

Research Progress of Phage Display System

Jianxin Wen*, Kunpeng Yuan

College of Veterinary Medicine, Qingdao Agricultural University, Qingdao, China

Email: *wenjianxin@126.com

How to cite this paper: Wen, J.X. and Yuan, K.P. (2021) Research Progress of Phage Display System. *Advances in Microbiology*, 11, 181-189.

<https://doi.org/10.4236/aim.2021.113013>

Received: January 29, 2021

Accepted: March 19, 2021

Published: March 22, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Phage display technology is a unique gene recombination expression technology, and it is also a simple and effective screening tool. Through panning, a protein or peptide with high affinity and selectivity to the target is obtained. Antibody phage display has become the first and most widely used *in vitro* screening technology. Phage display derivatives play an important role in the diagnosis and treatment of diseases. This article reviews the phage display system of phage display technology, the size and classification of antibody libraries and their applications, and discusses the application prospects and challenges of phage display technology. This thesis lays the foundation for the theoretical and experimental research of bacteriophages.

Keywords

Phage Display Technology, Phage Display System, Antibody Library, Prospects and Challenges

1. Introduction

Phage display technology was first proposed by Smith in 1985. Smith successfully integrated exogenous DNA into M13 phage, and expressed the exogenous polypeptide with the coat protein of M13 phage. McCafferty *et al.* described antibody phage display for the first time. They were able to fuse the encoding scFv to gene III. Without affecting the infectivity of the phage and maintaining the functionality of the antibody, the scFv was fused with the capsid protein and expressed and displayed on the surface of the phage. This method is different from other bacterial expression systems in that the protein (phenotype) is directly linked to its homologous gene (genotype) through phage.

The phage display system can display a sufficient level of antibodies on the surface of phage virus particles, so that phage that recognizes the antigen can be selected. The ultimate goal of phage display is to screen out a large number of

non-specific phage clones that bind to the target antigen with high affinity. Panning can enrich clones with high affinity, and the steps include binding, washing, rescue and reinfection. Firstly, the phage is bound to the target, and then the non-specific binding to the target is removed by washing, and then the remaining phage particles bound to the target are eluted and recovered. After each round, the eluted phage is infected and amplified by *E. coli*. After 3 to 4 rounds of panning, a large number of clones are obtained and identified. Finally, the final round of panning phage monoclonal antibodies is screened and the selected monoclonal antibodies are obtained by DNA sequencing.

2. Phage Display System

The protein expressed on the surface of the phage can be a polypeptide or an antibody fragment, such as scFv. Although there are other *in vitro* expression methods, such as yeast surface display, ribosome display or puromycin display, the phage display system is the most commonly used display system. The core of phage display technology is the phage used to display antibodies. Different phage can be used in the phage display system, including T4, λ and filamentous M13 phage. The most widely used is the M13 phage.

The filamentous phage has a filamentous structure and consists of a circular single-stranded DNA (ssDNA) genome. Among the filamentous phages that infect *E. coli*, the most typical phages are M13, f1 and fd, all of which belong to Ff phages. The ssDNA genome of Ff bacteriophage is approximately 6400 bp in length and encodes 11 different genes. These genes are divided into structural proteins (genes III, VI, VII, VIII, IX) and functional proteins required for phage replication and assembly (genes I, II, IV, V, X, XI). The phage genome also contains an ori site (origin of replication), which is responsible for the production of (+) and (-) DNA strands. Another site is called the “packaging signal” and is responsible for initiating phage assembly. The length of the Ff virion is approximately 900 nm and the diameter is 6.5 nm. The ssDNA genome of the Ff virion is completely enclosed in a cylinder composed of 2700 main protein pVIII molecules. The asymmetric end of the Ff phage contains 5 minor coat proteins pIII and pVI at one end, and pVII and pIX proteins at the other end. A distinctive feature of Ff bacteriophage is that it has the ability to replicate but does not kill the host bacteria. After the host bacteria are infected, the Ff phage will replicate as an extrachromosomal element instead of integrating itself into the host chromosome, so that the host bacteria can continuously shed virus particles.

