

Phage Display Technologies

Nicole Stricker, *Johns Hopkins University, Baltimore, Maryland, USA*

Min Li, *Johns Hopkins University, Baltimore, Maryland, USA*

Phage display is a method that uses bacterial virus (phage) as a vehicle to express diverse protein or peptide sequences as part of the phage coat protein by cloning deoxyribonucleic acid (DNA) fragments in frame with phage coat protein genes. Upon viral infection, the expressed peptides are displayed on the surface of the viral particle. Depending on the DNA fragments cloned, phage display allows for expression of either short peptides or large proteins, such as immunoglobulin, on the viral surface.

Introduction

Phage display is a powerful technique for identifying and characterizing interactions of recombinant polypeptides on the surface of bacterial phage particles with their binding targets, including, but not limited to, proteins. Its potential applications include identification of 'lead' molecules for therapeutical macromolecular targets, selection of high-affinity antibodies, characterization of antibody-binding sequences, and investigation of potential ligands for orphan molecules. The conceptual framework of this technology is essentially to tag individually a large number of peptides or proteins with their corresponding nucleotide coding sequences, thereby allowing for selective amplification of sequences that possess certain properties, such as binding to the target molecule. This article provides a general description of this technology, its common applications and modified versions of the technique.

Filamentous Phage Coats are Multivalent Surfaces

Filamentous phage, such as M13, infect hosts by injecting their deoxyribonucleic acid (DNA) through a protein fibre structure of bacteria, the pilus. To initiate infection, the phage bind to the tip of the F pilus on male bacteria. The binding induces retraction of the pilus and leads to injection of the single-stranded DNA of the phage into the bacterial cell. The phage DNA is replicated and subsequently packaged. Unlike most other phage, filamentous phage are not released by cell lysis but rather by secretion from the infected host, at roughly 100 copies per cell division. As a result, the infection of filamentous phage such as M13 slows down cell division considerably, thereby producing a turbid plaque containing roughly 10^8 infective phage particles. Primarily for this reason, the filamentous phage or its derivatives are commonly titred by the ability to confer antibiotic resistance as colony-forming units (CFUs) or transducing units (TUs).

Secondary article

Article Contents

- Introduction
- Filamentous Phage Coats are Multivalent Surfaces
- Gene III Protein
- Gene VIII Protein
- Library Construction and Identification of Recombinant Phage of Interest
- Peptide Ligands
- Phage as Alternatives to Antibodies
- Epitope Mapping
- Other Applications
- Summary

The protein capsid of filamentous phage comprises several coat proteins, including pIII and pVIII, which are encoded by genes III and VIII, respectively. The pVIII is the major coat protein that wraps around DNA. There are about 3000 copies of pVIII protein per phage particle, depending upon the length of phage genomic DNA. In contrast, the minor coat protein, pIII, is present at only 4–5 copies per phage, and it plays an important role in phage–host interaction during infection (**Figure 1**). Both proteins contain an N-terminal signal sequence that directs them to the bacterial inner membrane, where signal sequences are cleaved by signal peptidase. After cleavage, the mature proteins are recruited to specific sites where they are assembled to viral coat, with their C-terminal regions attached to the viral core and the N-terminal domain positioned distally from the phage particle.

The N-terminal portions of both major pVIII and minor pIII coat proteins are essential for targeting newly synthesized proteins to the inner membrane for viral assembly, but they do not directly contribute to the stability of the assembled phage particles. Foreign peptides

