# Bacteriophage module reshuffling results in adaptive host range as exemplified by the baseplate model of listerial phage A118

Christian Cambillau

Virology 484 (2015) 86–92

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Brief Communication

# Bacteriophage module reshuffling results in adaptive host range as exemplified by the baseplate model of listerial phage A118

Christian Cambillau [a,b,*]

[a] Architecture et Fonction des Macromolécules Biologiques, UMR 7257 CNRS, France
[b] AFMB, Aix-Marseille University, Campus de Luminy, Case 932, 13288 Marseille Cedex 09, France

ARTICLE INFO

*Article history:*
Received 9 April 2015
Returned to author for revisions
10 May 2015
Accepted 21 May 2015
Available online 15 June 2015

*Keywords:*
Bacteriophages
*Lactococcus lactis*
*Listeria monocytogenes*
Host adaptation
Baseplate
Receptor binding protein
Domain reshuffling

ABSTRACT

Each phage infects its specific bacterial host strain through highly specific interactions between the baseplate-associated receptor binding protein (RBP) at the tip of the phage tail and the receptor at the host surface. Baseplates incorporate structural core modules, Dit and Tal, largely conserved among phages, and peripheral modules anchoring the RBPs. Exploiting structural information from the HHpred program and EM data from the Bielmann et al. (2015) paper, a molecular model of the A118 phage baseplate was generated from different building blocks. This model implies the occurrence of baseplate module reshuffling and suggests that listerial phage A118 may have been derived from lactococcal phage TP901-1 through host species exchange. With the increase of available viral module structures, modelling phage baseplates will become easier and more reliant, and will provide insightful information on the nature of the phage host receptor and its mode of recognition.



## Introduction

Viruses that infect bacteria, commonly designated as bacteriophages or phages, can adopt a wide diversity of morphologies, although a large majority, the *Caudovirales*, possess an icosahedral capsid and a tail. *Caudovirales* are further classified as *Myoviridae*, which possess a contractile tail, *Podoviridae*, with a short tail, and *Siphoviridae*, exhibiting a long non-contractile tail (Ackermann, 2009). *Podoviridae*, due to their rigid structure, are among the best characterized phages from a structural perspective, the most prominent representatives being phages T7 (Hu et al., 2013), P22 (Lander et al., 2006), Phi29 (Simpson et al., 2000). The structure and infection mechanism of phage T4, the phage model of *Myoviridae*, have been extensively studied, yielding a detailed picture of its infection mechanism (Kanamaru et al., 2002; Kostyuchenko et al., 2003, 2005). In contrast, until recently, the flexible tail of *Siphoviridae* has hindered full phage structural studies. Structural investigations on phages infecting gram-positive bacteria have been conducted on phage SPP1 (Lhuillier et al., 2009; Plisson et al., 2007), that infects *Bacillus subtilis*, and on several phages that infect the dairy bacterium *Lactococcus lactis*, such as phages p2, TP901-1, Tuc2009 and 1358 (for a review see (Spinelli et al., 2014b)).

The infection process of viruses involves an initial interaction between the viral host recognition proteins and the host cell surface receptor, a protein, a polysaccharide, or both. Phage T5 uses the FhuA porin, an iron-importing membrane protein, as a receptor (Flayhan et al., 2012), while phage SPP1 attaches to YueB, a trans-membrane component of the type VII secretion system, as a receptor (Sao-Jose et al., 2006). Both bind to their receptor with sub-nanomolar Kds. Phage SPP1 (Plisson et al., 2007) as well as coliphages Lambda (Wang et al., 2000) and T5 (Boulanger et al., 2008) possess a straight spike at the distal end of the tail formed by the antireceptor and ancillary proteins. In contrast, some other phages, such as those infecting *L. lactis*, possess a large oligomeric structure at the tip of their tail containing, among other ORFs, the receptor binding protein (Spinelli et al., 2014b). This structure recognizes a saccharidic "pellicle" polymer on the host cell surface (Farenc et al., 2014; Mahony et al., 2012, 2013). The structure of several RBPs and baseplates from various lactococcal phages has been solved (Desmyter et al., 2013; Farenc et al., 2014; Ricagno et al., 2006; Sciara et al., 2010; Spinelli et al., 2006a, 2006b; Veesler et al., 2012). Among other findings, the lactococcal RBPs share a modular structure and a saccharide-binding site located at the C-terminal end that is critical for host binding. Furthermore, the baseplates structures and the infection studies of lactococcal phages p2, TP901-1 and 1358 have revealed that while some

*Abbreviations:* RBP, Receptor binding protein; Dit, Distal Tail protein; Tal, Tail-Associated Lysin; MCP, Major Capsid Protein; MTP, Major Tail Protein; T6SS, Type VI Secretion System

* Corresponding author at: Architecture et Fonction des Macromolécules Biologiques, UMR 7257 CNRS, France. Fax: +33 491 266 720.
*E-mail address:* ccambillau@gmail.com

http://dx.doi.org/10.1016/j.virol.2015.05.015
0042-6822/

phages (e.g. the 936 group) require activation by $Ca^{2+}$, leading to a large conformational change of the baseplate), other phages are "ready to infect" (e.g. the P335 group).

