# Lessons from bacteriophages part 1: Deriving utility from protein structure, function, and evolution

**Kunica Asija, Carolyn M. Teschke***

Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut, United States of America
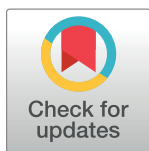
* carolyn.teschke@uconn.edu

## A historical introduction to research using phages

Historically, some of the most fundamental discoveries of modern molecular biology were revealed by examination of phage-infected cells [1]. Some examples include the use of bacteriophage T2 to show that DNA, not protein, was the genetic material by Hershey and Chase in 1952 [2]. Crick and colleagues showed that codons are degenerate and encode for single amino acids by making use of T4 bacteriophage. Brenner and colleagues discovered that mRNA acts as the messenger in relaying the information from DNA to ribosomes with the aid of bacteriophage T2 [3]. Szybalski and colleagues demonstrated the semiconservative mechanism of DNA replication using λ phage [4]. Restriction modification, a bacterial innate immune system that evolved as protection against invading mobile genetic elements such as plasmids and phages, was discovered in bacteriophage-infected cells. Restriction mapping was crucial for the completion of the human genome project. The mechanism of action of heat shock genes—which encode for molecular chaperones such as *dnaK*, *dnaJ*, and *grpE*, as well as the *groE* gene locus, which encodes for the molecular chaperones GroES and GroEL—were discovered in λ phage–infected *Escherichia coli* [5]. The first capsid assembly inhibitor, the fluorescent dye 4,4'–bis(1-anilinonaphthalene-8-sulfonic acid) (bisANS), was discovered in experiments investigating assembly of phage P22 [6]. These are just some of the examples of contributions made by bacteriophage research in the past few decades. Here, we discuss additional discoveries derived from studies of phages.

## Insights from phage capsid protein structures and assembly

Dennis Bamford and colleagues made a key observation that links eukaryotic, archaeal, and prokaryotic viruses: Despite the lack of sequence homology, viruses from all domains of life have capsid proteins that fall into four structural lineages [7]. These lineages are likely due to divergent evolution from common viral progenitors [7]. For example, the picornavirus-like lineage has a characteristic jelly roll or β-barrel roll fold in its capsid protein. This fold is seen in plus-sense single-stranded RNA (ssRNA), ssDNA, double-stranded RNA (dsRNA), and dsDNA viruses and phages, although it is most common in RNA viruses [7]. As another example, the dsDNA phages and herpesviruses have a conserved coat protein fold known as the Hong Kong 97 (HK97)-fold. We call this "nature's favorite building block" due its common occurrence in the dsDNA class of viruses, which are the most abundant entities on Earth [8]. With its perfect wedge shape that can readily form the hexons and pentons requisite for building an icosahedral capsid, the HK97-fold can be used to build an amazing array of capsids with triangulation numbers ranging from 1 to 52, yielding capsid diameters of 25

nm to 160 nm [9]. While the viruses in the HK97-fold lineage have coat proteins that share less than 10%–15% sequence similarity, they have similar assembly mechanisms. For instance, several parallels can be drawn in the assembly processes of Herpesviridae and members of the Caudoviridae family comprising the tailed bacteriophages, including P22, T4, Φ29, and HK97. These include the formation of an unstable procapsid intermediate, the use of an internal scaffolding protein (or delta domain) to aid in the formation of the procapsid that is absent from the mature virion, and the use of a unique portal vertex for the uptake of DNA into the empty procapsid during maturation [10]. In case of T4 and HK97 phages, the scaffolding protein is proteolytically cleaved during the maturation process, a feature seen in herpesvirus as well [11]. These proteases, belonging to the procapsid protease superfamily, are evolutionarily related [8, 12].

