

Review Article

Interaction Analysis through Proteomic Phage Display

Gustav N. Sundell and Ylva Ivarsson

Department of Chemistry-BMC, Uppsala University, P.O. Box 576, 751 23 Uppsala, Sweden

Correspondence should be addressed to Ylva Ivarsson; ylva.ivarsson@kemi.uu.se

Received 25 June 2014; Revised 11 August 2014; Accepted 19 August 2014; Published 11 September 2014

Academic Editor: Calvin Yu-Chian Chen

Copyright © 2014 G. N. Sundell and Y. Ivarsson. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Phage display is a powerful technique for profiling specificities of peptide binding domains. The method is suited for the identification of high-affinity ligands with inhibitor potential when using highly diverse combinatorial peptide phage libraries. Such experiments further provide consensus motifs for genome-wide scanning of ligands of potential biological relevance. A complementary but considerably less explored approach is to display expression products of genomic DNA, cDNA, open reading frames (ORFs), or oligonucleotide libraries designed to encode defined regions of a target proteome on phage particles. One of the main applications of such proteomic libraries has been the elucidation of antibody epitopes. This review is focused on the use of proteomic phage display to uncover protein-protein interactions of potential relevance for cellular function. The method is particularly suited for the discovery of interactions between peptide binding domains and their targets. We discuss the largely unexplored potential of this method in the discovery of domain-motif interactions of potential biological relevance.

1. Introduction

The human interactome is estimated to contain about 130,000 binary protein-protein interactions (PPIs), of which the majority remains to be discovered [1]. PPIs are crucial for cellular function and dysfunction and large efforts are therefore invested in their identification and in constructing PPI based networks [2]. Different high-throughput methods render complementary data. For example, affinity purification coupled to mass spectrometry (AP-MS) [3, 4] and luminescence-based mammalian interactome mapping (LUMIER) [5] provide information on complexes, and yeast-two-hybrid (Y2H) experiments give insights into binary PPIs [1], as summarized in Table 1. Despite the significant advances being made the last decade, the human interactome is still largely uncharted and the accumulated knowledge is biased towards well-studied proteins [1, 6].

Particularly elusive to high-throughput methods are the interactions between peptide binding domains and their target motifs, which are typically less than ten residues in length [7, 8]. The peptide motifs are typically located in regions of intrinsic disorder, which can be found in about 35% of the human proteins [9]. Currently, there are more than 2,400 instances reported in the eukaryotic linear motif (ELM)

resource for functional sites in proteins [10], including binding motifs and posttranslational modification sites. This, however, covers only a fraction of the motifs expected to be present in the human proteome [8].

Among the most abundant peptide binding domains in the human proteome are the PSD-95/Discs-large/ZO-1 (PDZ) domains that typically interact with C-terminal sequences of target proteins [11]. Other domains, such as the Src Homology 2 (SH2), bind to phosphorylated target motifs [12]. Domain-motif interactions tend to be of rather low affinities and hence are easily lost in methods such as AP-MS. Although difficult to capture experimentally, transient protein-peptide interactions are crucial for cell function and may be perturbed by disease-causing genetic variations or by viral interferences [13, 14].

Phage display is a powerful approach for establishing binding preferences of peptide binding domains and in extension to discover novel motifs. In combinatorial peptide phage display, highly diverse libraries are used to identify high-affinity ligands with potential to serve as inhibitor [15]. Consensus motifs are derived based on the retained sequences and can be used for predictions of potential ligands in a target proteome [16]. These predictions, however, are not always accurate, which can lead to tedious experimental validations

TABLE 1: Summary of high-throughput methods for identification of PPIs, types of interactions identified, and major advantages and disadvantages of the respective method.

Method	Type of interaction	Advantage	Disadvantage
AP-MS	Binary and complexes	Physiological	Bias towards stable interactions, limited to specific condition (e.g., cell type)
LUMIER	Binary and complexes	Physiological	Bias towards stable interactions
Y2H	Binary	Low-tech	Bias towards stable interactions, bias towards soluble proteins that can translocate to the nucleus
Combinatorial peptide phage display	Binary	Large library size (up to 10^{10}) Identification of consensus motifs	Need for bioinformatics, limited to natural amino acids, limited to protein-peptide interactions
Proteomic phage display	Binary	Identification of target proteins and consensus motif	Limited to natural amino acids

of putative targets. Luck and Travé demonstrated that predictions of human PDZ domain ligands based on results of combinatorial phage display may be hampered due to a bias towards overly hydrophobic (i.e., Trp containing) peptides [17].

A promising strategy to discover novel protein-motif interactions is to reduce the search space to comprise only sequences of a target proteome. In such proteomic phage display, expression products from genomic DNA, cDNA, open reading frames (ORFs), or from designed synthetic oligonucleotides are displayed on phage particles (Figure 1). Proteomic phage display has been used for the identification of allergens [18], antibody epitopes, tumor polypeptides producing immune response [19], and PPIs as well as for the identification of proteins binding to phospholipids and small chemical compounds [20, 21]. In this review, we survey the features, the development, and the applications of various phage display systems used for proteomic phage display, with a particular focus on the elucidation of cellular PPIs. For extensive reviews on cDNA/ORF display for antibody epitope mapping of antigen and pathogen research we refer the readers to dedicated reviews [22–26].

2. Phage Display Systems Used for Proteomic Phage Display

Phages that have been used for proteomic phage display include the filamentous M13 phage, the lytic T7 phage, and the temperate phage λ . The main advantage of the M13 phage display system is the ease of its manipulation and handling as detailed in the following section. The main drawback of the M13 phage is that the displayed proteins are secreted through the periplasmic space of the *Escherichia coli* membrane (Figure 2(a)), which may confer problems with unpaired cysteine residues that form undesired disulfide bridges and thereby compromise the successful display of designed sequences [27]. The T7 phage and the λ phage do not suffer from this issue, as they are not passing through the membrane but lysing the host cells (Figure 2(b)).

2.1. M13 Filamentous Phage. The M13 filamentous bacteriophage was the first developed phage display system [28].

The M13 phage consists of a circular single stranded DNA (ssDNA) that is covered by five different coat proteins (pIII, pVI, pVII, pVIII, and pIX). The 2,700 copies of the major coat protein pVIII cover the length of the phage [29]. The minor coat proteins pVII and pIX cover one end of the phage particle and pIII and pVI the other end (Figure 2(a)). The minor coat protein pIII is crucial for infection as it initiates the interaction with the F-pilus and TolA receptor [30]. For details on the structure and assembly of filamentous bacteriophages we refer the reader to an extensive review on the topic [31]. Typically, phage propagation is uncoupled from expression and display of desired peptide on the phage particle. This is accomplished through hybrid systems where a phagemid is used for library construction and helper phage is added to provide the information needed for assembly of the phage particle [32].