Therefore, M13 and related Ff phages are preferable to lytic phages (such as T4 and T7). M13 phage has high replication ability and can accommodate a large amount of foreign DNA, making it the most commonly used phage display vector. Infection of the host bacteria is achieved by attaching the phage pIII to the F'-fimbria of *E. coli*. After ssDNA enters the bacteria, it is converted into double-stranded DNA by host enzymes to produce ssDNA and phage protein. Phage assembly occurs in or near the inner membrane of the host bacteria and

in the periplasm [1]. Each M13 particle has 3-5 copies of pIII protein, and the most abundant coat protein pVIII has 2700 copies.

At present, phage display technology often uses phagemid vector system. Phagemid vector (4.6 kb) is a plasmid encoding several key elements, including the origin of replication of bacteria and phage, leader sequence, multiple cloning sites, antibiotics Resistance gene, phage coat protein gene (gIII or gVIII) and weak promoter (lacZ) [2]. The phagemid vector alone cannot produce infectious phage particles. With the help of helper phage (such as M13KO7 or VCSM13), the replication and packaging of the phage can be completed, and the mature phage can be released into the external environment [3]. Helper phages are modified Ff phages with defective packaging signals (M13 intergenic region). Therefore, compared with the phagemid vector carrying the M13 gene, the efficiency of replication and packaging is lower. In order to produce the fusion protein, the *Escherichia coli* with the phagemid vector is super-infected with helper phage. Initially, phagemid vectors were replicated in *E. coli* in the form of dsDNA. After co-infection with helper phage, ssDNA and phage coat protein are produced, which starts to package and release mature phage particles. Phagemid vectors are small and relatively easy to clone, which improves transformation efficiency and can generate larger libraries. Phagemid display enables the monovalent display of fusion proteins because the resulting phage particles usually only carry a single copy of the fusion protein. In addition, only 10% or less of phage particles will present the fusion protein.

3. The Size of the Antibody Library

The size of the antibody library is directly related to the probability of the appearance of the target antibody. The larger the antibody library, the greater the chance of screening high-affinity antibodies, and the greater the chance that antibodies will specifically bind to epitopes with higher affinity [4]. The antigen binding site is composed of six complementarity determining regions (CDR): LCDR1, LCDR2, LCDR3, HCDR1, HCDR2 and HCDR3. The average length of CDR is 10 amino acids. Each CDR is diversified by 20 natural amino acids, and the theoretical library will contain 1.2×10^{78} unique antibody variants. Obviously only a very small part can be displayed in the phage display antibody library. On the other hand, the transformation efficiency of *E. coli* and the number of repeated transformations limit the size of the phage display antibody library to 10^{10} - 10^{11} .

Ideally, the size of the library should be equal to its effective size. The factors affecting the effective size of the antibody library are as follows 1) the quality of gene synthesis [5]. 2) The nucleotide sequence contains a stop codon. 3) Insertion and deletion of nucleotides. 4) Antibody genes are not applicable to *E. coli* [6]. 5) The use of oligonucleotides or degenerate NNK codons may reduce the diversity of amino acids. The effective size of the library is usually expressed as a percentage. Generally, the effective size is an order of magnitude lower than the absolute size of the library. The effect of different CDRs of scFv on the antigen

binding site is CDRH3 > CDRH2 > CDRL3 > CDRH1 > CDRL1 > CDRL1 > CDRL2 > framework region. Since CDRH3 has the highest natural diversity and contributes the most to the antigen-binding site, HCDR3 is located in the center of the antigen-binding site and plays a key role in recognizing different targets. It may produce antibodies with reasonable affinity [7]. Therefore, in addition to the number of antibodies, the number of functional molecules that can recognize different targets determines the effective size of the antibody library.