The structural homology observed in coat proteins can be generalized and applied to other structurally related viruses. If we begin by considering the phage HK97 coat protein (Fig 1), the G-loops interact with the E-loop of an adjacent subunit to form a salt bridge that stabilizes the capsomer interactions in the procapsid across a 2-fold axis of symmetry [13]. In phage P22, an extra domain (I-domain) is inserted within the A-domain of the HK97 core of the coat protein (Fig 1) [14, 15]. The D-loops of the I-domain make stabilizing contacts across the 2-fold axis of symmetry with the adjacent subunit, similar to the G-loops (within the P domain) in the HK97 coat protein [16]. The roles of these loops in the proper procapsid formation have been elucidated experimentally in both HK97 and P22: Aberrant structures such as tubes result from substitutions in these loops. A similar theme can be seen with the E-loop of the HK97 fold, again with slight variations as we compare coat proteins. In the HK97 coat protein, E-loop interactions stabilize the capsids through making covalent isopeptide crosslinks with the P-domain of the adjacent subunit [13], as well as ionic interactions with G-loops [17]. In phage T4's coat protein, a domain is inserted into the E-loop. This inserted domain interacts with an adjacent subunit, stabilizing the capsid in a fashion analogous to the crosslinks in HK97 [18]. Structural evaluation of the phage P22 capsid indicates that the E-loop makes salt bridges with the spine helix of the neighboring subunit to stabilize the capsid [15]. Overlay of cryoEM density maps of herpes simplex virus-1 (HSV-1) base domain with HK97 coat protein highlights the importance of the E-loop in HSV-1 capsid as well. The E-loop of HSV-1 contacts the adjacent subunits at multiple points, including the spine helix and the middle and upper domains. The density maps also show that the E-loop of the HSV-1 capsid interacts with the N-arm of an adjacent subunit, similar to that observed in phage HK97 [19]. These data indicate that the E-loop plays a crucial role in making stabilizing contacts with other subunits, and it appears to be conserved between structurally similar phages and viruses.

## Phages and vaccines

The repeated use of the same subunit to build an icosahedral capsid, and the high stability of phage capsids, has made phages ideal vehicles for vaccine development. Phage particles with displayed antigens and virus-like particles (VLPs) robustly provoke immune responses. When phage F1 was conjugated with B2.1 peptide, the immune response generated in mice was much stronger than when the peptide was conjugated with ovalbumin [20], and when the filamentous bacteriophage *fd* was conjugated with peptide epitopes, both humoral and cellular immune responses were elicited by activating major histocompatibility complexes (MHC) I and II after being taken up by lymphocytes [21]. In an additional example, λ phage decorator protein gpD (the product of gene D) fused with fragments of aspartate β-hydroxylase, a protein
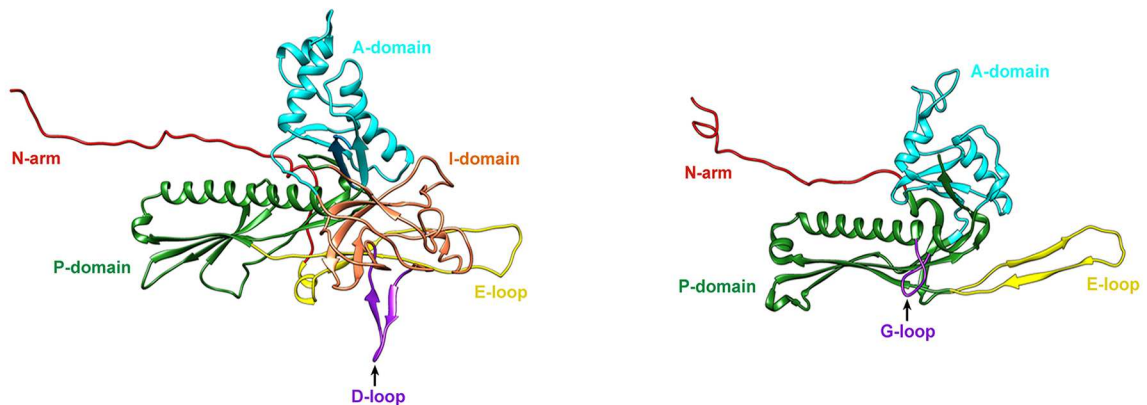
A
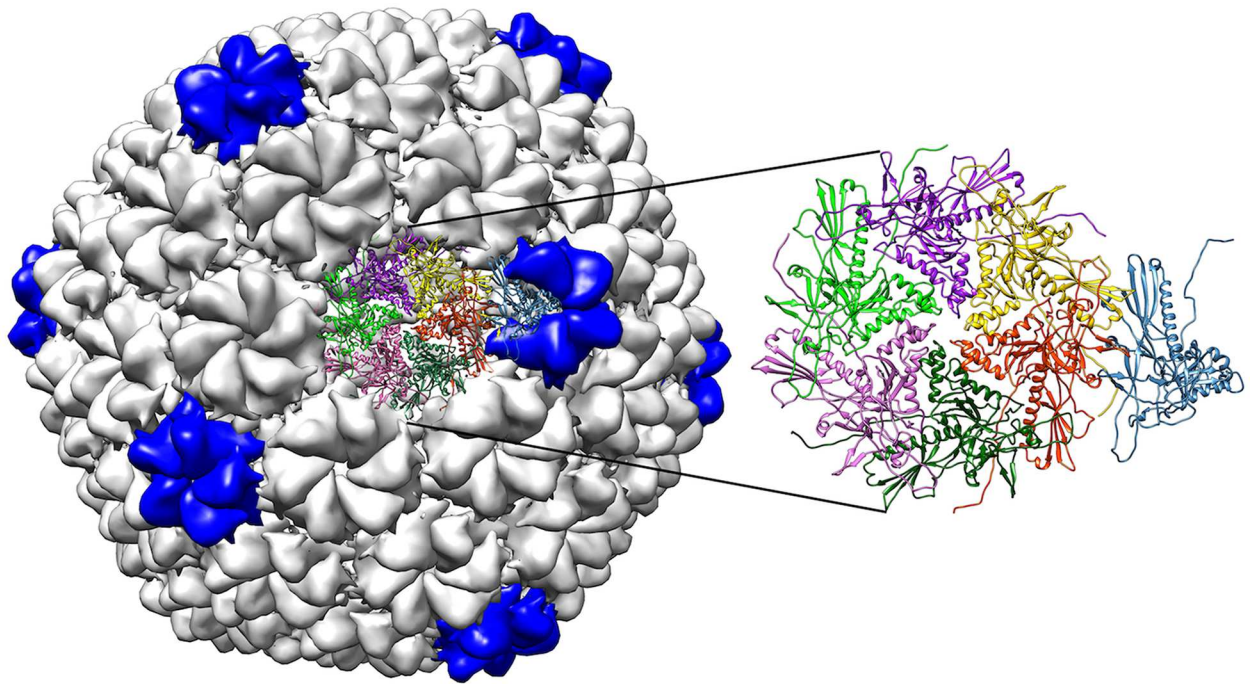