The M13 phage is a highly versatile system as distinct coat proteins can be used for N- or C-terminal display and for monovalent or multivalent display, respectively [33]. Commonly, the pIII protein is used for low valency display (one to five copies per phage) and the pVIII for high valency display, with up to 1,000 copies per phage in evolved hybrid systems [23]. Highly diverse M13 phage libraries (up to 10^{10}) can be constructed due to the fact that M13 has a circular ssDNA. For a typical library construction, an oligonucleotide library is designed complementary to the ssDNA with flanking regions corresponding to the phagemid vector. The oligonucleotides are then annealed to the vector and the complementary strand is synthesized and ligated together to form a circular, double stranded DNA vector, which is then electroporated into *E. coli* [34].

2.2. T7 Phage. T7 phage is an icosahedral virus of the Podoviridae family and has a linear double stranded (ds) DNA genome. In contrast to M13, T7 is not secreted but released from the host cell through lysis (Figure 2(b)). The T7 phage starts to reproduce immediately upon infection, which is continuous until the point of cell lysis. The major capsid protein (gp10) is encoded by gene 10 and makes up about 90% of the icosahedron capsid. This gene yields two products, 10A and 10B, in a nine-to-one ratio. The minor protein 10B results from a frame shift in the end of the gene that makes

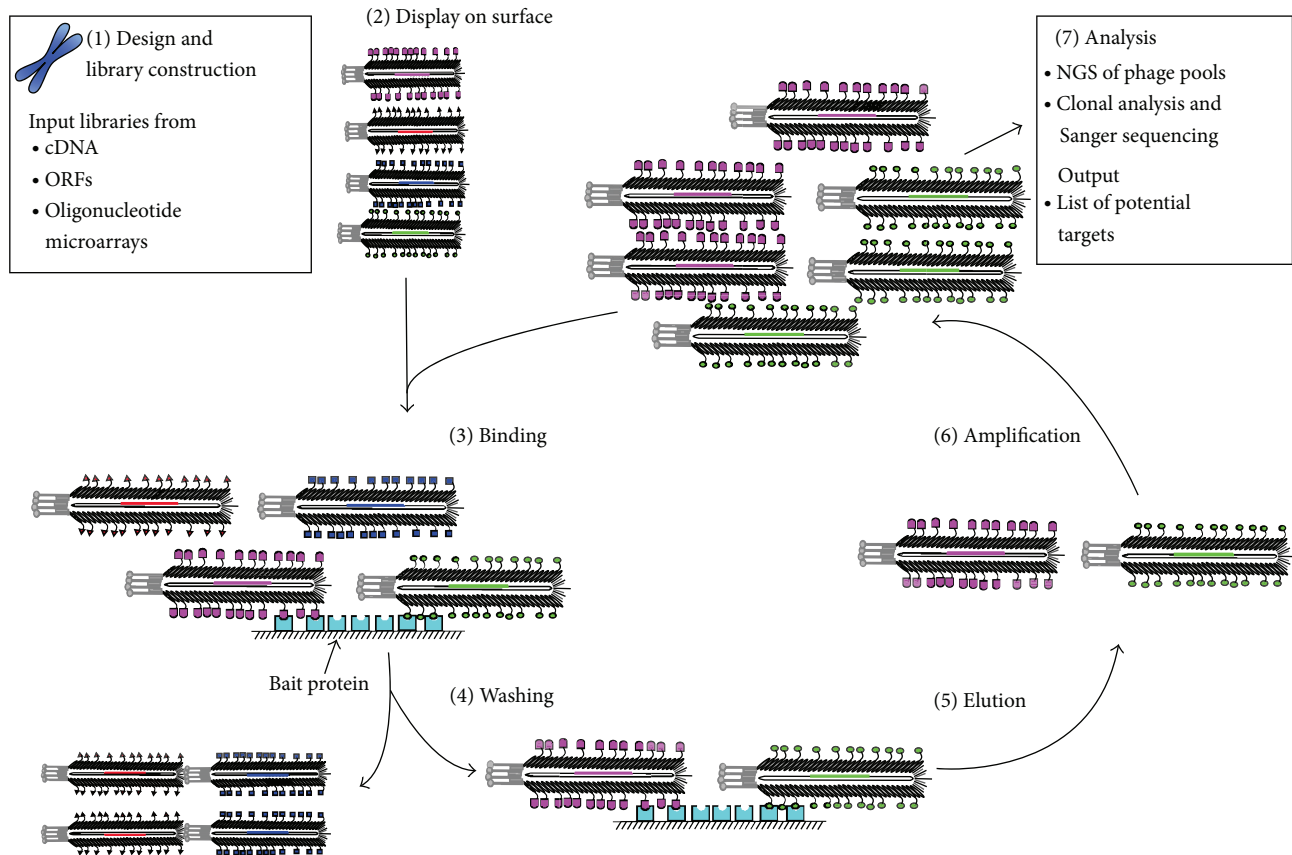


FIGURE 1: Schematic representation of proteomic phage display using the M13 phage. Input phage display libraries are constructed from cDNA, ORFs, or oligonucleotide arrays designed from a proteome of interest (1). Peptides are displayed on pVIII (2). Bait proteins are immobilized on a solid surface and incubated with the naïve input phage library (3). Binding of phage occurs through interactions between displayed peptides and bait proteins, but nonspecific interactions cause noise in the selection (not shown). Unbound phage is washed away (4) and bound phage is eluted through acidic or basic conditions or by the addition of actively growing host bacteria (5). Eluted phage is amplified (6) and used for repeated (typically 3–5) cycles of selection, which is necessary to amplify specifically bound phage over nonspecific binders. Sanger sequencing of confirmed binders and/or NGS of the retained phage pools provides lists of binders from the target proteome (7).

the capsid protein 52 residues longer [35]. Fusion proteins are displayed on protein 10B C-terminally of the 52 extra residues. Depending on the system used, up to 1,200 amino acid inserts can be displayed at low valency (5–15 copies per virion) or shorter inserts (up to 50 amino acids) at higher valency (up to 415 copies) [36]. The linear genome makes it more challenging to construct T7 phage libraries as compared to M13 libraries. Library construction includes two-step ligations and the *in vitro* packing of DNA into the phage, which in the T7 select system (Novagen) is accomplished by the addition of DNA to commercially available packaging extract. The packaging extract is sensitive to work with and rather costly if larger libraries are prepared [33].