4. Classification of Antibody Libraries

The phage single-chain antibody library mimics the process of B cell assembly and antibody production. In fact, every B cell is a self-replicating packaging system that contains antibody genes encoding antibodies displayed on its surface. First, in the natural process of B cell development, functional immunoglobulin (Ig) genes are randomly assembled from variable heavy (VH) and variable light (VL) gene fragments to generate all foreign antigens. Second, during antigen activation, each B cell undergoes a process called somatic hypermutation, during which, especially amino acid residues in the complementarity determining region (CDR) are mutated [8]. It plays a vital role in improving the affinity and selectivity of antibodies to their target antigens. In phage display, this *in vivo* step can be imitated by methods such as site-directed mutagenesis and recombination of heavy and light chains. After affinity maturation, B cells expressing antibodies with high affinity to the target antigen are cloned and expanded to generate memory B cells and plasma cells that secrete antibodies.

The variable regions of the heavy and light chains of an antibody are determined by the binding of the target. During the construction of the antibody library, different antibody libraries can be obtained by randomly combining VL and VH [9]. One antibody library may represent countless A gene combination method, phage display technology mimics the natural antibody library through the different antibodies presented by billions of phage particles [10]. The source of the antibody library used for library preparation has a great influence on the type of antibody library generated. It may come from different hosts, not humans [11], and different immune responses in health and disease states will have a profound impact on the diversity of antibody lineages.

Antibody libraries can be divided into two main types according to the sources of VH and VL gene fragments used in the library construction process. The natural antibody phage library is composed of V genes obtained from immunized [12] and non-immunized donors. The synthetic antibody phage library consists of V genes designed to be partially or fully synthesized *in vitro*.

4.1. Natural Antibody Library

The natural antibody library is a collection of immunoglobulins extracted from lymphoid tissues (bone marrow, spleen and tonsils) or B cells in peripheral blood. The gene combination of the natural antibody library can be from hu-

mans or other animals. Under normal circumstances, a person has at least 10⁸ antibody-producing B cell clones [9]. The clonal diversity of B cells allows it to produce different antibody populations that can target almost all types of antigens. Natural antibody libraries can be used to develop antibodies related to cancer, autoimmune diseases, infectious diseases and other diseases [13] for diagnostic or therapeutic applications. Compared with antibodies in immune libraries, natural antibody libraries generally have lower antibody affinity and higher cross-reactivity. This can be solved by constructing large-scale and diverse antibody libraries or by *in vitro* affinity optimization [14].

4.2. Immune Antibody Library

The ideal sample for the immune antibody library is patients who are in the acute infection or recovery phase, and who have recovered from a specific infection or disease [3]. A unique feature of the immune antibody library is that the sample comes from activated B cells, which are activated when the antigen invades, and then undergo an affinity maturation process. Therefore, the tendency of antibody clones to recognize certain antigens is obvious and allows the isolation of high-affinity binding partners of the target antigen. Unlike the immune library, the antibody library after exposure to the antigen has a preference for V genes, which will produce a large number of antigen-specific antibodies. Most importantly, antibody clones that have undergone the affinity maturation process (*i.e.*, somatic hypermutation and clonal selection amplification) will have high copy numbers, thus increasing the possibility of enriching high-affinity antibody clones. Therefore, the size of the immune library does not need to be as large as the natural library [15]. Compared with natural libraries, immune libraries are not suitable for targeting a large group of antigens, especially self-antigens. This is because the immune system has developed immune tolerance to self-antigens. The sample source of the immune library is limited to patients with disease infection and is difficult to obtain [3], and the immune library is mainly for antigens of specific diseases. When targeting antigens of different diseases, the library needs to be rebuilt [16].

