development of more effective vaccine and diagnostic tools for various diseases. Identification and characterization of epitopes is a complex process. Although various methods have been developed in this area, there still lacks a simple common approach which can be applied to all epitopes. Since its first introduction more than a decade ago, phage display technology has made a major impact in this area of research. With the exponential growth in this area, it is impractical to review the entire literature detailing all possible applications. Instead, this review aims to focus on specific applications related to the discovery and identification of epitopes which have potential as vaccine candidates or can be used in disease diagnosis.

Key Words: epitope, mimotope, phage display, random peptide library, gene fragment library, diagnosis, vaccine.

1. INTRODUCTION

In the broadest sense, an epitope can be defined as the contacting points of any molecular interaction in a biological system, which is not restricted to protein-protein interactions. However, the term epitope is more frequently used to represent antigenic epitopes, i.e., the parts of an antigen involved in specific binding to an antibody (B-cell epitope) or T-cell receptor (T-cell epitope). Furthermore, unless specifically defined, the terms epitope and B-cell epitope are frequently used interchangeably. The region in the antibody responsible for the specific interaction with an epitope present on the antigen is called a paratope.

Determining the sequence or structure of a given epitope, frequently termed epitope mapping, has gained increasing popularity since the birth of monoclonal antibodies (mAbs). Epitope mapping is no longer restricted to immunology, and is instead being applied in all fields of biology. Epitope mapping is used in the study of protein-protein interactions, immunoassay development, elucidation of neutralizing sites, production of protective peptide vaccines, investigation into pathogenesis of autoimmune diseases, and investigation of protein topology in intact cells or organelles, just to name a few [1].

Although the concept of an epitope may not be that difficult to comprehend, it is much harder to experimentally define or map an epitope. Generally speaking, epitopes can be classified into two major groups, discontinuous and continuous epitopes. Discontinuous epitopes are also known as assembled or conformational epitopes in which the contacting points (amino acid residues within the protein) are located far apart in the sequence and are brought together by protein folding. In contrast, continuous epitopes (also known as linear or sequential epitopes) are composed of amino acid residues in close proximity in their primary sequence. Epitope mapping is further complicated by the fact that some so-called “linear” epitopes are actually part of a larger conformational epitope. In practice, most experimental scientists tend to define linear epitopes as those capable of reacting with their cognate antibodies in a Western blot analysis. It must be remembered, however, that few proteins are completely denatured on Western blots, and that protein renaturation can occur during or after electroblootning. Epitopes identified by Western blotting may therefore contain a considerable conformational element.

Another important point any “epitope mapper” should be aware of is the distinction between the structural and functional characteristics of any given epitope. As pointed out by van Regenmortel [2], structural epitopes are defined by identifying all amino acid residues in contact with the paratope of an antibody without the information on the relative importance of these contact points. X-ray crystallography is often regarded as virtually the only method to define a structural epitope. On the other hand, functional epitopes are elucidated by identifying individual amino acid residues which are important for binding and cannot be replaced with alternative amino acids. Most methods developed for epitope mapping are tailored for studying functional epitopes.

From the early work of Geysen and co-workers [3, 4], it was demonstrated that certain peptides, although having different sequences from the native epitope of an antigen, were able to interact with the paratope in a similar way as the native epitope. They classified such peptides as mimotopes. The mimotope concept played a major role in the evolution of random peptide libraries using phage display (see more discussion later).

As pointed out above, X-ray crystallography remains the only method suitable to accurately perform structural analysis of any epitope. However, there have been a large
number of techniques developed for mapping epitopes of proteins. These vary from physical methods, chemical methods, enzymatic methods, to a variety of different genetic methods. It is beyond the scope of this review to discuss the pros and cons of all these methods. Readers interested in different methods of epitope mapping will find the Epite Mapping Protocols [1] to be an excellent resource book to start with. This review aims to focus on the specific use of phage display libraries in epitope identification and discovery, and their applications in the development of vaccines and diagnostics.

2. PHAGE DISPLAY TECHNOLOGY

In 1985, George Smith demonstrated that a foreign protein sequence could be fused to the coat protein of a bacterial phage, a filamentous phage in this case, so that it is accessible or displayed on the virus surface [5]. Three years later, he further introduced the concept of affinity selection of phage displayed foreign target sequence from a library of displayed peptides, which ignited a burst of interest in the field, culminating in the simultaneous publication of three seminal papers describing the first generation of phage display random peptide libraries in 1990 [6-8]. Since then, phage display has rapidly developed into one of the most widely used research tools and the technology has been applied to almost every discipline of biological research [9-27].

The huge success of phage display technology is underpinned by the following factors. Insertion of a foreign sequence at an appropriate site within the phage coat protein results in the display of that sequence on the phage surface. If random sequences are inserted, it is possible to obtain a large library containing a vast number of different peptide sequences, each displayed on a different phage. Most importantly, the peptide sequence displayed on the surface, the phenotype, is physically linked to the encoding sequence in the ssDNA phage genome, the genotype. It is this link of phenotype and genotype which made it possible to select for a specific peptide, followed by subsequent amplification by re-infecting E. coli to obtain a virtually unlimited supply of specific peptides. The problems intrinsic to analysis of very small amounts of a selected chemical (peptide in this case) are thus resolved by the simple and cost effective process of viral replication [19]. It is effectively using combinatorial chemistry in a biological system. Because of this biological nature, it is paramount for anybody interested in the technology to have a full appreciation of the biology involved.

2.1. The Biology of Filamentous Phage

Several comprehensive reviews of filamentous phage biology have been published previously [24, 28, 29], and only the most relevant aspects will be discussed here.

The FII class filamentous phages, including strains M13, f1 and fd, are about 1 µm in length. The phage contains a circular ssDNA genome which is wrapped inside several thousand copies of the major coat protein pVIII. For a wild type phage genome of 6408 nt, there are 2670 copies of pVIII. At the “head” region of the phage, there are 5 copies each of pIII and pVI, important for host cell binding. At the other end or the “tail” region, there are 5 copies each of pVII and pIX that are required for the initiation of phage assembly and for the maintenance of phage stability [19]. Unlike other E. coli phages such as lambda phage, the filamentous phage genome is quite flexible and can accommodate large inserts of foreign sequences without grossly affecting its replication and packaging. Mature filamentous phage particles are extremely stable, can tolerate temperatures up to 70 °C for 30 min, and are resistant to 0.2 N HCl. These factors are important for the overall success of phage display technology.