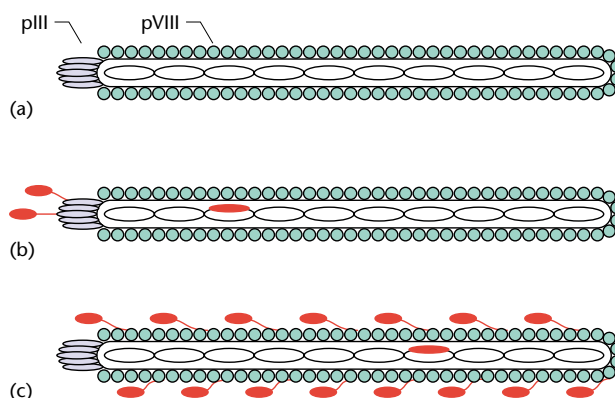


Figure 1 A schematic diagram showing (a) wild type filamentous phage; (b) pIII-based phage display; (c) pVIII-based phage display.

or proteins can therefore be inserted at the N-termini (after the first 3–5 N-terminal residues) of both pIII and pVIII, as long as the signal sequence is preserved. A common strategy for expressing a recombinant fusion coat protein is to insert a DNA coding fragment of interest at the position C-terminal to the signal peptide. Such a gene fusion, when synthesized in the bacterial host together with other coat and accessory proteins of filamentous phage, will be incorporated into the assembled phage. The display vehicle can be either phage or phagemid. Phagemids are plasmids with a phage replication origin, which can be incorporated into a phage particle when the other necessary protein components are available within the cell – for example, by infecting the cells with helper phage. In both formats of display (**Figure 1**), the released phage or phagemid particles expose the recombinant protein on the phage surface. Unlike phage, phagemids usually encode either pIII or pVIII recombinant protein. Thus, the assembled phagemid particle is normally a mosaic, containing both wild-type coat proteins from helper phage and recombinant coat protein from phagemid genome. Display in a phagemid format usually improves phage viability.

By inserting different DNA fragments, a library of phage particles bearing different recombinant coat proteins can be generated. For example, cloning of random oligonucleotides produced a phage library of random peptides, where each phage particle contains only one type of recombinant coat protein, encoded by the corresponding gene fusion physically present inside the same phage particle. Thus, DNA sequencing would allow for deduction of the amino acid sequence of a recombinant coat protein.

Gene III Protein

The pIII, a minor coat protein present only at the tip of the phage particle, mediates infection of bacteria. It contains 406 amino acids and can be divided into two distinct domains. The C-terminal domain interacts with pVIII proteins to maintain attachment to the phage coat, while the N-terminal two-thirds of the protein forms a knob-like structure that mediates the attachment to the F pilus essential for subsequent infection. Peptide insertions at the N-terminus of pIII typically do not interfere with pIII function or infectivity. Genetic manipulation of DNA insertion, such as use of a link peptide sequence (GGGSGGGSGGGGS) to separate the displayed epitope from the rest of the coat protein, allows for these N-terminal epitopes to be displaced from the rest of the phage particle, making them more accessible for binding interactions with a receptor of interest. A random peptide library that displays sequences as pIII fusion proteins maximally allows 4–5 copies of individual random peptide displayed at the tip of each phage particle.

Gene VIII Protein

The major coat protein, pVIII, is a polypeptide of only 50 residues that makes up the body of the phage. The number of pVIII per phage depends on the length of the phage particle, which depends on the size of the genome. Generally, one pVIII protein is required per 2.3 nucleotides, which corresponds to approximately 2800 copies of pVIII in an average phage particle of 6.5-kb genome. As with pIII, random oligonucleotides can be cloned into the pVIII gene and expressed as N-terminal fusion proteins that project away from the surface of the phage. Because pVIII composes the entire phage coat, fusion proteins will be distributed along the length of the phage.

Addition of more than five amino acids to the N-terminus of pVIII often becomes toxic to assembly of the phage capsid, resulting in a significant reduction of phage infectivity. If longer insertions are required, the phagemid system can be used to provide enough wild-type pVIII to make a hybrid phage capsid, containing both wild-type and peptide-fused pVIII proteins which are encoded by helper phage and phagemid DNA. Most pVIII proteins in the coats of viable phage will be wild-type, but many pVIII fusion proteins will also be incorporated. The relative abundance of pVIII fusion proteins incorporated into the phage coat is determined by the amount of wild-type and recombinant pVIII protein synthesized in the cell. Positioning the pVIII gene fusion downstream from an inducible promoter has allowed for regulation of copy number per phage particle.