B

Fig 1. (A) Capsid structure of bacteriophage P22. The left side of panel A shows the mature P22 capsid. The pentons are highlighted in blue, and hexons are shown in gray. The asymmetric unit (seven subunits) of the capsid is represented in ribbon format in the center of the capsid. The asymmetric unit is magnified and shown on the right of the panel. (B) Ribbon diagrams of coat monomers of P22 (left) and HK97 (right). The A-domains, P-domains, E-loops, and N-arms have been highlighted in cyan, green, yellow, and red, respectively. The I-domain of P22 coat protein has been highlighted in coral. The arrows indicate the D-loop (purple) of the monomer of P22 (left) and the G-loop (purple) on the monomer of HK97 (right).

overexpressed in hepatocarcinoma cells, provoked a strong antitumor immune response in mice [22].

VLPs are also used for vaccines. VLPs are composed of only the protein structural units of a virus and lack genetic material, thereby rendering them noninfectious (Fig 2). These particles illicit a strong immune response of both the innate as well as the adaptive immune systems

[23]. The size of VLPs ranges from 20–200 nm, allowing them to be endocytosed by different cells in the human body, including antigen-presenting cells such as dendritic cells and macrophages [23]. Both RNA and DNA phages (MS2, PP7, Qβ, P22) are being used to generate VLPs for vaccines [24]. VLPs from the RNA bacteriophage Qβ are being tested as vaccine carriers against Alzheimer disease, malignant melanomas, and type II diabetes mellitus, all of which are in different phases of clinical trials [25].