As shown in Fig. 1, the life cycle starts by binding of phage, via pII, to the cell receptor for the initiation of infection. Only male E. coli strains, i.e., F-plus strains or cells harboring an F episome, are susceptible to infection, although F-minus strains are still capable of producing phage particles once the phage genome is introduced into the cells by other means such as chemical transformation or electroporation [24, 28]. Once the ss genome (the + strand) is injected into the cell, a complementary (-) strand is synthesized by the host polymerase to form a double stranded phage genome called the replication form (RF). Following this, host mediated protein synthesis results in the production of all ten phage-encoded proteins, including structural proteins (coat proteins pII, pVI, pVII, pVIII and pIX), proteins for assembly and export (pl and pLV), and proteins for replication (pII, pV and pX). Levels of expression are regulated by the grouping of gene clusters into two regions. Genes for proteins required in large quantities (e.g., pVIII) are preceded by a strong promoter while genes coding for other proteins (e.g., pIII) are controlled by a less efficient promoter.

Phage genome replication is initiated by pII, which nick the (+) strand at the origin of replication (ori+), allowing the freed 3’ end to serve as a primer for rolling circle replication using the (-) strand as a template. Phage protein pII also has a ligase activity which is responsible for formation of the circular progeny or (+) strand genome. The newly formed (+) genome is initially converted into a RF as for the parental genome until later in the infection cycle when nascent (+) strands become sequestered by increasing levels of the ssDNA binding protein pV, which triggers the termination of ss-to-RF conversion.

Virion assembly and export require the intricate involvement of host membranes. Following synthesis, all five coat proteins are inserted into the inner membrane with the aid of the bacterial export system, but only pIII and pVIII are synthesized as precursors containing a signal peptide at the N-terminus of the protein. The secretion of pIII is sec-dependent while the major coat protein pVIII uses a sec-independent pathway. For this reason, peptides fused with pVIII are more sensitive to charge distribution and the presence of large hydrophobic residues, and this is reflected in the relative paucity of charged residues in random peptide libraries constructed using pVIII-fusion [19]. For display using pIII or pVIII (Fig. 2), the signal peptide has to be cleaved before the mature protein is assembled into the final virion particle. The efficiency of cleavage is largely dependent on the sequence of the signal peptide, but is also affected by residues immediately following the cleavage site [19].

While ease and low cost of the biological amplification was a key element to the success of phage display technology over other in vitro combinatorial chemical methods, it is
Fig. (1). Life cycle of Ff filamentous phage. The drawing is highly schematic and the scale not proportional. Each of the five different structural proteins is represented in the figure as shown, but not in the actual stoichiometry of the mature virion. The circles represent viral DNA genome, the sense of the single stranded genome is shown (+ or -), and the double stranded replication form (RF) is shown as a double circle.

Fig. (2). Schematic diagram of different fusion strategies used to display foreign peptide by filamentous phage. In each diagram, the black region represents foreign peptide, the open box stands for the C-terminal region of the coat protein involved in virus assembly, and the dotted region is the signal peptide. For pIX and pVII fusion, a heterologous signal peptide (usually that of the pelB gene) is supplied during gene construction since the native pIX and pVII proteins do not have a signal peptide. For pIII, the N-terminal region, which is not required for virus assembly, is indicated by the slanted box.
also to some degree a limitation. As discussed in the review by Rodi and Makowski [19], any of the biological processes in the phage life cycle can hinder the successful display of a foreign peptide. These include the efficiency of codon usage in protein synthesis, the insertion of fusion proteins into the inner membrane, the correct and efficient cleavage of the signal peptide, the proper assembly of the phage virion, the required viral stability, the productive binding of phage with the host cell receptor and the correct disassembly following infection.

It should be pointed out that since the early application of filamentous phage display technology based on coat proteins pIII and pVIII, display systems using other filamentous phage proteins such as pIX have also been reported [30-37]. In addition, a number of other systems have been developed for the surface display of peptides/proteins including those based on bacterial phages such as lambda [38, 39], T7 [40], T4 [41, 42], P4 [43], MS2 [44, 45], and on surface proteins of bacteria, yeast, baculovirus and mammalian cells (see review in [22]). The pIII- and pVIII-based display, however, still remains the platform of choice by far, due to its robust nature and ease of use, and as a major source of successful applications in the literature.

Regardless of whether pIII or pVIII is used as the display platform, two types of phage display libraries have been used for epitope identification. These are described below.

2.2. Random Peptide Libraries (RPLs)

Since the publication of the three pioneering papers on RPLs in 1990 [6-8], there has been an exponential growth of interest in these libraries. The major advantage of this type of library is its universal nature. The same RPL can be used for mapping of epitopes recognized by different antibodies, for isolating mimotopes of proteins or carbohydrates, and its range of applications is only limited by one’s imagination.

In constructing a RPL, the peptides are encoded by synthetic oligonucleotides with each random residue (X) encoded by a degenerate codon. To reduce the introduction of stop codons, degenerate codons NNN or NNS (N = A, C, G or T; K = G or T; S = G or C), instead of NNN, are used which code for all 20 natural amino acid residues plus one stop codon (TAG). For example, the theoretical complexity of a 6-mer (X,) RPL is 64 million (20^6) peptide sequences that are encoded from approximately 1 billion (32^5) nucleotide sequences specified by (NNK)_6. The random peptides are genetically linked to coat-protein genes present in the phage genome. This allows the construction of libraries containing up to billions of virions, with each displaying and encoding a different peptide.

Although RPLs with insert lengths of between 6 and 43 amino acid residues have been reported in the literature [46, 47], the library complexity is currently restricted to approximately 10^3 to 10^10 due to the limited transformation efficiency of E. coli cells. RPLs with longer inserts may have some advantage in certain applications, such as in identifying mimotopes of conformational epitopes or any structural determinants that can not be mimicked by shorter peptides. It is also possible to view a longer peptide insert as an overlapping (sliding) series of smaller peptides. For example, a 20-mer peptide could be considered as 15 different 6-mer peptides or 14 different 7-mer peptides or 11 different 10-mer peptides and so on. It should be pointed out, however, that the length of peptide inserts generally has an adverse effect on the randomness of RPLs, especially when a pVIII protein display system is used [19].

It is generally believed that short peptide sequences displayed on the phage surface are presented in a non-structured form and that this may restrict their potential to mimic highly structured ligands. For this reason, various structured or constrained RPLs have also been constructed, most of them modified by the introduction of flanking cysteine residues to enable the formation of a rigid loop structure, similar to that developed for the commercial Ph.D.™-C7C library available from New England Biolabs. Although there have been some reports suggesting that constrained RPLs seemed to be more successful in certain specific applications [48-50], there is no general rule applicable as to what type of RPL is more suitable for a specific application [51]. It goes without saying that the success rate will be enhanced if one has access to multiple RPLs and has the resources to do parallel panning experiments against all different RPLs available [18, 20]. The role cysteine residues play in RPLs has been analyzed in detail by Zwick et al. [52]. They found that randomly picked phage clones containing a single cysteine tend to form dimers on the phage coat whereas those with two cysteine residues usually formed an intramolecular disulfide bonds instead.