2.3. Lambda Phage. The temperate λ phage has an icosahedral head. The main structure of the shell is built from the major coat protein gpE (415 copies) and is stabilized by the capsid protein gpD (402–420 copies) [37]. The head is linked to a flexible helical tail constructed by disks of the major tail protein gpV. Its linear dsDNA is packed in the bacteriophage

head. The DNA is injected into the host bacteria and is stably integrated into the host chromosome during the lysogenic state. When triggered correctly, the λ phage starts a lytic cycle [38].

Both the tail protein gpV and the head protein gpD have been used for phage display. Initially, the λ foo vector was constructed for the C-terminal display on gpV, with a low display level that made it suitable for capturing high-affinity interactions [39]. Later, systems were developed for the display of peptides N-terminally or C-terminally to the major coat protein gpD [40–42]. Libraries with diversities in the range of 10^7 – 10^8 are constructed using commercially available *in vitro* packaging systems.

3. Proteomic Phage Display

Over the years, different approaches towards proteomic phage display have been taken, from cDNA and ORF display to the display of the expression products of highly defined synthetic oligonucleotide libraries, as detailed in the following section.

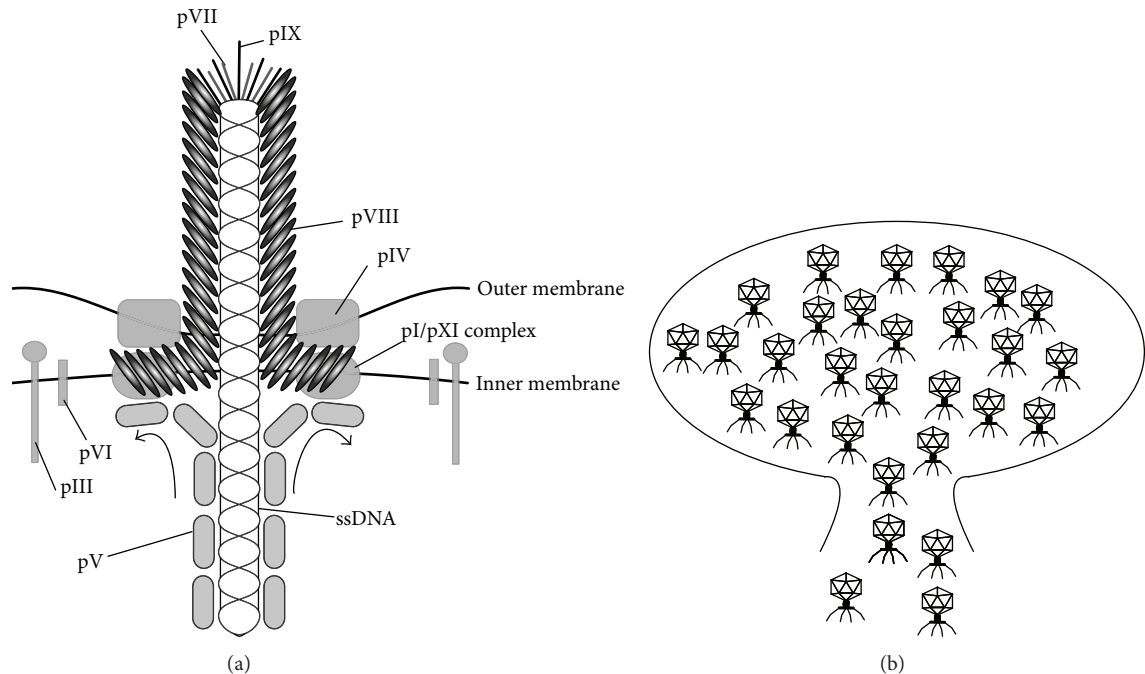


FIGURE 2: Schematic representation of the assembly and excretion of the M13 filamentous phages and the exit through cell lysis of the lytic T7 phage. (a) The filamentous M13 phage is assembled at the cell membrane of the bacterial host. In the host cell, the ssDNA is protected by association with protein pV, which detaches at the membrane upon assembly. At the start of the assembly, membrane-associated pVII/pIX bind a specific DNA packing signal. Membrane-bound protein pVIII binds to the DNA and is transferred across the membranes. The transport is facilitated by a complex of pI and pXI situated in the inner membrane and protein pIV that makes a pore through the outer membrane of the bacteria for the phage to pass through. As a final step, pVI and pIII that span the inner membrane are assembled on the phage. The figure was created based on [31]. (b) The lytic T7 phage, schematically shown with its typical icosahedral head, is assembled in the cytosol of the host cell. It is multiplied to such extent that the host cell finally bursts and the phage is released to the surrounding.

3.1. cDNA/ORF Display. In cDNA display, a gene, a cDNA, or a complete genome is displayed on phage particles. Theoretically, this is a straightforward technique. However, it suffers from difficulties in obtaining high-quality libraries [43]. This is a consequence of the transcriptional stop codons at the 3'-end of coding regions, the polyA tail of mRNA, and the often nondirectional cloning. The fraction of clones expressing peptides in frame in a naïve cDNA library may be as low as 6%. Additionally, phage with truncated constructs tends to outgrow clones with correctly displayed sequences [44, 45]. The quality can be improved by using ORF enriched DNA collections for library construction [43, 46]. Library quality can further be improved by fragmentation of the DNA by, for example, treatment with Deoxyribonuclease I or by sonication before cloning [26, 47]. Despite the quality issues, a variety of libraries based on human or pathogen cDNA/ORFs have been derived and used for PPI screening.

3.1.1. cDNA/ORF Display Using the M13 Filamentous Phage. A number of studies have employed the M13 filamentous phage system for cDNA display. In most cases, the expression products are displayed on the pIII protein either indirectly through the Jun-Fos system or directly [48]. There are also reports of N-terminal multivalent display on pVIII [49] and monovalent C-terminal display on pVI [50]. When displaying inserts N-terminally, a main limitation is that the inserts have

to be in the same reading frame as the pIII or pVIII proteins and that there can be no in frame stop codons. A way to enrich for ORFs and the correct presentation of encoded sequences and thereby improve the quality is the so-called Hyperphage [51]. In this system, the helper phage has a truncated g3 so that the phagemid pIII fusion is the only source of pIII, as originally described by Kristensen and Winter [52]. This strategy has been used for the successful identification of immunogenic polypeptides of *Mycoplasma hyopneumoniae* [53]. Recently, a novel trypsin-sensitive helper phage was derived for a similar purpose [54].

Indirect Display on pIII: The Jun-Fos System. A system for indirect cDNA display on pIII termed pJun-Fos was engineered in 1993. This system takes advantage of the strong association between the Leucine Zipper Jun and Fos [55]. The pIII-fused Jun and Fos-linked cDNA expression product is assembled in the *E. coli* periplasm, which leads to the indirect display of the functional expression products of cDNA on pIII. The complex is stabilized by disulfide bridges between cysteines engineered at the N- and C-termini of Jun and Fos. In the original publication, enzymatically active alkaline phosphatase was displayed on pIII and the authors discussed the potential of the system as a tool in PPI screening. The Jun-Fos system has since then been a popular cDNA display system for the discovery of antibody epitopes [56] as reviewed elsewhere [48].