Phages are being engineered for DNA vaccines, in which the gene encoding the antigenic epitope is cloned into the phage genome but is under the control of a eukaryotic expression cassette (Fig 2). A eukaryotic expression cassette is composed of a promoter, an open reading frame (ORF), and a polyadenylated 3' untranslated region (UTR). Phage DNA vaccines are easily taken up by antigen-presenting cells, and the phage capsids protect the DNA from nucleases. About 15–20 kb expression constructs can be carried within phages such as λ, which means multiple vaccines can be cloned in the phage genome at once. A DNA vaccine for the small surface antigenic (HBsAg) of hepatitis B virus has been developed using λ phage. This phage vaccine improves upon the previously used hepatitis B vaccines, which failed to elicit a strong immunogenic response. The phage vaccine has been successfully tested on rabbits, in which it is highly immunogenic, and is now awaiting human trials [26].

## Insights based on phage tail machinery

Evolutionary links also can be seen in the tail machinery of bacteriophages and the type VI secretion system (T6SS) of bacteria [27]. The T6SS system is used by many Gram-negative α-, β-, and γ-proteobacteria to deliver proteins to other bacteria and eukaryotes and has features that are structurally analogous to contractile myophage tails. The T6SS system is comprised of three structural components: the inner tube, sheath, and baseplate. The hemolysin coregulated protein (HCP), a crucial component of the T6SS, spontaneously assembles into hexamers to form the tube. The tip of the tube contains a trimer of the VgrG protein, another hallmark protein of the T6SS. A sheath wraps around the tube to form an inverted phage tail–like structure [28] and provides the tube with a contractile function to release effector molecules. HCP protein is structurally related to the tail tube protein (gpV) of λ bacteriophage [29]. The VgrG protein shares sequence as well as structural similarity to the tail needle machinery of the T4 bacteriophage. It also shares functional homology, since it is responsible for puncturing target cells [27, 30]. The general mechanism of assembly for the two systems is similar in the need for a baseplate and the use of the tube as a template for the polymerization of the sheath [31].