Although lactococcal phages are the only Siphoviridae for which the basplate has been structurally studied, low resolution studies reveal that a large number of siphophages possess a baseplte: listerial phages (P35, P40, A118 and many others; Klumpp and Loessner, 2013), staphilococcal phages (e.g. Phi11 Xia et al., 2011), Lactobacillus phages (Casey et al., 2014), etc.

In a recent paper, Bielmann et al. identified the RBPs from siphophages A118 and P35, infecting the Gram-positive bacterium *Listeria monocytogenes*, as well as their receptor specificity (Bielmann et al., 2015). These authors also performed a topological study of the A118 baseplate using EM-immuno labeling, making it possible to identify and localize the baseplate proteins. Using the program HHpred (Hildebrand et al., 2009; Soding et al., 2005) they identified similarities between the A118 structural proteins and the baseplate components of other phages (Bielmann et al., 2015).

Here, following full exploitation of the information from HHpred and from the Bielmann et al. paper (Bielmann et al., 2015), a molecular model of the A118 phage baseplate was generated from the individual building blocks. This model reveals the existence of baseplate module reshuffling, and suggests that listerial phage A118 has been derived from lactococcal phage TP901-1 following a host alteration event. Although this kind of topological modelling does not provide details of the structures, the present approach can been seen as an illustration of the power of structural bioinformatics, whose outcome makes it possible to visualize structural similarities in an easier way, compared to lists of sequences similarities, and to propose experimentally testable hypothesis.

## Results

A HHpred analysis was performed with the components of A118 baseplate, proteins pg17 through to pg21. Protein gp17 is identified as Dit with a calculated probability of 100% (Fig. 1a). Dit is a widespread component of siphophages and the hub on which the rest of the baseplate is assembled below the last tail MTP ring. It forms an interlaced hexameric ring exposing at its periphery a variable domain (Fig. 2a), the latter being a galectin-like domain in phages p2, TP901-1, SPP1 (Sciara et al., 2010; Veesler et al., 2010b, 2012) and other phages, or an OB-fold domain as reported for phage T5 (Flayhan et al., 2014).

A crystal structure of a protein from Listeria has been deposited in the PDB, that shares 95.5% sequence identity with A118gp18 (Fig. 2b). This protein was the first in the list of hits reported by HHpred (Fig. 1b), identified as members of the Tal family. The Tal family comprises proteins belonging to *Myoviridae* (e.g. phage T4 gp27 (Kanamaru et al., 2002)), *Siphoviridae* (lactococcal phages p2, TP901-1 (Sciara et al., 2010; Veesler et al., 2010b, 2012), phage SPP1 (Veesler et al., 2010b)), and also to the type VI secretion system (T6SS), the VgrG1 protein (Leiman et al., 2009). Tal proteins can be subdivided into two classes: the first class encompasses short Tal proteins that exist as a trimeric globular structure of ~350 residues per monomer, e.g. those found in lactococcal phages p2 (Sciara et al., 2010), 1358 (Spinelli et al., 2014a), and some members of the P335 family (Stockdale et al., 2013). In the second class, Tals represent large proteins up to 1100 residues, and consist of a N-terminal trimeric domain, similar to short Tal's, plus a long C-terminal extension, seen in EM as a fiber, and bearing a cell hydrolyzing enzyme (Plisson et al., 2007; Stockdale et al., 2013). Noteworthy, the A118 Tal is a member of the first class, i.e. the short Tal's, thus devoid of a fiber extension, as can be discerned on electron micrographs of A118 (Bielmann et al., 2015). Tal is known to exist in an open or closed form (Sciara et al., 2010), the former corresponding to the infection prone baseplate (Sciara et al., 2010). Dit and Tal form a complex at the tip of the phage tail, and constitute the central core of the baseplate (Goulet et al., 2011; Sciara et al., 2010).