A Jun-Fos system modified to ensure cloning in all three reading frames was used to identify host-pathogen protein-protein interactions between the ribonucleoproteins of influenza virus and the expression products of a human cDNA library (inserts >750 bp) [57]. In this study, the authors pinpointed a direct interaction between the A domain of human high mobility group box proteins and the viral bait protein.

Direct Display on pIII. Expression products of cDNA/ORF have also been displayed directly on pIII. In an early study, the plasminogen-activator inhibitor 1 was fragmented into 50–200 bp and cloned into the M13 gpIII phagemid vector [58]. This library was used for epitope mapping of a monoclonal antibody raised against this protein. For PPI screening, Hertveldt et al. constructed a phage library by fusing genomic *S. cerevisiae* DNA (100–1,500 bp) to gpIII lacking the N1 domain [59]. From panning of the yeast cDNA library against GAL80, fragments of the known binder GAL4 and three other ligands of potential physiological relevance were retrieved, thereby demonstrating that the system can be used for identification of biologically relevant targets. Around the same time, Yano et al. constructed a fragmented genomic *E. coli* pIII library and identified binders to alkaline phosphatase [60].

Two other studies demonstrate that ORF enriched cDNA display on pIII can be used to identify targets of potential biological relevance. In the first study, the interactomes of the high mobility group A proteins HMG1 and HMG2 were elucidated using an ORF enriched murine cDNA M13 pIII library displaying 200–500 base pair fragments [61]. For these nuclear chromatin factors, four targets were identified, namely, TBP associated factor 3b and chromatin assembly factor I, subunit A, and two previously uncharacterized proteins. For the first two proteins, interactions were confirmed between the full-length proteins through GST-pull down assays and coaffinity purification of overexpressed proteins in HEK293T cells [61].

In the second study, an ORF enriched and fragmented cDNA library displayed on pIII was used for interactome mapping of transglutaminase 2 (TG2) [62]. Through next-generation sequencing (NGS) of selected phage pools a list of potential targets was retrieved. The most frequently occurring ligands interactions were validated through protein complementarity assays with 80% success rate, thus demonstrating the power of the combination of ORF enriched cDNA display and NSG in interactome mapping.

Posttranslational Modifications and ORF Display on pIII. PPIs are often controlled by posttranslational modifications, with the most common modifications in eukaryotic proteomes being phosphorylation of Ser/Thr/Tyr residues [63]. These modifications can create or abrogate binding sites or modulate function by more indirect means. A few attempts have been made towards investigating PPIs relying on posttranslational modifications through proteomic phage display. In particular, Cochrane and coworkers used the fyn tyrosine kinase to *in vitro* phosphorylate a fragmented leukocyte cDNA library (10^8) displayed on pIII. The phosphorylated

library was used in selections against the phosphotyrosine binding tandem of SH2 domains of SHP-2 [64]. Nonspecific binding clones were removed before *in vitro* phosphorylation and selection using SHP-2 Sepharose. Through competitive ELISA experiments using phosphorylated phage and synthetic peptides, double phosphorylated PECAM-1 was identified and confirmed as a SHP-2 ligand. It thus appears possible to identify natural interactions relying on posttranslational modification through cDNA phage display. However, given the lack of follow-up studies it does not seem like a feasible way to go for high-throughput analysis of PPIs depending on posttranslational modification. Other attempts to tackle posttranslational modification involved the system for the production and enrichment of phage displaying N-glycoproteins [65].

cDNA Display on pVI. A limited set of studies has employed C-terminal cDNA display on pVI, thus circumventing issues related to the presence of premature stop codons. Using a pVI cDNA library of the hookworm *Ancylostoma caninum* ligands were identified for two serine proteases [50]. A few years later, a rat liver cDNA library fused to pVI was used for the identification of peroxisomal proteins by panning the library against antibodies raised against peroxisomal subfractions [66]. In another study, a pVI cDNA library from the colorectal cancer cell line HT-29 was used to identify a panel of candidate tumor antigens [67]. Other studies have reported the discovery of autoantigens for diseases such as multiple sclerosis [68] and rheumatoid arthritis [69]. However, at this stage there are no studies that have applied pVI cDNA display for the explicit purpose of interactome analysis. The monovalent display on pVI makes it less suited for capturing low-affinity interactions.

3.1.2. T7 Phage Display. T7 phage display has become a popular system for cDNA/ORF display, starting from the identification of RNA binding proteins from cDNA displayed on the C-terminus of the capsid protein 10B [70]. It has typically been used for antigen discovery [71]. For example, T7 cDNA display of sea snake venom gland mRNA identified rabbit anti-sea snake venom IgGs as well as new toxins [72]. T7 cDNA display has also been used to explore interactions between parasite proteins and host enterocytes [73, 74].

High-quality ORF T7 display libraries have been used for interactome analysis. In particular, Caberoy and coworkers created a library by combining dual phage display with specific elution of bound phage by protease cleavage [75]. In this system a biotin tag is expressed C-terminally of the inserts and thus is only present when the inserts are in frame. The tag is biotinylated by the *E. coli* BirA enzyme, which enables the selection of ORF clones using immobilized streptavidin. Bound phage is eluted by cleavage with 3C protease. Following this approach, novel tubby binding proteins were identified and then validated through complementary approaches. Of 14 potential target proteins tested, 10 were confirmed as ligands by Y2H and/or pull down assays [75]. The same group used their T7 high-quality ORF library to identify tubby and tubby-like protein 1 as eat-me signals stimulating phagocytosis [76] as well as substrates for the protease calpain 2 [77].

A final example is provided by a study on the suppressor of cytokine signaling 3 (SOC3) [78]. A potential ligand of SOC3, an 11-mer C-terminal peptide of the very long chain acyl-CoA dehydrogenase (VLAD), was found through selections against a human liver cDNA T7 phage library. The interaction was confirmed *in vitro* and in cell-based experiments and was further validated in animal experiments. Based on the results, the authors proposed that SOC3 is an important factor for lipid metabolism.

3.1.3. Phage λ cDNA/ORF Display. Phage λ cDNA/ORF display has found use in antigen discovery, as reviewed elsewhere [79]. Already in 1997, the λ foo system was used for epitope mapping of human galectin [80]. In this study, a library was constructed from fragmented cDNA of galectin-3 and screened against immobilized monoclonal antibodies, leading to the identification of two distinct epitopes of nine and eleven amino-acid residues. This method was shown to outperform a random peptide phage library. Other studies report on epitope mapping of monoclonal and polyclonal antibodies with cDNA phage λ libraries from human brain and mouse embryo [41, 81, 82]. However, to our knowledge there are at this stage no papers on interactome analysis using phage λ .