4.3. Semi-Synthetic Antibody Library and Synthetic Antibody Library

Unlike natural and immune libraries, semi-synthetic libraries are composed of a mixture of natural and chemically synthesized sequences [14]. Gene synthesis methods are usually used to provide random CDR combinations. The advantage of semi-synthetic antibody libraries is that the framework sequence to be used is predetermined. This will help the development and application of specific antibodies. The semi-synthetic library is carefully designed by assuming the number of scaffolds, the positions to be diversified, the types of amino acids to be included in the design, and the ratio of amino acids to be diversified at each position. These assumptions are not always correct, especially in HCDR3. In order to

avoid making assumptions about the structure of CDRs, a combination of synthetic scaffolds and natural CDRs was used [17].

The main difference between a semi-synthetic library and a synthetic library is that the entire antibody sequence used in the synthetic library is chemically derived. Unlike other antibody libraries, synthetic antibody libraries are purely artificial libraries, and V genes are usually reconstructed *in vitro* through CDR randomization. Because it has no natural prejudice and redundancy caused by evolutionary influence, and it can target many different kinds of antigens. Synthetic libraries are also useful for targeting non-immunogenicity, toxicity, and autoantigens [18]. Synthetic libraries can be further classified according to the type of framework used, the origin and design of sequence diversity in CDRs, and the method of library generation. The huge diversity in the synthetic library is provided by the pre-determined framework regions and CDR diversity. The synthetic antibody library is basically generated from a series of bioinformatics analyses using existing experimental data, including antibody epitopes, antigen-antibody interactions, affinity maturation design, variable gene fragment recombination, and variable region structure prediction. These studies provide valuable information for the amino acid dominance and variation in the CDR regions [19]. In order to improve the quality of the synthetic library, the framework region was designed with high frequency and consensus sequence and optimized, and the 6 CDRs were randomly grouped using trinucleotide mutation (TRIM) technology [20], according to different lengths. The use frequency of amino acids in HCDR3 is different to redesign HCDR3 and optimize amino acids. The CDR is designed according to the known CDR concept and standard structure [21]. The design of HCDR3 is an eternal topic and may also be a major opportunity for improvement. HCDR3 is a key factor in determining antibody specificity and affinity, but it is also the most diverse region of antigen binding sites so far, so it is difficult to design. The 3D modeling method can predict the structure of all CDRs except HCDR3. However, there is currently no method that can reliably predict the structure of HCDR3, which limits our ability to correctly design the diversity of this antigen binding site region.

5. Prospects and Challenges

Phage display technology can produce antibodies *in vitro* without host immunization. Bacteriophages can withstand the influence of factors such as high temperature, pH, denaturant, ultraviolet radiation, and proteolytic enzymes. This function allows the separation of different types of antibodies that are stable or least affected by harsh environments. Therefore, in phage display technology, the conditions and selection pressure can be customized according to needs, and a selection method that is impossible in *in vivo* antibody production can be obtained.

Through the study of phage display and the chimeric antibody cetuximab, cytuzumab (a human antibody used to treat cancer) was discovered. Another ex-

ample is Adalimumab, which is the first fully humanized antibody to enter the market and the world's best-selling drug. Adalimumab was discovered and optimized through phage display through guided selection of mouse antibodies. In fact, until 2017, six fully human therapeutic antibodies discovered or engineered through phage display have been approved by the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA), and there are hundreds of antibodies [13].