Before constructing a phage-displayed random peptide library, a decision must be made regarding the use of either pIII or pVIII fusion proteins. Both pIII- and pVIII-displayed libraries are commonly used; the difference resides mainly in the desired affinity of the binding peptides isolated, because there are more copies of displayed peptide per phage in pVIII-based system. The multiple binding peptides per particle permit multivalent interaction between phage and target proteins. Binding conditions and valence can be altered to isolate phage with modest to high binding affinities (K_D s from 500 μ M to 5 pM). Generally, higher valence of displayed peptides reduces the affinity threshold. A pIII-displayed library is a low polyvalent system, as there are a total of 4–5 copies of pIII per phage and all displayed peptides on a given phage are in close proximity to each other, whereas a pVIII-displayed phagemid library is less polyvalent, as the displayed peptides on a given phage particle are relatively isolated from each other (**Figure 1**). However, since each pVIII fusion phage displays many copies of its random peptide, multivalent interactions with the receptor surface may result in avidity effects that allow preferential selection of phage with modest affinity over higher-affinity clones. In order to isolate a small number of high-affinity peptides, a monovalent display system is often more desirable. This can be accomplished by using the phagemid system with

pIII recombinant phage. The helper phage will provide a large excess of wild-type coat protein, while the phagemid encodes the recombinant fusion proteins. The vast majority of the phage population will display limited copies of fusion proteins, thereby reducing the effect of multivalent binding.

Library Construction and Identification of Recombinant Phage of Interest

A phage-displayed random peptide library is constructed by first synthesizing oligonucleotides that encode the random peptides. Because the synthetic chemistry for each base is essentially identical, one can create oligonucleotides encoding completely random or biased amino acid sequences by mixing different combinations of nucleotide triphosphates at different molar ratios. For example, a common strategy to generate random amino acids is to synthesize oligonucleotides with (NNK)_{*n*}, where N represents an equal mixture of all four nucleotides, and K is a 1:1 mixture of guanine (G) and thymine (T). Incorporation of these oligonucleotides in frame with gIII or gVIII will produce a stretch of *n* residues of random amino acids as part of the coat proteins. If a stop codon occurs, the resultant coat protein will terminate prematurely. Consequently, the truncated polypeptide will not be assembled into the phage particle.

Library construction represents one of the key steps because starting with a good library allows for a better chance of isolating peptides of interest. In addition to length and constraint factors, the quality of a random peptide library is often judged by its sequence complexity, i.e. how many independent peptide sequences are present in the library. The current size of a given peptide library is in the range 10⁹–10¹¹. Library complexity is primarily limited by ligation and transformation efficiency. The complexity of 10¹¹ independent clones would in theory allow for a complete representation of all sequences for 8-residue random peptides. Of course, one should bear in mind that sequences toxic to bacterial growth or phage assembly will be underrepresented or completely eliminated from the library.

In theory, peptides of interest can be selected based on their properties using either physical or physiological procedures. The most commonly used approach is to enrich clones with binding affinity to a target of interest by a technique known as panning. The basic steps involve mixing the random peptide-displaying phage pool with the immobilized receptor of interest. The target can be attached to a column matrix or microtitre wells by means of antibody binding or other high-affinity interactions. Recombinant phage displaying peptides that bind to the receptor will be retained on this surface, while nonbinding phage will be removed during a series of wash steps. Bound

phage particles can be then eluted with a buffer (for example, low pH) which disrupts the binding interaction, or via a specific proteolytic cleavage. Because the phage encode the sequence of the displayed peptide, amplification of the eluted phage in bacteria would enrich clones with interacting peptides. The amplified clones can then be used for subsequent rounds of panning and amplification.