Another type of constrained library is designed by taking advantage of existing protein scaffolds of known structure. Random peptides are inserted into the more flexible loop region of an otherwise rigid folded structure or domain so that the conformation of the random peptide is “constrained”. This type of library is therefore also known as a domain library [49, 53-56].

In addition to the choice of coat proteins (pIII or pVIII) used as a display scaffold, there are also different formats of display that can be adopted to suit different applications. For example, libraries can be constructed (see Fig. 3) so that all five copies of the pIII fusion protein contain the peptide insert or alternatively only a fraction of the structural proteins contain the insert. The latter can be constructed by the inclusion of a second non-fusion gene either in the same viral genome or through the duel vector system involving a fusion gene encoded on a phagemid and a non-fusion gene supplied by a helper phage. To distinguish these three formats of display, Smith named them “3”, “33” and “3+3”, respectively. The same nomenclature also applies to pVIII-based display formats, i.e., “8”, “88” and “8+8”, respectively [57, 58].

Due to the limitation of library complexity as stated above, it is often impossible to isolate peptide ligands of high affinity from any single RPL (i.e., original or primary RPLs). One way to improve the affinity is by the construction and screening of secondary libraries or sublibraries. In a successful panning experiment using primary RPLs, a consensus sequence usually containing 4-5 common residues can be derived. If a tighter-binding peptide is desired, further optimization can be achieved by constructing a secondary library, in which the conserved residues are fixed, and the
rest are totally randomized [59]. This \textit{in vitro} evolution process of peptide ligands has proven to be very powerful in increasing the specificity and affinity of antibody-mimotope interactions [60, 61].

2.3. Gene or Genome Fragment Libraries (GFLs)

GFLs differ from RPLs in several aspects. GFLs are constructed from randomly generated fragments of a specific gene or genome [62-66]. The size of inserts in GFLs is less uniform due to the fact that the inserts are derived from random digestion of genes or genomes by DNase or from random shearing by sonication. The insert sizes are usually larger than those in RPLs, and for that reason, GFLs are constructed almost exclusively using pIII-fusion. Unlike RPLs which are “universal”, GFLs are gene- or genome-specific and a new GFL must be constructed for each individual application. The major advantages of GFLs are that successful outcomes can be achieved with a relatively small library, especially when gene-specific application is of interest. GFLs are also useful in isolating conformation-dependent epitopes and can sometimes be more productive in isolating linear epitopes of a known protein [63, 67, 68].

2.4. Key Elements of a Successful Panning Experiment

As the saying goes, “you only find what you are looking for”. The design of a panning strategy is extremely important for the success of any experiment dealing with phage display libraries. In this context the saying “you are only going to find what is already in the library” is also true. The following is a brief discussion of the key elements and considerations in conducting a successful panning experiment.

\textbf{Library Complexity}

The application of a phage display library is often referred to as a “numbers game”. As pointed out above, if the desired peptide mimotope is absent from the library being used, even the best-designed panning strategy will fail. Library complexity is one of the most important factors determining the success rate of any applications. The size of any given library is a good measure of library complexity, but it is certainly not the only indicator. Due to the biological selection exerted on phage clones during the construction of a library, it is conceivable, and proven, that certain peptide sequences will not be present in a particular library. So the best way to maximize library complexity is to 1) increase library size (to the limit of current transformation technology), and 2) use libraries of different format as suggested by Scott and colleagues [18, 20]. One other consideration is the insert size. It is generally agreed that the minimum size of a linear epitope is 5-7 aa in direct contact with a paratope [2, 4] and this was one of the reasons that the 6-mer peptides inserts were used in the first generation of phage display peptide libraries [6-8]. However, a 6-mer library may not be ideal for isolating mimotopes of discontinuous epitopes or functional ligands for more
complex systems, such as discovery of cell surface receptors by *in vivo* panning [23].

**Panning Strategy**

The principle of a general panning strategy (show in Fig. 4) is simple and easy to follow, but to achieve the expected outcome of any panning experiment, careful planning is crucial. One common complaint encountered in this field is the isolation of “false positives”. For example, in identifying mimotopes of a mAb, it is often possible to find peptides which bind to the contaminating protein in the mAb preparation, to blocking reagents, to the plastic matrix used as the solid phase or even to the Fc region of the antibody, but not to the paratope of the antibody which was meant to be the “only” target of interest. All these could be “specific” binding peptides which were enriched under the particular experimental conditions used. So the problem lies at the difference between “what the experiment design was looking for” and “what one thought the experiment was looking for”. Other than cell-based panning or *in vivo* panning, it is paramount to make sure that the purity of the selector molecule (mAb, for example) is at the highest level possible. In addition, it is also worthwhile to consider the following:

1) using different solid matrix for different rounds of panning;
2) using alternative blocking agents;
3) using different elution strategies, and
4) using different infection methods for phage amplification.

**Affinity, Avidity and Copy Numbers**

The affinity of a specific ligand-selector interaction plays an important role in the overall success of any panning experiment. For most random peptide libraries, the “apparent affinity” or the avidity is determined by two factors. First, the absolute affinity of the ligand-selector interaction, which, albeit crucial, is difficult to measure experimentally. Second, the copy number of the peptides displayed per virion. Depending on the display format used (see Fig. 3), this number can vary from one to several hundred. For most applications, it is desirable to have a ligand with true high affinity, rather than one with apparently high avidity. One way to achieve this is to reduce the copy number by using the phagemid display format “3+3”. It should be pointed out that it is now possible to modulate the displayed peptide copy number without having to use a different display format. In the “3+3” display system, the peptide copy number can be adjusted simply by using different helper

![Diagram showing a typical panning experiment to identify peptide sequences sharing a similar motif.](image)

Fig. (4). Diagram showing a typical panning experiment to identify peptide sequences sharing a similar motif. The selector molecule (the “Y” shaped sign representing an antibody molecule in this case) is immobilized to a solid matrix (ELISA plate well, for example). After the addition of the phage peptide library, incubation is carried out to allow sufficient binding between selector molecules and mimotope peptides. This is followed by extensive washing to remove the unbound phage clones, followed by elution of the bound phage. The eluted phage are used to infect *E. coli* for amplification for use in the next round of panning. After the last round of panning, individual phage clones are isolated for further characterization. Specific binding of each clone with the selector molecule can be determined by a variety of techniques, including ELISA, Western blot and Biocore analysis. The sequence of a peptide insert in each clone is deduced from the DNA sequence corresponding to the coding region of the foreign insert. If successful, common motifs can then be identified as shown.
phage. With the development of HyperPhage™, it is possible to increase the copy number to five in the same “3+3” display format [69]. This has made it feasible to select for high avidity binders in the first round of panning by using HyperPhage™ and then bias towards high affinity binders in subsequent pannings by using conventional helper phage [69].