3.2. Proteomic Peptide Phage Display Libraries from Oligonucleotide Array. Recently, the advances in oligonucleotide microarray synthesis [83] in combination with bioinformatics and NGS have opened new avenues for the construction of highly defined phage libraries. The pioneering study published in 2011 by Larman and coworkers reported the creation of a T7 library displaying 36-mer peptides representing the complete human proteome, with seven amino acids overlaps [84]. With this library, the authors developed a phage immunoprecipitation sequencing platform for the discovery of autoantibodies. They also demonstrated a more general use for interactome mapping by identifying targets for thering protein A2.

In a recent study, a previously engineered pVIII phagemid for multivalent C-terminal display [85] was used to create two distinct proteomic peptide phage display (ProP-PD) libraries. The first library was designed to contain all human C-terminal 7-mer peptides whereas the second library contained all C-termini of known viral proteins. After confirmation of composition and coverage of the libraries through NGS they were used in selection against nine PDZ domains of densin-180, DLG1, erbin, and scribble. Phage pools retained after different selection rounds were analyzed through NGS, which provided detailed information on the progress of the selections. Between two and thirty ligands were obtained for each PDZ domain after the fifth round of selection. Of these, more than 50% of the ligands retained for DLG1, densin-180, and erbin were previously known targets. In contrast, only 13% of the scribble ligands were known since previously. Interactions between full-length scribble and the novel ligands plakophilin-4, mitogen-activated protein kinase 12, and guanylate cyclase soluble subunit alpha-2 were confirmed through colocalizations and coimmunoprecipitations, suggesting that ProP-PD identified biologically relevant targets and that the approach can be used to complement

PPI networks. The ligands retrieved from the selections against the library designed from virus proteins were mostly established biologically relevant ligands, thus demonstrating that the approach can efficiently identify host-pathogen PPIs of biological relevance. Taken together, the proteomic peptide phage display appears to be a highly useful tool for proteome wide screening of domain-motif interactions.

4. Concluding Remarks and Further Perspectives

Various systems for proteomic phage display have been evolved over the last 20 years, with different approaches taken to improve the quality of the displayed sequences. The preferred systems have been the filamentous M13 and the lytic T7 system. The displayed regions range from 7 to 1,500 bps, thus allowing the proteomic identification of peptide ligands as well as interactions involving folded domains. Although most studies have focused on mapping antibody epitopes, it appears as if cDNA/ORF phage display has the potential to successfully identify PPIs of putative biological relevance. This is evident from the validation range of 50–80%, which is considerably better than for techniques such as Y2H [86]. However, cDNA/ORF phage display has had limited use as a method for interactome analysis. The main issue of cDNA/ORF phage display is the lack of control over the displayed sequences, which affects library quality and likely results in the display of a high percentage of unfolded/misfolded proteins and of stretches that are typically inside of folded proteins and not normally available for binding. In most cases, there is a lack of information on the library quality and the coverage of the target genome. Despite the advent of NGS we did not find any publication reporting on the complete sequencing of a cDNA/ORF library, which would provide valuable insights into the quality of the libraries and a better understanding of the interaction space covered during the experiments.

Proteomic peptide phage libraries, created using a combination of bioinformatics and synthetic oligonucleotide libraries, and analyzed through NGS, offer the advantage of full control of displayed regions [87]. At this stage, the cost of highly diverse high-quality oligonucleotide libraries is still rather high. However, given the rapid advances in large-scale *de novo* DNA synthesis [83] we foresee that the cost will go down and that this approach will become increasingly popular for proteomic screening of domain-motif interactions. This will be particularly feasible as phage display can be scaled to hundreds of proteins in parallel [88] and can be paired with NGS of the naïve phage libraries [89] as well as the selected pools, thus providing comprehensive information on the library composition as well as on the retained targets.

By performing proteomic phage display in parallel with other high-throughput methods such as AP-MS or Y2H it is possible to enrich PPI networks with additional interactions and insights on the domain-motif level. Such attempts have previously been made using combinatorial peptide phage display, with an excellent example provided by the Tong et al. study that elucidated the yeast SH3 interactome [90]. More recently, the binding specificities of the worm *Caenorhabditis*

elegans SH3 domains were elucidated via high-throughput peptide phage display. The results were combined with the SH3 interactome that was mapped through Y2H experiment [91]. The use of proteomic libraries rather than combinatorial phage libraries for this kind of analysis will obviate the need for predictions as it directly identifies the target protein based on the selected ligands. ProP-PD will be particularly useful in providing unbiased information on domain-motif interactions. This will give novel insights into the function of unexplored motifs in the human proteome. As these experiments can be performed in high-throughput the limiting factor for elucidating domain-motif interactions will be the access to recombinant proteins and the downstream cell biological validations.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by grants from the Swedish Research Council and the Åke Wiberg Foundation (Sweden).