The R- and F-type pyocins produced by *Pseudomonas aeruginosa* are structurally similar to phage tails and are thought to have evolved from defective phages owing to the high sequence homology and the presence of a lysis cassette in the pyocin gene cluster [32, 33]. The R-type pyocins share a high similarity to the contractile, nonflexible tail structures of the P2 phages, whereas the F-type pyocins are related to the flexible and noncontractile λ phage tail [32]. As a final example of phage proteins being absconded for alternative uses by other organisms, Photorhabdus virulence cassettes (PVCs) have been found in the insect pathogen *Photorhabdus* and in the amber disease-associated plasmid (pADAP) of *Serratia entomophila*. PVCs contain phage-derived toxin genes that inhibit the larvae of beetles from feeding [27]. Apart from gene products that confer insect larvae with antifeeding activity, PVCs also encode a bacteriocin that is structurally similar to the R-type pyocin [27]. This shows how widespread the phage genes are and that these genes are likely just a fraction of the ones that get fixed in the host chromosome or their extra chromosome elements.
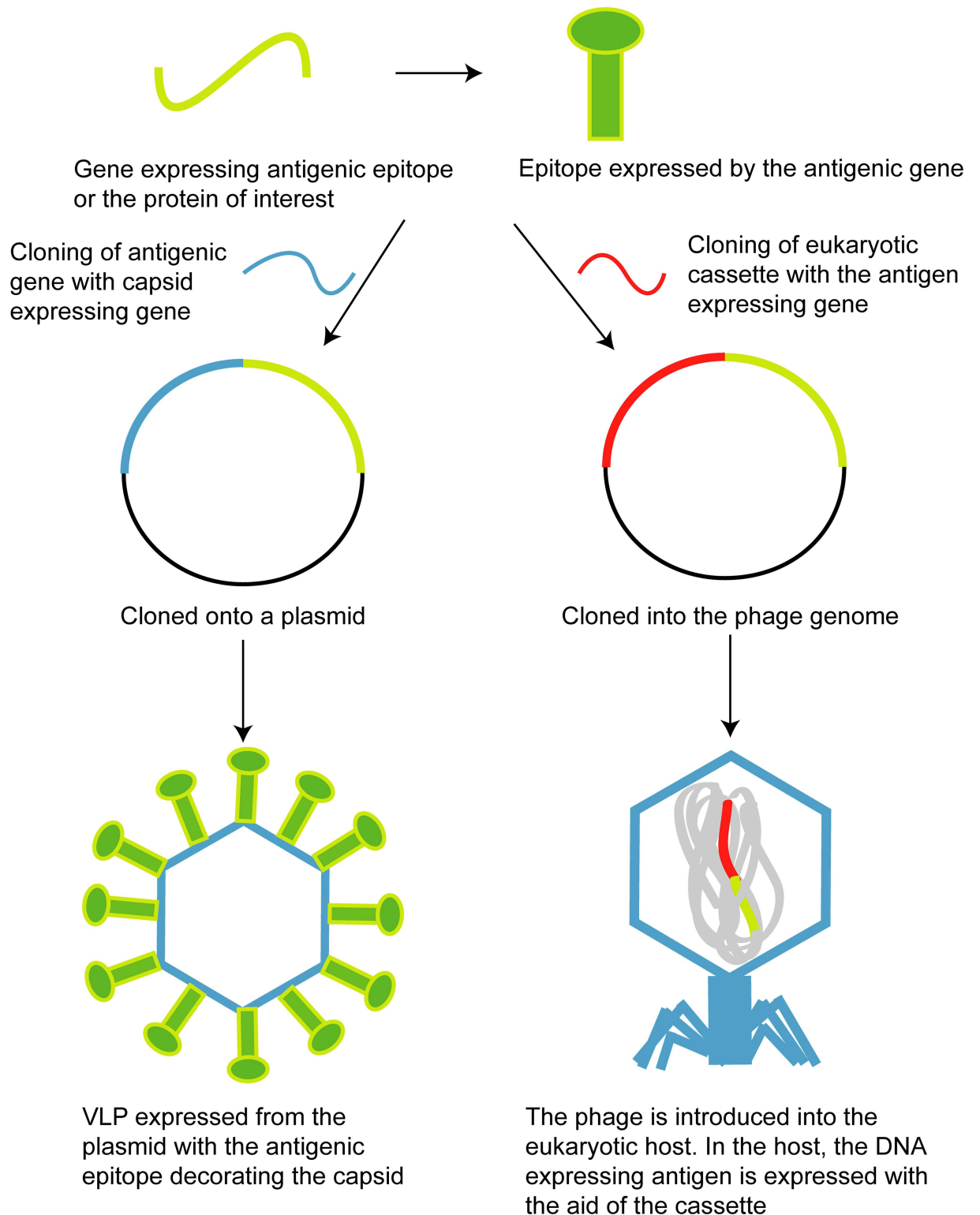
Gene expressing antigenic epitope or the protein of interest

Epitope expressed by the antigenic gene

Cloning of antigenic gene with capsid expressing gene

Cloning of eukaryotic cassette with the antigen expressing gene

Cloned onto a plasmid

Cloned into the phage genome

VLP expressed from the plasmid with the antigenic epitope decorating the capsid

The phage is introduced into the eukaryotic host. In the host, the DNA expressing antigen is expressed with the aid of the cassette

**Fig 2. Phage use in vaccine design and delivery.** The left panel shows the use of a VLP to express an antigenic gene product on its capsid. The right panel shows an antigenic gene being cloned into the phage genome and its subsequent use as a DNA vaccine carrier. VLP, virus-like particle.

https://doi.org/10.1371/journal.ppat.1006971.g002

## Conclusion

Time and again, bacteriophages have helped further scientific progress in biology. Evolutionary links can be seen between viruses such as phage P22 and HSV-1 and also between bacterial viruses and their hosts. This makes bacteriophages ideal as model systems to understand mechanisms such as viral assembly and evolution of pathogenic bacteria. Certainly, further studies will find even more links between viruses infecting hosts from different domains and reveal new relationships between viruses and their host cells.

## References

1. Keen EC. A century of phage research: bacteriophages and the shaping of modern biology. Bioessays. 2015; 37(1):6–9. https://doi.org/10.1002/bies.201400152 PMID: 25521633.

2. Crick FH, Barnett L, Brenner S, Watts-Tobin RJ. General nature of the genetic code for proteins. Nature. 1961; 192:1227–32. PMID: 13882203.

3. Brenner S, Jacob F, Meselson M. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. Nature. 1961; 190:576–81. PMID: 20446365.