**Affinity Selection Versus Biological Selection**

Inherent to the phage display technology is the bias of growth competition found within any biological system, i.e., certain phage clones grow much better than others. In an ideal system, the enrichment of specific binders would be solely determined by the affinity of the ligand-selector interaction and the subsequent infection of and amplification in *E. coli* would only increase the total number of each binder selected without changing the ratio of different clones represented. However, it is well documented that growth competition poses a serious challenge in the overall success of phage display technology [19]. In each round of panning, affinity selection is followed by growth competition of the selected phage clones. It is possible that good affinity-selected binders are poor growers and will be overgrown by poor binders or non-specific binders which have much better growth characteristics. To solve this problem, several modifications have been reported to minimize this growth competition effect. These include direct screening of a large number of first round binders using the colony or plaque lift method [70], reducing rounds of amplification and carrying out statistical analysis on a large number of sequences [71], replacing liquid medium with solid media (on plate) for amplification and growing individual clones separately and then pooling them together for subsequent panning [72]. All of these methods are labor intensive and tedious to operate. So far there is really no easy alternative to circumvent this obstacle.

**3. PHAGE DISPLAY IN VACCINE DEVELOPMENT**

Among the various applications of phage display technology, vaccine development is one area that has attracted a great deal of interest, which is also the main focus of this review. Phage display in vaccine development has a wide range of applications, including the direct use of phage as an immunogen, identification of mimotopes, and study of T-cell specificity. Each of these is discussed separately in the following sections.

**3.1. Phage as Immunogenic Carriers for Vaccination**

Since the first demonstration that recombinant phage could be used as immunogenic carriers for foreign peptide antigens by de La Cruz *et al.* [73], there have been numerous reports of similar applications [20, 22, 26].

Use of recombinant phage as immunogenic carriers offers several advantages over more traditional protein carriers, such as ovalbumin, tetanus toxoid and keyhole limpet hemocyanin. Firstly, it is known that B cell epitopes on phage are limited to the outer 10-12 aa residues of pVIII [74], and the outer domain of pIII (only 4-5 copies per virion). Thus the antibody response to the carrier is much more restricted than to the other more frequently used protein carriers. Secondly, the conjugation through genetic fusion is easier, less expensive and more uniform than chemical conjugation methods. Thirdly, it is believed that filamentous phage of *E. coli*, a common bacterium found in humans and animals, may have evolved its proteins to induce a low-level response, so as to evade mucosally generated antibody responses [20].

In an early report by Greenwood *et al.* [75] it was demonstrated that adjuvant is not necessary for the induction of strong anti-phage antibody responses. More recently, a study by Demangel *et al.* [76] arrived at a similar conclusion. This has been attributed to the particulate nature of the phage virion as well as to the presence of significant amounts of *E. coli* LPS associated with purified phage preparations [20]. The effect of LPS, however, can be removed by treating phage with detergent before precipitation by polyethylene glycol and sodium chloride [20, 77]. For production of high affinity antibodies using low doses of phage (e.g., 10-25 µg per subcutaneous injection), it is recommended that adjuvant (such as the Ribbi Adjuvant System, Sigma) be used to enhance antibody responses [20].

Table 1 lists selected examples of applications in which phage have been successfully used as an immunogen carrier in immunization studies. These examples demonstrated that phage displaying epitope or mimotope could be used directly for the development of vaccines or as vaccine-delivery tools. It is generally believed that polyvalent display is more effective in eliciting an efficient immune response than monovalent display. The studies by Bastien *et al.* [78] and Zuércher *et al.* [79] further demonstrated that effective immunization could be achieved by direct oral administration of recombinant phage.

**3.2. Identification of Peptide and Peptide Mimics as Vaccine Candidates**

Among the many applications of RPLs and GFLs, one main application is to identify a peptide which can act as an immunogenic mimic eliciting an antibody response *in vivo* that is similar or identical to that used to isolate the peptide mimotope. While both RPLs and GFLs are useful for this kind of application, the use of GFLs is generally limited to the identification of continuous epitopes whereas RPLs can be used to isolate mimotopes for both continuous and discontinuous epitopes.

Although most of the published work describe the selection of mimotopes using monoclonal antibodies, polyclonal antibodies from patient sera have also been used successfully to identify mimotopes as potential vaccine candidates. For example, Folgori *et al.* [80] were able to isolate mimotopes of both continuous and discontinuous epitopes of the hepatitis C virus (HCV) using pooled sera from HCV-infected patients. They further demonstrated that immunization of rabbits with mimotopes produced antibodies which were able to react with HCV native antigens. In a similar approach, Scala *et al.* [81] used sera from HIV-infected patients to identify mimotopes from two different RPLs. Immunization in mice using phage clones displaying these mimotopes resulted in the production of antibodies capable of neutralizing HIV-1 virus in an *in vitro* infection assay using human peripheral blood mononuclear cells.
In identifying immunogenic peptide mimics, it is important to differentiate between functional and structural mimics. A functional mimic peptide binds to the same paratope as the native epitope, and can be demonstrated by competition binding assay using the native antigen. However, the exact contact points of the mimotope-paratope interaction may not be the same as that in the epitope-paratope binding. On the other hand, a truly structural mimicking peptide has identical contact points to that of the native epitope, especially for the so-called critical binding residues (CBRs). The significance of peptides which act as structural mimics of an epitope, versus purely functional mimics which make contacts with the antibody molecule via different residues from the native epitope, has yet to be clearly defined as it pertains to immunogenic mimicry [20]. Several groups have carried out studies using both RFL and GFL to compare their effectiveness in identifying epitope/mimotope sequences and to assess the affinity of such epitopes/mimotopes [62, 63, 82, 83]. In addition to the determination of linear epitopes in these studies, it has been shown that by combining both RPLs and GFLs it is possible to determine elements of a discontinuous epitope [83]. In another study, the immunogenicity of RFL- and GFL-derived peptides was compared. While peptide mimics of a model “pathogen” (T4 bacteriophage) were found from both types of libraries, the authors showed that only GFL-derived peptides were able to elicit antibodies that had a high affinity for the T4 phage tested using a stringent in vitro assay [84]. This study suggested that affinity-optimization (-maturation) may be required for RFL-derived peptides to achieve the immunogenic properties required for vaccine development.

In the development of vaccines for human or animal use, it is important to determine whether the use of peptides identified from RPLs or GFLs will be more advantageous than the use of whole proteins or killed/attenuated pathogens. Peptide or mimotope based vaccines are preferred when whole protein antigens or intact pathogens are too difficult or costly to produce, and when neutralization sites (epitopes) are not immunodominant. HIV is one such example. Although HIV infection induces strong antibody responses against the viral envelope proteins, very low levels of neutralizing antibodies are produced [85]. There have been numerous reports of the successful isolation of mimotopes using either mAb or patient sera. Many of them have been discussed in reviews [20, 22, 23, 26, 86].