References

- [1] K. Venkatesan, J. F. Rual, A. Vazquez et al., "An empirical framework for binary interactome mapping," *Nature Methods*, vol. 6, no. 1, pp. 83–90, 2009.
- [2] M. Vidal, M. E. Cusick, and A.-L. Barabási, "Interactome networks and human disease," *Cell*, vol. 144, no. 6, pp. 986–998, 2011.
- [3] A.-C. Gavin, K. Maeda, and S. Kühner, "Recent advances in charting protein-protein interaction: Mass spectrometry-based approaches," *Current Opinion in Biotechnology*, vol. 22, no. 1, pp. 42–49, 2011.
- [4] W. H. Dunham, M. Mullin, and A.-C. Gingras, "Affinity-purification coupled to mass spectrometry: basic principles and strategies," *Proteomics*, vol. 12, no. 10, pp. 1576–1590, 2012.
- [5] M. Barrios-Rodiles, K. R. Brown, B. Ozdamar et al., "High-throughput mapping of a dynamic signaling network in mammalian cells," *Science*, vol. 307, no. 5715, pp. 1621–1625, 2005.
- [6] J.-F. Rual, K. Venkatesan, T. Hao et al., "Towards a proteome-scale map of the human protein-protein interaction network," *Nature*, vol. 437, no. 7062, pp. 1173–1178, 2005.
- [7] B. A. Liu, B. W. Engelmann, and P. D. Nash, "High-throughput analysis of peptide-binding modules," *Proteomics*, vol. 12, no. 10, pp. 1527–1546, 2012.
- [8] P. Tompa, N. E. Davey, T. J. Gibson, and M. M. Babu, "A million peptide motifs for the molecular biologist," *Molecular Cell*, vol. 55, no. 2, pp. 161–169, 2014.
- [9] J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton, and D. T. Jones, "Prediction and functional analysis of native disorder in proteins from the three kingdoms of life," *Journal of Molecular Biology*, vol. 337, no. 3, pp. 635–645, 2004.
- [10] H. Dinkel, K. van Roey, and S. Sushama Michael, "The eukaryotic linear motif resource ELM: 10 years and counting," *Nucleic Acids Research*, vol. 42, pp. D259–D266, 2014.
- [11] Y. Ivarsson, "Plasticity of PDZ domains in ligand recognition and signaling," *FEBS Letters*, vol. 586, no. 17, pp. 2638–2647, 2012.
- [12] J. Jin and T. Pawson, "Modular evolution of phosphorylation-based signalling systems," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, no. 1602, pp. 2540–2555, 2012.
- [13] N. E. Davey, G. Travé, and T. J. Gibson, "How viruses hijack cell regulation," *Trends in Biochemical Sciences*, vol. 36, no. 3, pp. 159–169, 2011.
- [14] B. Uyar, R. J. Weatheritt, H. Dinkel, N. E. Davey, and T. J. Gibson, "Proteome-wide analysis of human disease mutations in short linear motifs: neglected players in cancer?" *Molecular BioSystems*, 2014.
- [15] Y. Zhang, B. A. Appleton, C. Wiesmann et al., "Inhibition of Wnt signaling by Dishevelled PDZ peptides," *Nature Chemical Biology*, vol. 5, no. 4, pp. 217–219, 2009.
- [16] J. Teyra, S. S. Sidhu, and P. M. Kim, "Elucidation of the binding preferences of peptide recognition modules: SH3 and PDZ domains," *FEBS Letters*, vol. 586, no. 17, pp. 2631–2637, 2012.
- [17] K. Luck and G. Travé, "Phage display can select over-hydrophobic sequences that may impair prediction of natural domain-peptide interactions," *Bioinformatics*, vol. 27, no. 7, pp. 899–902, 2011.
- [18] C. Rhyner, M. Weichel, S. Flückiger, S. Hemmann, T. Kleber-Janke, and R. Cramer, "Cloning allergens via phage display," *Methods*, vol. 32, no. 3, pp. 212–218, 2004.
- [19] A. Fosså, L. Alsøe, R. Cramer, S. Funderud, G. Gaudernack, and E. B. Smeland, "Serological cloning of cancer/testis antigens expressed in prostate cancer using cDNA phage surface display," *Cancer Immunology, Immunotherapy*, vol. 53, no. 5, pp. 431–438, 2004.
- [20] B. van Dorst, J. Mehta, E. Rouah-Martin et al., "cDNA phage display as a novel tool to screen for cellular targets of chemical compounds," *Toxicology in Vitro*, vol. 24, no. 5, pp. 1435–1440, 2010.
- [21] W. Guo, S. Liu, J. Peng et al., "Examining the interactome of huperzine A by magnetic biopanning," *PLoS ONE*, vol. 7, no. 5, Article ID e37098, 2012.
- [22] L. J. Matthews, R. Davis, and G. P. Smith, "Immunogenically fit subunit vaccine components via epitope discovery from natural peptide libraries," *Journal of Immunology*, vol. 169, no. 2, pp. 837–846, 2002.
- [23] G. A. Weiss and S. S. Sidhu, "Design evolution of artificial M13 coat proteins," *Journal of Molecular Biology*, vol. 300, no. 1, pp. 213–219, 2000.
- [24] Y. Weiss-Ottolenghi and J. M. Gershoni, "Profiling the IgOme: meeting the challenge," *FEBS Letters*, vol. 588, no. 2, pp. 318–325, 2014.
- [25] L. M. Mullen, S. P. Nair, J. M. Ward, A. N. Rycroft, and B. Henderson, "Phage display in the study of infectious diseases," *Trends in Microbiology*, vol. 14, no. 3, pp. 141–147, 2006.
- [26] J. Kugler, J. Zantow, T. Meyer et al., "Oligopeptide m13 phage display in pathogen research," *Viruses*, vol. 5, no. 10, pp. 2531–2545, 2013.
- [27] L. R. H. Krumpke, A. J. Atkinson, G. W. Smythers et al., "T7 lytic phage-displayed peptide libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide libraries," *Proteomics*, vol. 6, no. 15, pp. 4210–4222, 2006.
- [28] G. P. Smith, "Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface," *Science*, vol. 228, no. 4705, pp. 1315–1317, 1985.