After three or four rounds of affinity panning, the binding population will typically contain a small subset of peptides that interact with the target protein, or sometimes with bovine serum albumin (BSA), a control protein used in buffer to reduce nonspecific interactions. These phage clones can be tested for specific binding to the receptor by enzyme-linked immunosorbent assay (ELISA). DNA sequencing of the coding region of individual phage in this population allows for deduction of the amino acid sequence of their displayed peptides. Comparison of the peptide sequences may reveal consensus motifs required for binding to the receptor. Based on the amino acid sequence conservation or degeneracy at a given position, one may predict the role of that residue in the peptide–receptor interaction.

To further improve the desired properties of an isolated peptide, one can employ a mutagenesis screen by constructing a peptide library with degenerate oligonucleotides that encode a large pool of homologous peptides. When this procedure is used to screen for peptides with higher binding affinity, one may perform the panning selection in the presence of synthetic peptide corresponding to the initial sequence. Under these competitive panning conditions, higher-affinity binding peptides may be obtained. This procedure is often referred to ‘affinity maturation’. In addition to the competition panning procedure, one may consider modifying other steps. For example, a reduction of multivalency on the phage capsid would also reduce the effect of avidity; consequently, only peptides of higher affinity would be retained in complex with the target receptor.

Peptide Ligands

Random peptide libraries can be displayed in a variety of ways in order to isolate the most useful peptides. Libraries are generated with the length of the displayed peptides varying from 6 to 30 amino acids. Constrained libraries are designed with cysteine residues that allow for the formation of a disulfide bond. In this format, the peptides are displayed in a more constrained conformation.

Isolation of peptides that bind to a protein of interest has a number of applications. First, many proteins interact with other proteins via a peptide-binding mechanism. Identification of their peptide-binding specificity would allow for better understanding of their interaction with targets (Sparks *et al.*, 1994). For example, SH3 domains

(Src-homology 3) bind proteins containing a polyproline motif and most PDZ domains (PSD-95, Discs large, ZO-1) bind target proteins that contain a stereotypic amino acid consensus (i.e. T/S-X-V-COOH, where X can be any amino acid, and T, S and V represent threonine, serine and valine, respectively). The peptide-binding motif of an unknown protein can be rapidly determined by random peptide library panning. Second, identification of peptide-binding consensus for an 'orphan' protein (one of unknown binding partners) can help identify its interacting proteins, thereby providing information on its function (Stricker *et al.*, 1997). Third, although peptides generally make poor pharmaceuticals, binding peptides identified by phage display can serve as lead molecules, which are used to design pharmacologically active drugs (Wrighton *et al.*, 1996). And finally, random peptide phage display can be used to identify peptides that bind to nonprotein targets. These peptides may be useful for a variety of purposes, such as biosensor.

Phage as Alternatives to Antibodies

Phage display technologies can also be used to generate high-affinity (nM range K_d) antibodies to proteins of interest, essentially bypassing conventional immunization and hybridoma technology (Winter *et al.*, 1994). Antibody isolation is accomplished by panning a library that displays variable regions of immunoglobulin genes. There are several ways to construct a phage-based antibody library, the most common of which is to clone the variable region of both immunoglobulin (Ig) light (V_L) and heavy (V_H) chain genes upstream of the pIII phage coat protein gene. The V_L and V_H genes are amplified from lymphocytes using the polymerase chain reaction (PCR) and connected via a linker region so that, upon expression, the fusion proteins, also known as single-chain antibodies, resemble the antigen-binding region of a traditional antibody. Panning of the library against an antigen of interest allows for isolation of phage displaying single-chain antibody that binds to the antigen. Each antibody-displaying phage particle can be amplified and retained indefinitely, similar to a hybridoma cell line for a monoclonal antibody. Additionally, a process analogous to affinity maturation, termed chain shuffling, can further improve the binding affinity of an isolated phage. This chain shuffling process involves the generation of two libraries, one in which the V_L gene of the best binding phage is held constant and displayed with a library of V_H genes, and another where the best antigen-binding V_H region is displayed with a library of V_L genes. Panning of both libraries and assembling optimal V_H and V_L into one molecule may produce antibodies with significantly higher binding affinity. Antibodies with higher affinity are usually more useful in research and therapeutic applications.