3.3. Identification of Carbohydrate Mimics

An ideal vaccine should be able to effectively stimulate both humoral and cellular immune responses. This, however, is not the case for carbohydrate antigens. Most carbohydrates are intrinsically T cell-independent (TI) antigens because they do not require mature T cells to elicit a humoral response. Table 1. Induction of Immune Response by Immunization Using Recombinant Phage

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Displayed on</th>
<th>Results</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeat regions of circumsporozoite protein of Plasmodium falciparum</td>
<td>pIII</td>
<td>The displayed peptide was both antigenic and immunogenic</td>
<td>[73]</td>
</tr>
<tr>
<td>Peptides from major antigens Plasmodium falciparum</td>
<td>pVIII</td>
<td>The immune responses were highly specific to the displayed epitopes, and T-cell dependant and underwent class switching from IgM to IgG.</td>
<td>[132]</td>
</tr>
<tr>
<td>HIV-1 gp 120 antigen V3 loop</td>
<td>pVIII</td>
<td>High titer of antibodies, which were able to cross-react with other HIV strains and capable of neutralizing the viruses</td>
<td>[133]</td>
</tr>
<tr>
<td>Hepatitis B virus surface antigen mimotopes</td>
<td>pIII and pVIII</td>
<td>Mimotopes could induce antibody response resembling that induced by the original antigen, and the phage displayed mimotopes proved to be a better immunogen than synthetic mimotopes or the HBV core peptide</td>
<td>[134]</td>
</tr>
<tr>
<td>Mimotopes of Hepatitis B and C viruses</td>
<td>pIII and pVIII</td>
<td>Immune response induced by oral administration is specifically cross-reactive with the original antigen</td>
<td>[135]</td>
</tr>
<tr>
<td>Respiratory syncytial virus G protein epitope</td>
<td>pIII</td>
<td>High level of circulating antibodies detected, and achieved complete protection as indicated by the lack of virus in the lungs following intranasal challenge with live RSV virus</td>
<td>[78]</td>
</tr>
<tr>
<td>T-cell epitopes of HIV reverse transcriptase</td>
<td>pVIII</td>
<td>Strong cytolytic responses were obtained either by priming human cell lines or by immunization of transgenic mice.</td>
<td>[106]</td>
</tr>
<tr>
<td>Mimotopes</td>
<td>pIII and pVIII</td>
<td>All phage clones induced anti-IgE antibodies following oral immunization. After feeding a single dose of mimotope-displaying phage, phage DNA could be detected in mouse feces for 10 days.</td>
<td>[79]</td>
</tr>
<tr>
<td>Anti-idiotypic Fab of anti-IgE monoclonal antibody</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta-amyloid (Abeta) antigen for Alzheimer’s disease</td>
<td>pVIII</td>
<td>The immunization carried out in transgenic mice evoked effective auto-immune antibodies and provoked a considerable reduction in the number of Abeta amyloid plaques in the brain.</td>
<td>[136]</td>
</tr>
</tbody>
</table>
response in vivo, which diminishes their efficacy as vaccine antigens [87]. This remains one of the hurdles that hinder the development of carbohydrate-based vaccines even though carbohydrate antigens have long been identified as immune targets associated with a variety of pathogens and tumor cells. The other difficulties in developing carbohydrate vaccines include the lack of a means of genetic manipulation and the lack of de novo synthesis or design methodologies.

Traditionally, the most common approach relies upon chemical conjugation of a carbohydrate to a protein carrier to convert from a totally TI antigen to a TD (T-dependent) antigen. This approach requires a large amount of purified carbohydrate, which is not always feasible, suffers with batch-to-batch variation, and sometimes alters the immune response to the carbohydrate because of the TD response to the carrier protein [88].

There has been increasing interest in molecular mimicry to facilitate vaccine design and development for carbohydrate immune targets. There are currently three approaches being adopted in this area of research, (1) anti-idiotypic (anti-Id) antibodies that mimic carbohydrate antigens, (2) use of combinatorial peptide libraries to identify peptide mimotopes for carbohydrates, and (3) 3-D modeling to design structural mimics of carbohydrates which can be de novo structure-based design, or more frequently lead molecule-assisted design. Screening of peptide libraries remains the most effective method of deriving lead molecules [87].

Rapid identification of peptide mimics using phage displayed RPLs represents the most cost-effective and efficient method to date. Although the majority of applications uses carbohydrate-specific mAbs to select for peptide mimics, it is also possible to use polyclonal serum to select disease-related carbohydrate mimotopes (S.S. Tang, K.-L. Thong, L.-F. Wang, unpublished results).

Vaccines developed using peptide mimics provide two major advantages over traditional carbohydrate-protein carrier approaches. Firstly, peptide mimotopes are more likely to be formulated to enhance not only humoral, but also cell-mediated immune responses. Secondly, peptide mimotopes are amenable to the development of DNA-based vaccines [89, 90]. Ease of manipulation at the DNA level will also open up the ways for co-delivery of multiple copies of the same or different carbohydrate mimotopes using technologies such as the polytope approach [91, 92].

A large number of carbohydrate mimotopes have been identified using phage displayed peptide libraries. Although most of the reports focus on the identification and demonstration of similar binding characteristics to the native carbohydrate antigen, there have been many reports of successful immunizations using peptide mimotopes. A few selected examples are listed in Table 2, and more examples can be found in the literature [87, 93].

### 3.4. Identification of Immuno Targeting Binders

Although the factors determining the immunogenicity of particular antigens in vivo are still not fully understood, it is believed that targeting of antigens to lymphatic organs can enhance immune responses if delivered properly. Over the