- [29] P. Malik, T. D. Terry, L. R. Gowda et al., "Role of capsid structure and membrane protein processing in determining the size and copy number of peptides displayed on the major coat protein of filamentous bacteriophage," *Journal of Molecular Biology*, vol. 260, no. 1, pp. 9–21, 1996.
- [30] J. Lubkowski, F. Hennecke, A. Plückthun, and A. Wlodawer, "Filamentous phage infection: crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA," *Structure*, vol. 7, no. 6, pp. 711–722, 1999.
- [31] D. A. Marvin, M. F. Symmons, and S. K. Straus, "Structure and assembly of filamentous bacteriophages," *Progress in Biophysics and Molecular Biology*, vol. 114, no. 2, pp. 80–122, 2014.
- [32] J. Vieira and J. Messing, "Production of single-stranded plasmid DNA," *Methods in Enzymology*, vol. 153, pp. 3–11, 1987.
- [33] K. Fukunaga and M. Taki, "Practical tips for construction of custom peptide libraries and affinity selection by using commercially available phage display cloning systems," *Journal of Nucleic Acids*, vol. 2012, Article ID 295719, 9 pages, 2012.
- [34] S. Rajan and S. S. Sidhu, "Simplified synthetic antibody libraries," *Methods in Enzymology*, vol. 502, pp. 3–23, 2012.
- [35] B. G. Condrón, J. F. Atkins, and R. F. Gesteland, "Frameshifting in gene 10 of bacteriophage T7," *Journal of Bacteriology*, vol. 173, no. 21, pp. 6998–7003, 1991.
- [36] S. C. Sharma, A. Memic, C. N. Rupasinghe, A.-C. E. Duc, and M. R. Spaller, "T7 phage display as a method of peptide ligand discovery for PDZ domain proteins," *Biopolymers*, vol. 92, no. 3, pp. 183–193, 2009.
- [37] F. Yang, P. Forrer, Z. Dauter et al., "Novel fold and capsid-binding properties of the λ -phage display platform protein gpD," *Nature Structural Biology*, vol. 7, no. 3, pp. 230–237, 2000.
- [38] J. Nicastro, K. Sheldon, and R. A. Slavcev, "Bacteriophage lambda display systems: developments and applications," *Applied Microbiology and Biotechnology*, vol. 98, no. 7, pp. 2853–2866, 2014.
- [39] I. N. Maruyama, H. I. Maruyama, and S. Brenner, " λ foo: a λ phage vector for the expression of foreign proteins," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 17, pp. 8273–8277, 1994.
- [40] N. Sternberg and R. H. Hoess, "Display of peptides and proteins on the surface of bacteriophage λ ," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 5, pp. 1609–1613, 1995.
- [41] E. Santi, S. Capone, C. Mennuni et al., "Bacteriophage lambda display of complex cDNA libraries: a new approach to functional genomics," *Journal of Molecular Biology*, vol. 296, no. 2, pp. 497–508, 2000.
- [42] Y. G. Mikawa, I. N. Maruyama, and S. Brenner, "Surface display of proteins on bacteriophage λ heads," *Journal of Molecular Biology*, vol. 262, no. 1, pp. 21–30, 1996.
- [43] W. Li, "ORF phage display to identify cellular proteins with different functions," *Methods*, vol. 58, no. 1, pp. 2–9, 2012.
- [44] Z. Kalniņa, K. Siliņa, I. Meistere et al., "Evaluation of T7 and lambda phage display systems for survey of autoantibody profiles in cancer patients," *Journal of Immunological Methods*, vol. 334, no. 1–2, pp. 37–50, 2008.
- [45] H.-S. Lin, H. S. Talwar, A. L. Tarca et al., "Autoantibody approach for serum-based detection of head and neck cancer," *Cancer Epidemiology Biomarkers and Prevention*, vol. 16, no. 11, pp. 2396–2405, 2007.
- [46] P. Zacchi, D. Sblattero, F. Florian, R. Marzari, and A. R. M. Bradbury, "Selecting open reading frames from DNA," *Genome Research*, vol. 13, no. 5, pp. 980–990, 2003.
- [47] P. H. Faix, M. A. Burg, M. Gonzales, E. P. Ravey, A. Baird, and D. Larocca, "Phage display of cDNA libraries: enrichment of cDNA expression using open reading frame selection," *BioTechniques*, vol. 36, no. 6, pp. 1018–1029, 2004.
- [48] Y. Georgieva and Z. Konthur, "Design and Screening Of M13 phage display cDNA libraries," *Molecules*, vol. 16, no. 2, pp. 1667–1681, 2011.
- [49] K. Jacobsson and L. Frykberg, "Gene VIII-based, phage-display vectors for selection against complex mixtures of ligands," *BioTechniques*, vol. 24, no. 2, pp. 294–301, 1998.
- [50] L. S. Jaspers, J. H. Messens, A. de Keyser et al., "Surface expression and ligand-based selection of cdnas fused to filamentous phage gene VI," *Nature Biotechnology*, vol. 13, no. 4, pp. 378–382, 1995.
- [51] M. Hust, M. Meysing, T. Schirrmann et al., "Enrichment of open reading frames presented on bacteriophage M13 using hyperphage," *BioTechniques*, vol. 41, no. 3, pp. 335–342, 2006.
- [52] P. Kristensen and G. Winter, "Proteolytic selection for protein folding using filamentous bacteriophages," *Folding and Design*, vol. 3, no. 5, pp. 321–328, 1998.
- [53] J. Kügler, S. Nieswandt, G. F. Gerlach, J. Meens, T. Schirrmann, and M. Hust, "Identification of immunogenic polypeptides from a *Mycoplasma hyopneumoniae* genome library by phage display," *Applied Microbiology and Biotechnology*, vol. 80, no. 3, pp. 447–458, 2008.
- [54] A. Gupta, N. Shrivastava, P. Grover et al., "A novel helper phage enabling construction of genome-scale ORF-enriched phage display libraries," *PLoS ONE*, vol. 8, no. 9, Article ID e75212, 2013.
- [55] R. Cramer and M. Suter, "Display of biologically active proteins on the surface of filamentous phages: a cDNA cloning system for the selection of functional gene products linked to the genetic information responsible for their production [Gene 137 (1993) 69–75]," *Gene*, vol. 160, no. 1, 139 pages, 1995.
- [56] S. Cramer, R. Jaussi, G. Menz, and K. Blaser, "Display of expression products of cDNA libraries on phage surfaces. A versatile screening system for selective isolation of genes by specific gene-product/ligand interaction," *European Journal of Biochemistry*, vol. 226, no. 1, pp. 53–58, 1994.
- [57] D. Moisy, S. V. Avilov, Y. Jacob et al., "HMGB1 protein binds to influenza virus nucleoprotein and promotes viral replication," *Journal of Virology*, vol. 86, no. 17, pp. 9122–9133, 2012.
- [58] A.-J. van Zonneveld, B. M. M. van den Berg, M. van Meijer, and H. Pannekoek, "Identification of functional interaction sites on proteins using bacteriophage-displayed random epitope libraries," *Gene*, vol. 167, no. 1–2, pp. 49–52, 1995.
- [59] K. Hertveldt, M. L. Dechassa, J. Robben, and G. Volckaert, "Identification of Gal80p-interacting proteins by *Saccharomyces cerevisiae* whole genome phage display," *Gene*, vol. 307, no. 1–2, pp. 141–149, 2003.
- [60] K. Yano, T. Yoshino, M. Shionoya, S. Y. Sawata, K. Ikebukuro, and I. Karube, "Preparation of a whole genome phage library using fragmented *Escherichia coli* genome and its characterization of protein binding properties by surface plasmon resonance," *Biosensors and Bioelectronics*, vol. 18, no. 10, pp. 1201–1207, 2003.
- [61] E. Malini, E. Maurizio, S. Bembich, R. Sgarra, P. Edomi, and G. Manfioletti, "HMGA interactome: new insights from phage display technology," *Biochemistry*, vol. 50, no. 17, pp. 3462–3468, 2011.