4. Taylor K, Hradecna Z, Szybalski W. Asymmetric distribution of the transcribing regions on the complementary strands of coliphage lambda DNA. Proc Natl Acad Sci U S A. 1967; 57(6):1618–25. PMID: 5231395.

5. Georgopoulos C. Toothpicks, serendipity and the emergence of the Escherichia coli DnaK (Hsp70) and GroEL (Hsp60) chaperone machines. Genetics. 2006; 174(4):1699–707. https://doi.org/10.1534/genetics.104.68262 PMID: 17182732.

6. Teschke CM, King J, Prevelige PE Jr. Inhibition of viral capsid assembly by 1,1'-bi(4-anilinonaphthalene-5-sulfonic acid). Biochemistry. 1993; 32(40):10658–65. PMID: 8399211.

7. Abrescia NG, Bamford DH, Grimes JM, Stuart DI. Structure unifies the viral universe. Annu Rev Biochem. 2012; 81:795–822. https://doi.org/10.1146/annurev-biochem-060910-095130 PMID: 22482909.

8. Suhanovsky MM, Teschke CM. Nature's favorite building block: Deciphering folding and capsid assembly of proteins with the HK97-fold. Virology. 2015; 479–480:487–97. https://doi.org/10.1016/j.virol.2015.02.055 PMID: 25864106.

9. Hua J, Huet A, Lopez CA, Toropova K, Pope WH, Duda RL, et al. Capsids and Genomes of Jumbo-Sized Bacteriophages Reveal the Evolutionary Reach of the HK97 Fold. MBio. 2017; 8(5):e01579–17. https://doi.org/10.1128/mBio.01579-17 PMID: 29042498.

10. Brown JC, Newcomb WW. Herpesvirus capsid assembly: insights from structural analysis. Curr Opin Virol. 2011; 1(2):142–9. https://doi.org/10.1016/j.coviro.2011.06.003 PMID: 21927635.

11. Fokine A, Rossmann MG. Common Evolutionary Origin of Procapsid Proteases, Phage Tail Tubes, and Tubes of Bacterial Type VI Secretion Systems. Structure. 2016; 24(11):1928–35. https://doi.org/10.1016/j.str.2016.08.013 PMID: 27667692.

12. Cheng H, Shen N, Pei J, Grishin NV. Double-stranded DNA bacteriophage prohead protease is homologous to herpesvirus protease. Protein Sci. 2004; 13(8):2260–9. https://doi.org/10.1110/ps.04726004 PMID: 15273316.

13. Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE. Topologically linked protein rings in the bacteriophage HK97 capsid. Science. 2000; 289(5487):2129–33. PMID: 11000116.

14. Teschke CM, Parent KN. 'Let the phage do the work': using the phage P22 coat protein structures as a framework to understand its folding and assembly mutants. Virology. 2010; 401(2):119–30. https://doi.org/10.1016/j.virol.2010.02.017 PMID: 20236676.

15. Hryc CF, Chen DH, Afonine PV, Jakana J, Wang Z, Haase-Pettingell C, et al. Accurate model annotation of a near-atomic resolution cryo-EM map. Proc Natl Acad Sci U S A. 2017; 114(12):3103–8. https://doi.org/10.1073/pnas.1621152114 PMID: 28270620.

16. D'Lima NG, Teschke CM. A Molecular Staple: D-Loops in the I Domain of Bacteriophage P22 Coat Protein Make Important Intercapsomer Contacts Required for Procapsid Assembly. J Virol. 2015; 89(20):10569–79. https://doi.org/10.1128/JVI.01629-15 PMID: 26269173.

17. Hasek ML, Maurer JB, Hendrix RW, Duda RL. Flexible Connectors between Capsomer Subunits that Regulate Capsid Assembly. J Mol Biol. 2017; 429(16):2474–89. https://doi.org/10.1016/j.jmb.2017.07.002 PMID: 28705762.

18. Fokine A, Leiman PG, Shneider MM, Ahvazi B, Boeshans KM, Steven AC, et al. Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common

ancestry. Proc Natl Acad Sci U S A. 2005; 102(20):7163–8. https://doi.org/10.1073/pnas.0502164102 PMID: 15878991.