<table>
<thead>
<tr>
<th>Carbohydrate</th>
<th>Mimotope</th>
<th>Immunization</th>
<th>Findings</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>O-antigen of <em>Shigella flexneri</em> serotype 5a</td>
<td>9-mer peptides identified from two libraries using two mAbs</td>
<td>Phage displaying mimotopes used directly for immunization in mice</td>
<td>Out of 19 mAb-reactive phage clones, two were able to induce O-antigen-specific antibodies, representing the first demonstration of a specific antibody response to carbohydrate antigen from phage-mimotope immunization</td>
<td>[137]</td>
</tr>
<tr>
<td><em>Neisseria meningitidis</em> serogroup C polysaccharide (MCPS)</td>
<td>Peptides derived from the heavy chain CDR3 region of an anti-ID mAb</td>
<td>Peptides with lauroyl group were complexed with OMP as carrier and adjuvant and used to immunize hu-PMBC/SCID mice</td>
<td>Specific human antibody responses to the MCPS were detected by inhibition ELISA and bactericidal assay.</td>
<td>[138]</td>
</tr>
<tr>
<td>Tumor-associated carbohydrate (TAC)</td>
<td>Mimotopes isolated from library or designed using the W/YXX motif</td>
<td>Multiple-antigen peptides (MAP) were used to immunize mice with adjuvant QS-21</td>
<td>In addition to specific antibody responses to human TAC antigens, vaccination of mice reduced tumor growth and prolonged host survival in a murine tumor model</td>
<td>[139]</td>
</tr>
<tr>
<td>Outer-membrane lipooligosaccharide (LOS) of group B <em>Neisseria meningitidis</em></td>
<td>7-mer peptides identified using a mAb</td>
<td>Peptide alone, peptide-diphtheria toxoid conjugate or phage displaying peptide were used as immunogen</td>
<td>Mice immunized with the peptide-carrier gave the best responses, with a total IgG response to LOS 2-4 times higher than the control group</td>
<td>[140]</td>
</tr>
<tr>
<td>Polysaccharide capsule of <em>Cryptococcus neoformans</em></td>
<td>High affinity mimotopes were identified from an evolutionary library</td>
<td>Mice were immunized with peptide-tetanus toxoid (TT) conjugate with adjuvant</td>
<td>No specific antibodies to the glucuronoxylomannan (GXM) component were detected using peptide-TT alone. However, mice primed with GXM-TT, and boosted with peptide-TT developed significant anti-GXM antibody with a maximum titer of 1:180,000</td>
<td>[61]</td>
</tr>
</tbody>
</table>
years, many different approaches have been attempted by various groups to achieve this. These include direct inoculation of antigens into the lymph nodes [94] and spleen [95], or targeting of specific receptors expressed on the surface of dendritic cells [96] and other antigen-presenting cells [97]. Although the results are mixed with regard to the effectiveness of these approaches, there have been some encouraging reports on novel targeting strategies.

In 1998, Boyle et al. [98] demonstrated that an enhanced immune response could be achieved by directing antigens to the sites involved in immune induction. Using human IgG (hIgG) as a “model antigen”, they constructed different fusion protein constructs so that the CH2 and CH3 domains of hIgG were linked to specific targeting molecules, L-selectin and cytotoxic T-lymphocyte antigen 4 (CTLA4), respectively. L-selectin is expressed on the surface of naïve lymphocytes and, by binding to CD34 located on high endothelial venule cells, initiates lymphocyte entry into lymph nodes [99]. CTLA4 present on activated T cells binds to B7-expressing cells, including antigen-presenting cells (APCs), which are potent initiators of immune responses [100]. After immunization of BALB/c mice with DNA constructs expressing these fusion proteins, a significant increase in the immune response to the “model antigen” was observed, with CTLA4-targeting giving the highest response. At two weeks post immunization, the CTLA4-targeted group recorded specific antibody levels 10,000-fold greater than the non-targeted group, and the levels reached a plateau at week 8 with a 1,000-fold difference between the CTLA4-targeted and non-targeted groups. They also demonstrated an enhanced T-cell proliferative response in the targeted group [98]. A similar strategy, i.e., DNA vaccination combined with CTLA4 targeting, was used to conduct a vaccine trial in sheep against Corynebacterium pseudotuberculosis infection [101]. The study demonstrated that the targeted DNA vaccine performed better than the non-targeted group, and that the protection level was similar to that provided by a formalin-inactivated subunit vaccine. Indeed, this was the first animal trial that clearly showed the efficacy of DNA vaccination in an outbred domestic species such as sheep [101].

Since the first report of the so-called “in vivo panning” for selection of organ-targeting peptides using RPLs by Pasqualini and Ruoslahti in 1996 [102], there has been a great deal of interest generated in this area of research. Although much of the attention has been directed towards applications in gene therapy and cancer therapy [103], the work by Trepel et al. [104] demonstrated that it is possible to modulate the immune responses by systemic targeting of antigens to lymph nodes. Using the in vivo panning strategy developed in the same group, the authors identified two lymph node-targeting peptides from a RPL with the insert X_CX_CX (C, cysteine, X, any amino acid residue). When phage displaying these peptides were used to immunize mice via intravenous injection into the tail vein, a significant increase in anti-phage antibody titer was obtained in comparison to control mice receiving the control phage containing no peptide insert. They further demonstrated that the observed increase in antibody response was dependent on the function of the targeting peptides. This was done by intravenous injection of 1 mg cognate synthetic peptide 5 min before the injection of phage. The results indicated that the increased antibody response could be “neutralized” by pretreatment with the cognate peptides corresponding to the displayed targeting peptide sequences.

The proof of principle for peptide-based immunotargeting in the above-described work may have broad implications in the development of improved vaccines, production of antibodies, and immunotherapy. The next challenge will be to find commercially viable methodology to link these immunotargeting peptides with vaccine antigens. There are several approaches to achieve this. One approach is to use phage as a delivery vehicle, and protein-based vaccine antigen can be delivered through the co-display strategy shown previously [105, 106] so that the targeting peptides are displayed via pIII fusion and the vaccine antigens are displayed via pVIII fusion. The advantage of this approach is that the functionality of the targeting peptides will not be influenced by conjugation with different vaccine antigens. The other approach is chemical conjugation. The targeting peptides, in a synthetic form, can be linked to various antigens by direct chemical conjugation. The advantage of this approach is that it is applicable to all antigen forms, and not limited to protein-based antigens. However, the targeting capability of synthetic peptide may be affected by the conjugation process and/or by the size and nature of the molecules attached to it. The third approach is genetic fusion, i.e., linking of the targeting peptide with a protein-based antigen by recombinant DNA means and achieving immunization either by DNA immunization or by injection of purified recombinant fusion proteins.

3.5. Study of T-cell Epitopes

To date, there is no direct way to determine the antigenic specificity of T-cells. While B-cell epitopes, whether continuous or discontinuous, can be selected from RPLs and/or GFLs, equivalent molecular tools for mapping T-cell epitopes are not yet available. However, phage display technology has nevertheless been used in different aspects of T-cell studies. One of the early works by Hammer et al. [107] used a purified recombinant HLA-DR1 molecule to select for peptides from a RPL. First, they demonstrated that recombinant phage clones containing peptides known to bind HLA-DR1 were able to bind HLA-DR1. They then conducted four rounds of panning using the same HLA-DR1 molecule and a 9-mer RPL, and found that 80% of the phage clones selected were able to bind the selector molecule. Sequencing analysis of 60 individual clones revealed two potential anchor positions, the first with an aromatic residue (Y, F or W) at the N-terminus, and the second (M or L) three residues downstream [107]. In an extension of this work [108], RPL was used to identify destabilizing residues in HLA class II-selected phage clones edited by HLA-DM. HLA-DM functions as a peptide editor by catalyzing the release of class II-associated invariant chain peptides and other unstable peptides, thus supporting the formation of stable class II-peptide complexes for presentation. Using a novel double selection procedure, the authors selected clones capable of binding to HLA-DR1 first, followed by specific elution with purified HLA-DM molecules. Their results demonstrated that HLA-DM susceptibility was dependant not only on the number and nature of anchor residues and peptide length, but more importantly also on subtle sequence...
characteristics which are less obvious to define [108]. RPL was also used to validate a peptide motif responsible for binding to the insulin-dependent diabetes mellitus-associated class II MHC molecule I-A\(^\beta\) [109].