Phage-displayed antibody libraries are particularly useful for the isolation of antibodies that may be difficult to generate by traditional means, such as those retained in the lumen of the endoplasmic reticulum, and surface markers of lymphocytes (e.g. CD₄), tumour cells and red blood cells. Screening of phage-displayed antibody libraries is also a convenient way to obtain human antibodies which may be therapeutically useful.

Epitope Mapping

The basic concept of peptide display was initially prompted by epitope mapping to determine linear amino acid sequences that mediate the interaction between antibody and its binding protein (Smith, 1985). The experimental steps involve screening of a peptide library against antibody in order to determine the epitope specificity. Binding phage may encode peptides sufficient for antibody binding. This approach can be useful for determining the specificity of a particular antibody. Two types of library are commonly used. To determine specificity of a given antibody, one can screen for binding peptides from a random peptide library. Based on the amino acid conservation, a binding consensus may be determined. Alternatively, one can construct a peptide library using fragmented coding sequences of the antibody-binding protein. Screening such a library against the antibody may allow for direct identification of native epitope(s).

Other Applications

Since the development of phage display technology, several variations have been developed to provide other advantageous features complementary to phage display. This includes peptide display on bacteriophage λ (Sternberg and Hoess, 1995), selectively infective phage (SIP) (Spada *et al.*, 1997), and peptides-on-plasmids (Cull *et al.*, 1992) and polysome display (Mattheakis *et al.*, 1994). The key concept shared by all variations is the physical linkage between displayed proteins or peptides and their corresponding DNA coding sequences. Thus, it provides a means of selectively amplifying clones of binding epitopes.

The SIP system is a variation of filamentous phage display technology (Spada *et al.*, 1997). Here, a specific protein–ligand interaction is required to restore the infectivity of the phage. This technique exploits the fact that the N-terminal portion of the pIII coat protein is required for infection; the extreme N-terminus (N1) is involved in bacterial membrane penetration, while the second N-terminal domain (N2) is responsible for binding to the bacterial F pilus (Spada *et al.*, 1997). Phage lacking either domain are unable to infect bacteria. SIP technology separates these domains (N1 or N1-N2) from the C-

terminal anchoring domain of pIII, rendering the phage noninfective. The library to be displayed is fused to the C-terminal portion, while the protein of interest is fused to the N-terminal region of pIII. Interaction of a peptide in the library with the protein of interest brings the N- and C-terminal domains together and restores the infectivity of that phage particle. Although the size limitations of the protein fused to the N-terminal domain have not been thoroughly tested, this approach is reported to be faster than traditional phage display (Spada *et al.*, 1997).

Peptide display on bacteriophage λ is another variation of phage display technology (Sternberg and Hoess, 1995). As with filamentous phage display, peptides displayed on the surface of λ phage are encoded by the DNA inside each particle. Here, foreign proteins are fused to either the N-terminus of the capsid D protein, or to the C-terminus of the tail protein pV. Since the phage λ particle is assembled inside the bacterial cell and released upon cell lysis, epitopes displayed as pV or D protein fusion proteins need not be filtered by membrane targeting. Thus foreign proteins which may not be compatible with the secretion process can be efficiently displayed, making this approach a valuable complement to the filamentous display procedure.