19. Huet A, Makhov AM, Huffman JB, Vos M, Homa FL, Conway JF. Extensive subunit contacts underpin herpesvirus capsid stability and interior-to-exterior allostery. Nat Struct Mol Biol. 2016; 23(6):531–9. https://doi.org/10.1038/nsmb.3212 PMID: 27111889.

20. van Houten NE, Zwick MB, Menendez A, Scott JK. Filamentous phage as an immunogenic carrier to elicit focused antibody responses against a synthetic peptide. Vaccine. 2006; 24(19):4188–200. https://doi.org/10.1016/j.vaccine.2006.01.001 PMID: 16488517.

21. Gaubin M, Fanutti C, Mishal Z, Durrbach A, De Berardinis P, Sartorius R, et al. Processing of filamentous bacteriophage virions in antigen-presenting cells targets both HLA class I and class II peptide loading compartments. DNA Cell Biol. 2003; 22(1):11–8. https://doi.org/10.1089/104454903321112451 PMID: 12590733.

22. Iwagami Y, Casulli S, Nagaoka K, Kim M, Carlson RI, Ogawa K, et al. Lambda phage-based vaccine induces antitumor immunity in hepatocellular carcinoma. Heliyon. 2017; 3(9):e00407. https://doi.org/10.1016/j.heliyon.2017.e00407 PMID: 28971150.

23. Schwarz B, Douglas T. Development of virus-like particles for diagnostic and prophylactic biomedical applications. Wiley Interdiscip Rev Nanomed Nanobiotechnol. 2015; 7(5):722–35. https://doi.org/10.1002/wnan.1336 PMID: 25677105.

24. Frietze KM, Peabody DS, Chackerian B. Engineering virus-like particles as vaccine platforms. Curr Opin Virol. 2016; 18:44–9. https://doi.org/10.1016/j.coviro.2016.03.001 PMID: 27039982.

25. Huang X, Wang X, Zhang J, Xia N, Zhao Q. Escherichia coli-derived virus-like particles in vaccine development. NPJ Vaccines. 2017; 2:3. https://doi.org/10.1038/s41541-017-0006-8 PMID: 29263864.

26. Clark JR, Bartley K, Jepson CD, Craik V, March JB. Comparison of a bacteriophage-delivered DNA vaccine and a commercially available recombinant protein vaccine against hepatitis B. FEMS Immunol Med Microbiol. 2011; 61(2):197–204. https://doi.org/10.1111/j.1574-695X.2010.00763.x PMID: 21204995.

27. Taylor NMI, van Raaij MJ, Leiman PG. Contractile injection systems of bacteriophages and related systems. Mol Microbiol. 2018; 108(1):6–15. https://doi.org/10.1111/mmi.13921 PMID: 29405518.

28. Brunet YR, Henin J, Celia H, Cascales E. Type VI secretion and bacteriophage tail tubes share a common assembly pathway. EMBO Rep. 2014; 15(3):315–21. https://doi.org/10.1002/embr.201337936 PMID: 24488256.

29. Pell LG, Kanelis V, Donaldson LW, Howell PL, Davidson AR. The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. Proc Natl Acad Sci U S A. 2009; 106(11):4160–5. https://doi.org/10.1073/pnas.0900044106 PMID: 19251647.

30. Cascales E, Cambillau C. Structural biology of type VI secretion systems. Philos Trans R Soc Lond B Biol Sci. 2012; 367(1592):1102–11. https://doi.org/10.1098/rstb.2011.0209 PMID: 22411981.

31. Zoued A, Brunet YR, Durand E, Aschtgen MS, Logger L, Douzi B, et al. Architecture and assembly of the Type VI secretion system. Biochim Biophys Acta. 2014; 1843(8):1664–73. https://doi.org/10.1016/j.bbamcr.2014.03.018 PMID: 24681160.

32. Nakayama K, Takashima K, Ishihara H, Shinomiya T, Kageyama M, Kanaya S, et al. The R-type pyocin of Pseudomonas aeruginosa is related to P2 phage, and the F-type is related to lambda phage. Mol Microbiol. 2000; 38(2):213–31. PMID: 11069649.

33. Michel-Briand Y, Baysse C. The pyocins of Pseudomonas aeruginosa. Biochimie. 2002; 84(5–6):499–510. PMID: 12423794.