Encouraged by the findings that some T-cell receptors (TCRs) can specifically bind to their cognate peptides in the absence of MHC molecules, Onda et al. [110] experimented with a phage display approach to demonstrate TCR-peptide interaction. They displayed the TCR\(\alpha\) subunit via pVIII fusion, and demonstrated that specific enrichment could be achieved, mimicking a panning process, by mixing two TCR\(\alpha\)-phage clones derived from T cells with different peptide specificities. While this study illustrated the possibility of being able to select and analyze specific TCR clones from TCR-displaying phage libraries, it should be pointed out that binding of TCR\(\alpha\) with peptide in the absence of TCR\(\beta\) and MHC is not a common feature of TCR-peptide interaction, hence the application of such a library may be limited.

More recently, Le Doussal et al. [111] demonstrated that it is possible to express a functional peptide-MHC complex on the phage surface, thus leading to the possibility of direct T-cell selection. They achieved this by expression of a peptide fused with a single chain MHC molecule (H-2K\(^d\)), displayed on phage surface via the pIII protein. Specific binding was demonstrated using purified and immobilized recombinant T cell receptor (TCR) as well as TCRs expressed on the surface of a T-cell hybridoma. A similar study was also reported by Kurokawa et al. [112] for the functional display of peptide-MHC class I molecules HLA-B*5101 and H-2D\(^\beta\). The functional display of scMHC-I and its binding with the antigenic peptide was elegantly demonstrated by protease cleavage, followed by E-tag mediated capture as shown in Fig. (5). Although both groups used defined peptide sequences of known TCR specificity in their proof-of-principle work, it is envisaged that libraries of random peptides fused to MHC molecules should be applicable for determining the spectrum of peptides that are recognized by a particular T-cell or TCR, complementing the tetrameric MHC-peptide approach. These peptide-MHC phage clones may also be useful for inducing immune responses in vivo [20].

4. IDENTIFICATION OF DISEASE-SPECIFIC MARKERS

Most diseases, especially infectious diseases and autoimmune diseases, leave a specific imprint on the complex mixture of antibody specificities present in the total serum antibody population. Within this collection of myriad antibodies, some are specific to a particular disease, some are elicited directly by antigens (e.g., an invading pathogen), while others may be recognizing antigens that reflect the disease process more indirectly (e.g., antibodies in autoimmune diseases) [21]. Disease-specific antibodies have long been used in clinical diagnosis, but such applications are traditionally limited to those antibodies for which their cognate antigens are easy to identify and prepare.

With the advances in phage display technology, it is now possible to make “epitope discovery” using patient sera with or without previous knowledge about the original antigen

Fig. (5). Functional display of single chain MHC-peptide fusion protein. A gene cassette containing the following gene segments (from N- to C-terminus) is constructed: signal peptide, T-cell epitope (peptide), E-tag, thrombin cleavage sequence (XIII), flexible linker, MHC-I heavy chain (\(\alpha_1, \alpha_2\) and \(\alpha_3\)), human \(\beta_2\)-microglobulin (\(\beta_2M\)) and pII gene. The functional association of scMHC-I and peptide on phage surface is demonstrated by cleavage using thrombin digestion to release the peptide from the single chain MHC molecule (scMHC-I), followed by capture of the whole phage complex by the anti-E-tag monoclonal antibody. Diagram modified from [112].
that induced the disease-specific antibodies. The key for success relies in the design of a selection strategy that ensures the enrichment of disease-specific mimotopes. As one can imagine, for most patient sera there are more non-disease-specific antibodies than disease-specific ones in any given serum antibody population. Counter-selection or pre-absorption with control sera is often essential in such panning experiments [80].

Both RPLs and GFLs have been used for the identification of disease-specific peptides depending on whether the original antigen(s) is known or not. In the case of infectious diseases caused by a known pathogen (e.g., HCV), either RPLs or GFLs can be used. On the other hand, for autoimmune diseases for which the autoantigen is unknown, RPLs will be the only choice.

Epitope identification or discovery using phage displayed peptide libraries has several advantages over other approaches. First, as stated above, it is not essential to have any previous knowledge about the “immunogen” which induced the disease-specific antibodies in the first place. Second, the same set of RPLs can be used for any and every disease of interest, making them highly “universal” or “portable”. Third, this technology is cheap and easy to run, and does not require expensive equipment. Fourth, it is often possible to identify multiple mimotopes that have slight differences in reactivity with different patient sera, making it possible to perform “finger printing” analysis of immune responses by different individuals. Fifth, it has also been demonstrated that it is possible to improve the qualities of a mimotope to better suit their diagnostic suitability, by a process similar to that used for the affinity maturation of antibodies, which has never been attempted by any other approach on antigen molecules [60, 113].

There have been numerous reports in the literature detailing the application of phage display technology for the identification of disease markers and for improvement of disease diagnosis. Three different types of applications are discussed below with selected examples.

4.1. Epitope Discovery for Diseases Caused by an Unknown Etiological Agent

One of the key obstacles in the investigation of autoimmune diseases is the difficulty in elucidating the etiology of the disease process. Phage display technology has made a significant contribution in this area. The first proof-of-principle showing specific identification of “disease-related” mimotopes was obtained in 1994 by Folgori et al. [80]. As a model system, they used human sera from a normal population and from a “diseased” population (in this case, the “diseased” population refers to those who have received vaccination against hepatitis B). After selection and counter-selection of a 9-mer RPL, they were able to identify two disease-specific mimotopes that cross reacted with sera from a large population of immunized individuals. They further demonstrated that mice immunized with these mimotopes elicited a strong antibody response specific for the hepatitis B antigen.

In order to extend this work to other diseases with unknown etiology, Cortese et al. [114] screened an RFL with the cerebral spinal fluid (CSF) of multiple sclerosis (MS) patients. Although MS patients are known to have high levels of oligoclonal antibodies, their specificity and role in the disease process are poorly understood. After panning, phage displaying selected mimotopes were shown to react not only with the CSF, but also with the serum of the original MS patients whose CSF was used in panning. However, when CSF antibodies of other MS patients were tested, only some showed positive cross reactivity, indicating that the selected mimotopes were not shared by all MS patients. In a further study, Cortese et al. [115] identified dominant mimotopes recognized by different MS patients which contained the common motif LPPNP. In searching for a potential viral etiological link, they tested cross reactivity of the MS antibodies with a panel of neurotropic viruses and found cross reactivity with one epitope in the envelope glycoprotein B of HSV-1. They also demonstrated that rabbit antisera raised against one mimotope peptide was able to bind to a brain-specific protein in Western blot, identifying a putative target for HSV-1-induced autoantibodies. This important discovery may, for the first time, point to the potential link of a virus infection acting as a trigger for the production of CSF antibodies in MS patients.