- [62] R. di Niro, A. M. Sulic, F. Mignone et al., "Rapid interactome profiling by massive sequencing," *Nucleic Acids Research*, vol. 38, no. 9, pp. e110–e110, 2010.
- [63] J. A. Ubersax and J. E. Ferrell Jr., "Mechanisms of specificity in protein phosphorylation," *Nature Reviews Molecular Cell Biology*, vol. 8, no. 7, pp. 530–541, 2007.
- [64] D. Cochrane, C. Webster, G. Masih, and J. McCafferty, "Identification of natural ligands for SH2 domains from a phage display cDNA library," *Journal of Molecular Biology*, vol. 297, no. 1, pp. 89–97, 2000.
- [65] C. Drr, H. Nothaft, C. Lizak, R. Glockshuber, and M. Aebi, "The *Escherichia coli* glycopHage display system," *Glycobiology*, vol. 20, no. 11, pp. 1366–1372, 2010.
- [66] M. Fransens, P. P. van Veldhoven, and S. Subramani, "Identification of peroxisomal proteins by using M13 phage protein VI phage display: molecular evidence that mammalian peroxisomes contain a 2,4-dienoyl-CoA reductase," *Biochemical Journal*, vol. 340, part 2, pp. 561–568, 1999.
- [67] V. A. Somers, R. J. Brandwijk, B. Joosten et al., "A panel of candidate tumor antigens in colorectal cancer revealed by the serological selection of a phage displayed cDNA expression library," *Journal of Immunology*, vol. 169, no. 5, pp. 2772–2780, 2002.
- [68] C. Govarts, K. Somers, R. Hupperts, P. Stinissen, and V. Somers, "Exploring cDNA phage display for autoantibody profiling in the serum of multiple sclerosis patients: optimization of the selection procedure," *Annals of the New York Academy of Sciences*, vol. 1109, pp. 372–384, 2007.
- [69] K. Somers, P. Geusens, D. Elewaut et al., "Novel autoantibody markers for early and seronegative rheumatoid arthritis," *Journal of Autoimmunity*, vol. 36, no. 1, pp. 33–46, 2011.
- [70] S. Danner and J. G. Belasco, "T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 23, pp. 12954–12959, 2001.
- [71] Y. Kim, N. B. Caberoy, G. Alvarado, J. L. Davis, W. J. Feuer, and W. Li, "Identification of Hnrph3 as an autoantigen for acute anterior uveitis," *Clinical Immunology*, vol. 138, no. 1, pp. 60–66, 2011.
- [72] T. Tan, X. Xiang, H. Qu, S. Zhu, and Q. Bi, "The study on venom proteins of *Lapemis hardwickii* by cDNA phage display," *Toxicology Letters*, vol. 206, no. 3, pp. 252–257, 2011.
- [73] S. B. Lauterbach, R. Lanzillotti, and T. L. Coetzer, "Construction and use of *Plasmodium falciparum* phage display libraries to identify host parasite interactions," *Malaria Journal*, vol. 2, article 1, 2003.
- [74] H. J. Ren, R. D. Liu, Z. Q. Wang, and J. Cui, "Construction and use of a *Trichinella spiralis* phage display library to identify the interactions between parasite and host enterocytes," *Parasitology Research*, vol. 112, no. 5, pp. 1857–1863, 2013.
- [75] N. B. Caberoy, Y. Zhou, X. Jiang, G. Alvarado, and W. Li, "Efficient identification of tubby-binding proteins by an improved system of T7 phage display," *Journal of Molecular Recognition*, vol. 23, no. 1, pp. 74–83, 2010.
- [76] N. B. Caberoy, D. Maignel, Y. Kim, and W. Li, "Identification of tubby and tubby-like protein 1 as eat-me signals by phage display," *Experimental Cell Research*, vol. 316, no. 2, pp. 245–257, 2010.
- [77] N. B. Caberoy, G. Alvarado, and W. Li, "Identification of calpain substrates by ORF phage display," *Molecules*, vol. 16, no. 2, pp. 1739–1748, 2011.
- [78] B. Luo, T. Zou, N. Lu et al., "Role of suppressor of cytokine signaling 3 in lipid metabolism: analysis based on a phage-display human liver cDNA library," *Biochemical and Biophysical Research Communications*, vol. 416, no. 1-2, pp. 39–44, 2011.
- [79] E. Beghetto and N. Gargano, "Lambda-display: a powerful tool for antigen discovery," *Molecules*, vol. 16, no. 4, pp. 3089–3105, 2011.
- [80] I. Kuwabara, H. Maruyama, Y. G. Mikawa, R. I. Zuberi, F. T. Liu, and I. N. Maruyama, "Efficient epitope mapping by bacteriophage lambda surface display," *Nature Biotechnology*, vol. 15, no. 1, pp. 74–78, 1997.
- [81] E. Beghetto, N. Gargano, S. Ricci et al., "Discovery of novel *Streptococcus pneumoniae* antigens by screening a whole-genome λ -display library," *FEMS Microbiology Letters*, vol. 262, no. 1, pp. 14–21, 2006.
- [82] H. Ansuini, C. Cicchini, A. Nicosia, M. Tripodi, R. Cortese, and A. Luzzago, "Biotin-tagged cDNA expression libraries displayed on lambda phage: a new tool for the selection of natural protein ligands," *Nucleic acids research*, vol. 30, no. 15, p. e78, 2002.
- [83] S. Kosuri and G. M. Church, "Large-scale de novo DNA synthesis: technologies and applications," *Nature Methods*, vol. 11, no. 5, pp. 499–507, 2014.
- [84] H. B. Larman, Z. Zhao, U. Laserson et al., "Autoantigen discovery with a synthetic human peptidome," *Nature Biotechnology*, vol. 29, no. 6, pp. 535–541, 2011.
- [85] G. Fuh, M. T. Pisabarro, Y. Li, C. Quan, L. A. Lasky, and S. S. Sidhu, "Analysis of PDZ domain-ligand interactions using carboxyl-terminal phage display," *The Journal of Biological Chemistry*, vol. 275, no. 28, pp. 21486–21491, 2000.
- [86] R. Mrowka, A. Patzak, and H. Herzel, "Is there a bias in proteome research?" *Genome Research*, vol. 11, no. 12, pp. 1971–1973, 2001.
- [87] Y. Ivarsson, R. Arnold, M. McLaughlin et al., "Large-scale interaction profiling of PDZ domains through proteomic peptide-phage display using human and viral phage peptidomes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 7, pp. 2542–2547, 2014.
- [88] H. Huang and S. S. Sidhu, "Studying binding specificities of peptide recognition modules by high-throughput phage display selections," *Methods in Molecular Biology*, vol. 781, pp. 87–97, 2011.
- [89] W. L. Matochko, K. Chu, B. Jin, S. W. Lee, G. M. Whitesides, and R. Derda, "Deep sequencing analysis of phage libraries using Illumina platform," *Methods*, vol. 58, no. 1, pp. 47–55, 2012.
- [90] A. H. Y. Tong, B. Drees, G. Nardelli et al., "A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules," *Science*, vol. 295, no. 5553, pp. 321–324, 2002.
- [91] X. Xin, D. Gfeller, J. Cheng et al., "SH3 interactome conserves general function over specific form," *Molecular Systems Biology*, vol. 9, article 652, 2013.