Ten years ago, the concept of developability was applied to the development of antibody drugs [22]. Developability includes design principles and experimental evaluation of the characteristics that the molecule should meet to further develop or manufacture, formulate and stabilize antibodies to achieve the desired therapeutic effect. Developability is affected by many factors, including the intrinsic physical and biochemical properties of the molecule, as well as external parameters, such as ionic strength, pH, and additive formulation. As more antibodies enter the market, many antibodies have failed to play a role in pre-clinical development and clinical trials. People realized that antibodies selected from different libraries, despite having the required specificity and affinity, are in the formulation and in the process of manufacturing and development, it often fails due to inappropriate physical and chemical properties. This includes glycosylation of residues in the antigen binding site, which may impair the ability of the antibody to bind to the antigen. This is especially important for antibodies selected for phage display libraries, because *E. coli* does not glycosylate proteins. Therefore, in the phage display discovery activity, variants with glycosylation sites on CDRs can be selected, but they will lose binding when transformed and expressed in mammalian cells. Unpaired cysteine can cause disulfide bonds to be disrupted, resulting in covalent aggregates. In addition, antibodies undergo post-translational modifications (PTM) of amino acids, such as asparagine (N) deamidation, methionine (M) oxidation, and aspartate (D) isomerization [23]. If the amino acids are involved in antigen interactions, chemical modifications of these amino acids can lead to failure of antibody preparation and treatment.

Acknowledgements

This work was supported by a grant from the National Key Research and Development Program of China awarded to YKP (Grant No. 2017YFD0501004) and Special Economic Animal Innovation Team of Shandong Modern Agricultural Industrial Technology System (SDAIT-21).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Straus, S.K. and Bo, H.E. (2018) Filamentous Bacteriophage Proteins and Assembly. In: Harris, J. and Bhella, D., Eds., *Virus Protein and Nucleoprotein Complexes*, Vol.

- 88, Springer, Singapore, 261-279. https://doi.org/10.1007/978-981-10-8456-0_12
- [2] Qi, H., Lu, H.Q., Qiu, H.J., Petrenko, V. and Liu, A.H. (2012) Phagemid Vectors for Phage Display: Properties, Characteristics and Construction. *Journal of Molecular Biology*, **417**, 129-143. <https://doi.org/10.1016/j.jmb.2012.01.038>
- [3] Bazan, J., Całkosiński, I. and Gamian, A. (2012) Phage Display—A Powerful Technique for Immunotherapy. *Human Vaccines & Immunotherapeutics*, **8**, 1829-1835. <https://doi.org/10.4161/hv.21704>
- [4] Perelson, A.S. (1989) Immune Network Theory. *Immunological Reviews*, **110**, 5-36. <https://doi.org/10.1111/j.1600-065X.1989.tb00025.x>
- [5] Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellenhofer, G., *et al.* (2000) Fully Synthetic Human Combinatorial Antibody Libraries (HuCAL) Based on Modular Consensus Frameworks and CDRs Randomized with Trinucleotides. *Journal of Molecular Biology*, **296**, 57-86. <https://doi.org/10.1006/jmbi.1999.3444>
- [6] Lin, B., Renshaw, M.W., Autote, K., Smith, L.M., Calvey, P., Bowdish, K.S., *et al.* (2008) A Step-Wise Approach Significantly Enhances Protein Yield of a Rationally-Designed Agonist Antibody Fragment in *E. coli*. *Protein Expression & Purification*, **59**, 55-63. <https://doi.org/10.1016/j.pep.2008.01.002>
- [7] Zadeh, A.S., Grässer, A., Dinter, H., Hermes, M. and Schindowski, K. (2019) Efficient Construction and Effective Screening of Synthetic Domain Antibody Libraries. *Methods and Protocols*, **2**, Article No. 17. <https://doi.org/10.3390/mps2010017>
- [8] Tonegawa, S. (1983) Somatic Generation of Antibody Diversity. *Nature*, **302**, 575-581. <https://doi.org/10.1038/302575a0>
- [9] Lerner, R.A. (2016) Combinatorial Antibody Libraries: New Advances, New Immunological Insights. *Nature Reviews Immunology*, **16**, 498-508. <https://doi.org/10.1038/nri.2016.67>
- [10] Beerli, R.R. and Rader, C. (2010) Mining Human Antibody Repertoires. *mAbs*, **2**, 365-378. <https://doi.org/10.4161/mabs.12187>
- [11] Koti, M., Saini, S. and Sachan, A.K. (2014) Engineered Bovine Antibodies in the Development of Novel Therapeutics, Immunomodulators and Vaccines. *Antibodies*, **3**, 205-214. <https://doi.org/10.3390/antib3020205>
- [12] Mccafferty, J., Griffiths, A.D., Winter, G. and Chiswell, D.J. (1990) Phage Antibodies: Filamentous Phage Displaying Antibody Variable Domains. *Nature*, **348**, 552-554. <https://doi.org/10.1038/348552a0>
- [13] Frenzel, A., Schirrmann, T. and Hust, M. (2016) Phage Display-Derived Human Antibodies in Clinical Development and Therapy. *mAbs*, **8**, 1177-1194. <https://doi.org/10.1080/19420862.2016.1212149>
- [14] Hoogenboom, H.R. (1992) By-Passing Immunization. Human Antibodies from Synthetic Repertoires of Germline V_H Gene Segments Rearranged *in Vitro*. *Journal of Molecular Biology*, **227**, 381-388. [https://doi.org/10.1016/0022-2836\(92\)90894-P](https://doi.org/10.1016/0022-2836(92)90894-P)
- [15] Rader, C., Steinberger, P. and Barbas 3rd., C. (2001) Selection from Antibody Libraries. Cold Spring Harbor Laboratory, Cold Spring Harbor, 10.12-10.20.
- [16] Winter, G., Griffiths, A.D., Hawkins, R.E. and Hoogenboom, H.R. (1994) Making Antibodies by Phage Display Technology. *Annual Review of Immunology*, **12**, 433-455. <https://doi.org/10.1146/annurev.iy.12.040194.002245>
- [17] Hoet, R.M., Cohen, E.H., Kent, R.B., Rookey, K., Schoonbroodt, S., Hogan, S., *et al.* (2005) Generation of High-Affinity Human Antibodies by Combining Donor-Derived and Synthetic Complementarity-Determining-Region Diversity. *Nature Biotechnology*, **23**, 344-348. <https://doi.org/10.1038/nbt1067>