Similar approaches have been used in discovering mimotopes associated with a variety of other diseases including rheumatoid arthritis [116], systemic lupus erythematosus (SLE) [117], insulin-dependent diabetes mellitus (IDDM) [118], primary biliary cirrhosis (PBC) [119,120], and multiple myeloma [121], just to list a few.

4.2. Identification of Biomarkers for Diseases with Complex Etiology

For certain diseases, such as cancer, diagnosis has not been easy due to the difficulty in identifying disease-specific biomarkers, although the cause of the disease (tumor, in this case) is well defined. It has long been believed that cancer-specific antibodies may exist in the complex antibody repertoire of cancer patients, but identification of such antibodies is usually difficult, if not impossible. In a recent report by Somers et al. [122], the authors demonstrated that it is possible to rapidly identify antigen markers which are recognized by the humoral immune response in patients with colorectal cancer (CRC). First, they constructed a cDNA phage display library using cDNA derived from the CRC cell line HT-29. The expression and display of polypeptides encoded by random cDNA fragments was achieved by fusion with the minor coat protein pV1. Pooled CRC patient sera was used to enrich for clones containing CRC-specific antigens. A total of 19 clones were identified which were reactive with the serum pool. Seventeen clones (89%) showed reactivity with one or more of the eight sera in the pool. In addition, six of the eight sera (75%) were reactive with at least one of the 19 clones. These newly identified antigens may be used as candidates for sero-diagnosis of CRC, as prognostic markers, as well as vaccine candidates or as probes for monitoring tumor cell-based vaccination trials.

There are some infectious diseases for which the causative pathogen may be hard to grow or the immunodominant epitopes are difficult to define due to the lack of specific antibodies and/or antigens. In these cases, serum antibodies
could be used directly to identify disease-specific peptide markers as a fast, specific and inexpensive diagnostic tool. One such example is the identification of Lyme disease-specific mimotopes using patient sera and RPLs. Kouzmitcheva et al. [123] used serum IgG isolated from Lyme disease patients to screen 12 different RPLs. Although all the mimotopes selected represented mimics of discontinuous epitopes (thus none matched any known protein sequences of the causative agent Borrelia burgdorferi), they were shown to be reactive with a number of patient sera which were not used in the initial selection. This demonstrated that peptides selected from RPLs might provide a simple diagnosis for infection in Lyme disease.

### 4.3. Improvement of Disease Diagnosis Using Engineered Epitopes or Mimotopes

Since the discovery of hepatitis C virus (HCV) as the major etiological agent of non-A, non-B hepatitis, different approaches have been explored to develop a sensitive and specific antibody detection test to monitor HCV infection. Various phage display strategies have been used to identify disease-specific mimotopes for diagnostic applications. The early work in Monaci’s group demonstrated that it is feasible to select HCV-specific mimotopes using patient sera [113, 124]. However, the diagnostic potential of these mimotopes was limited due to the fact that different mimotopes reacted with different sets of patient sera. Two subsequent strategies were adopted to improve their performance in detecting HCV-specific antibodies. The first utilized a process of in vitro maturation by constructing and screening sublibraries in which the DNA insert coding for the binding peptides were randomly mutated by error-prone PCR followed by the introduction of additional random sequences flanking the original inserts. New phage clones with higher affinity and broader reactivity were obtained which reacted with a larger panel of HCV sera than the original clones [60, 125]. The second strategy involved the use of multiple antigen peptides (MAPs) that link different mimotopes derived from the above studies [126]. Thus by the combination of improved affinity and increased diversity, the authors were successful in developing a new concept in antibody detection, which was termed ADAM (antibody detection by antigen mimics).

Using a different strategy, Pereboeva et al. [127] demonstrated that improved diagnostic antigens could also be obtained using the gene fragment library approach. Using a linker-ligation PCR method [128], a GFL was constructed using cDNA coding for the NS3, NS4 and NS5 proteins of HCV. After panning with a mixture of sera from five HCV-seropositive individuals, epitope fragments were identified for each of the three non-structural proteins. Using the information derived from the panning experiments, an engineered chimeric NS4 antigen was produced which incorporated two separate epitope fragments revealed in the study. When tested in an ELISA assay with 35 HCV positive sera, the new chimeric antigen not only increased the assay sensitivity, but also detected more HCV-positive sera than when these epitope fragments were assayed separately.

As pointed out earlier, prior to the introduction of phage display technology, epitope mapping was, to a large degree, limited to those defined by monoclonal antibodies. This approach limited use in disease investigation since antibody responses in the mice are not always the same as those induced in the target species, such as humans or other animals. For this reason, identification of disease-specific epitopes or mimotopes using polyclonal sera has been an important area of research in recent years. In an early study, our group mapped several epitope fragments using a GFL of the major protein P72 of African swine fever virus and sera from pigs infected with the virus [64]. More recently, several groups have also reported success in identifying mimotopes from RPLs using post-infection animal sera, for example for equine arteritis virus [129], equine herpesvirus type 1 [130] and HIV [131].

### CONCLUSION

Although it has been more than a decade since the introduction of phage display libraries, their full potential in vaccine development and disease diagnosis is yet to be realized. As briefly summarized in this review, the contribution of phage display technology to epitope studies goes far beyond just another approach in this field. The wide ranging impact of this technology originates from its simplicity, universality and the ability of discovering as well as identifying epitopes. The power to research into previously not readily answerable questions distinguishes phage display from other methods. However, as for most emergent new technologies, phage display also has its problems and difficulties. The limitations posed by phage and host biology may be greater than most of us have expected or hoped for. Its application for T-cell epitope investigation is less than impressive at the present time. It is clear that a better understanding of structural and functional features of mimotope (peptide)-antibody interaction and further elucidation of the restrictions imposed by the biology of the system will enhance the success rate and applicability of this already very powerful technology.

### ACKNOWLEDGEMENTS

We are grateful to Drs. Sandra Sapats and Hans Heine for critical reading of the manuscripts, and to Eric Hansson, Kaylene Selleck, Nadia Mayfield and Julia Hammond for technical assistance in various phage display-related projects.

### REFERENCES

Identification and Application of Peptide Mimotopes

Current Drug Targets, 2004, Vol. 5, No. 1

Identification and Application of Peptide Mimotopes