Other peptide display techniques have been developed that do not involve phage particles at all. Peptides-on-plasmids is a random peptide library, where peptides are fused to the C-terminus of the *lac* repressor protein, LacI (Cull *et al.*, 1992). Since LacI has a high affinity for a DNA sequence known as the *lac* operator, genetic linkage of LacI fusion proteins is accomplished by including an operator sequence in the plasmid DNA that encodes the LacI fusion proteins. Thus as each repressor-peptide fusion protein is expressed individually in *Escherichia coli*, it will bind to the operator sequence on the plasmid DNA that encodes it. This type of library is screened by lysis of the bacterial cells containing the peptide-repressor-plasmid complexes, followed by incubation of the crude lysate with an immobilized protein of interest. After washing, binding clones are eluted by the addition of isopropyl- β -D-thiogalactopyranoside (IPTG), which causes LacI to lose affinity for the operator sequence, thus releasing the plasmids that encode binding peptides. These plasmids are amplified in bacteria and the process is repeated. Sequencing of the plasmid DNA in the region of the *lac* gene reveals the amino acid sequence of the binding peptide.

Another type of display technology has been developed which presents peptides on polysomes and does not rely on bacterial cells at all (Mattheakis *et al.*, 1994). Since the size of a cell-based random peptide library is limited to the number of recombinants recovered from the initial transformation step, a completely *in vitro* system has the potential to screen peptide libraries several orders of magnitude larger than those previously described. An additional advantage is that the displayed peptides will not

be subjected to constraints of the host cell nor the ligation efficiency that normally limits the size of library. Polysome display is accomplished by incubating a DNA pool encoding the random peptide library in an *E. coli* transcription/translation system. Polysomes are isolated and incubated with the protein of interest. Those containing nascent peptides that can interact with the protein are retained, while nonbinding polysomes are washed away. The messenger ribonucleic acid (mRNA) from binding complexes is recovered and converted into complementary DNA (cDNA) by reverse transcription. The cDNA pool can then be amplified by PCR to generate templates for the next round of synthesis and selection. Here, genetic linkage is accomplished by means of the association between the ribosome that synthesizes a nascent peptide and the encoding mRNA.

Summary

Phage display and other biologically amplifiable display systems have revolutionized the study of protein-protein interactions. These techniques now make it possible to identify rapidly specific interactions among protein, peptide and chemical moiety. The molecular understanding of novel interactions and binding ligands provides important insights into biological mechanisms in both health and disease.

References

- Cull MG, Miller JF and Schatz PJ (1992) Screening for receptor ligands using large libraries of peptides linked to the C terminus of the *lac* repressor. *Proceedings of the National Academy of Sciences of the USA* **89**: 1865–1869.
- Mattheakis LC, Bhatt RR and Dower WJ (1994) An *in vitro* polysome display system for identifying ligands from very large peptide libraries. *Proceedings of the National Academy of Sciences of the USA* **91**: 9022–9026.
- Smith GP (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**: 1315–1317.
- Spada S, Krebber C and Pluckthun A (1997) Selectively infective phages (SIP). *Journal of Biological Chemistry* **378**: 445–456.
- Sparks AB, Quilliam LA, Thorn JM, Der CJ and Kay BK (1994) Identification and characterization of Src SH3 ligands from phage-displayed random peptide libraries. *Journal of Biological Chemistry* **269**: 23853–23856.
- Sternberg N and Hoess R (1995) Display of peptides and proteins on the surface of bacteriophage. *Proceedings of the National Academy of Sciences of the USA* **92**: 1609–1613.
- Stricker N, Christopherson K, Yi B *et al.* (1997) PDZ domain of neuronal nitric oxide synthase recognizes novel C-terminal peptide sequences. *Nature Biotechnology* **15**: 336–342.
- Winter G, Griffiths AD, Hawkins RE and Hoogenboom HR (1994) Making antibodies by phage display technology. *Annual Review of Immunology* **12**: 433–455.
- Wrighton N, Farrell F, Chang R *et al.* (1996) Small peptides as potent mimetics of the protein hormone erythropoietin. *Science* **273**: 458–463.