- [18] Clementi, N., Mancini, N., Solforosi, L., Castelli, M., Clementi, M. and Burioni, R. (2012) Phage Display-Based Strategies for Cloning and Optimization of Monoclonal Antibodies Directed against Human Pathogens. *International Journal of Molecular Sciences*, **13**, 8273-8292. <https://doi.org/10.3390/ijms13078273>
- [19] Zhao, J., Nussinov, R., Wu, W. and Ma, B. (2018) In Silico Methods in Antibody Design. *Antibodies*, **7**, Article No. 22. <https://doi.org/10.3390/antib7030022>
- [20] Sondek, J. and Shortle, D. (1992) A General Strategy for Random Insertion and Substitution Mutagenesis: Substoichiometric Coupling of Trinucleotide Phosphoramidites. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 3581-3585. <https://doi.org/10.1073/pnas.89.8.3581>
- [21] Chothia, C. and Lesk, A.M. (1987) Canonical Structures for the Hypervariable Regions of Immunoglobulins. *Journal of Molecular Biology*, **196**, 901-917. [https://doi.org/10.1016/0022-2836\(87\)90412-8](https://doi.org/10.1016/0022-2836(87)90412-8)
- [22] Chennamsetty, N., Voynov, V., Kayser, V. and Trout, B.L. (2010) Prediction of Aggregation Prone Regions of Therapeutic Proteins. *The Journal of Physical Chemistry B*, **114**, 6614-6624. <https://doi.org/10.1021/jp911706q>
- [23] Gilliland, G.L., Luo, J., Vafa, O. and Carlos Almagro, J. (2012) Leveraging SBDD in Protein Therapeutic Development: Antibody Engineering. In: Tari, L., Ed., *Methods in Molecular Biology*, Vol. 841, Humana Press, Totowa, 321-349. https://doi.org/10.1007/978-1-61779-520-6_14