The most interesting part (by far) of phage A118 baseplate component is gp20. HHpred reports two hits, with overlapping sequences, from phage TP901-1 BppU and BppL (the RBP) (Fig. 1c). Starting from the N-terminus, the two first hits cover the full length BppU and its isolated N-terminal β-stranded domain (Fig. 1c, blue box; Fig. 2c). In BppU, this domain is followed by a helical stretch projected radially (Fig. 1c, magenta box; Fig. 2c), followed by a kink and another helical stretch roughly perpendicular (Fig. 1c, gray box; Fig. 2c). In full-length BppU, this helix is followed by a C-terminal β-stranded domain accommodating a BppL (RBP) trimer. In phage A118, this helix extends, encompassing the helical domain of BppL (Fig. 1c, orange box; Fig. 2c). This domain is followed by a β-stranded interlaced domain (Fig. 1c, gray box; Fig. 2c) and by the receptor-binding domain (termed head domain in phages p2 and TP901-1; Fig. 1c, green box; Fig. 2c). The overall interpretation from this HHpred comparative analysis is that gp20 represents a fusion between functional analogs of BppU and BppL (Fig. 2d and e). Compared with the TP901-1 BppU/BppL arrangement ($3 \times$ BppU$+3 \times 3$ BppL; Fig. 2f) where each BppU trimer accommodates three RBP trimers, the gp20 configuration will only allow (because it is a fusion) the association of a single RBP head trimer (Fig. 2e).

Due to the similarity between gp17 and Dit proteins and gp19 N-terminal domain with the N-terminal β-sandwich domain of TP901-1 BppU, the assembly of gp17 and gp19 should resemble that of the similar component in the TP901-1 baseplate (Veesler et al., 2012). The same should apply to the gp17 and gp18 (Tal) assembly although this structure is not documented in TP901-1, but only in the SPP1 (Goulet et al., 2011) and the p2 phages (Sciara et al., 2010). As a result, the three components, gp17, gp18 and gp20 are predicted to assemble as a complete baseplate of $6 \times$ gp17 (Dit), $3 \times$ gp18 (Tal) and $6 \times 3$ gp20 (BppU/BppL fusion) using TP901-1 baseplates as a template (Fig. 2g). The main topological differences between TP901-1 and A118 baseplate are (i) the fused BppU/BppL configuration of gp20 results into to a single RBP trimer instead of three, (ii) a shortening of the horizontal arm of the BppU, as can be deduced from the HHPRED figure and the width of the baseplate as seen by EM (200 Å for A118 vs. 280 Å for TP901-1) and (iii) a longer span of the vertical helix bundle (Figs. 3 and 4). A topological model of the full gp20 was constructed using the molecular graphics program Coot (Emsley et al., 2010) by deleting the final part of TP901-1 BppU, before the β-stranded C-terminal domain, and grafting the helical bundle of TP901-1 BppL, the RBP, after deleting the extended 17 amino-acids long N-terminal stretch. Because the baseplate looked less large on the EM view, the length of the helical bundle between BppU N-terminal domain and the kink was shortened (24 residues shorter). To adhere with the A118 baseplate dimensions, as seen on the low resolution EM view, the vertical helical bundle, between the kink and the receptor binding domain, had to be elongated by duplicating part of the helix bundle. The total length of the gp20 helix bundle, including the kink, was adjusted to the HHpred prediction as well as to the low EM view of the baseplate.

Two other proteins, gp19 and gp21, have been shown to be part of the A118 baseplate (Bielmann et al., 2015). HHpred analysis reports that of the 339 amino acids of gp19, ~220 are related to pectinases, esterases that hydrolyze pectin. This suggests that gp19 encompasses an enzymatic activity that hydrolyzes components of the bacterial wall using its two catalytic aspartic-acid residues, either saccharidic or peptidic linkages. The EM immuno labeling of gp19 is consistent with two positions: one at the upper end of the


| No | Hit | Prob | E-value | P-value | Score | SS | Cols | Query HMM | Template HMM |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2x8k_A Hypothetical protein 19 | 100.0 | 5.9E-40 | 1.7E-44 | 275.1 | 21.8 | 230 | 7-270 | 12-252 (252) |
| 2 | 4div_S ORF46; distal tail prot | 100.0 | 3.1E-35 | 9E-40 | 246.7 | 18.6 | 226 | 1-270 | 18-253 (253) |

| No | Hit | Prob | E-value | P-value | Score | SS | Cols | Query HMM | Template HMM |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 3gs9_A Protein GP18; NP_465809 | 100.0 | 9.9E-72 | 2.9E-76 | 524.7 | 33.4 | 341 | 1-341 | 2-342 (342) |
| 2 | 3d37_A Tail protein, 43 kDa; s | 99.2 | 1.1E-08 | 3.2E-13 | 97.0 | 26.7 | 289 | 23-320 | 22-336 (381) |
| 3 | 3cdd_A Prophage MUSO2, 43 kDa | 98.9 | 5.4E-07 | 1.6E-11 | 84.6 | 24.8 | 290 | 23-320 | 19-339 (361) |
| 4 | 1wru_A 43 kDa tail protein; ba | 98.8 | 3E-07 | 8.7E-12 | 87.1 | 20.2 | 304 | 1-320 | 1-334 (379) |
| 5 | 4mtk_A VGRG1; beta-barrel, OB- | 95.5 | 2.7 | 7.8E-05 | 42.4 | 22.7 | 269 | 25-308 | 22-323 (643) |

| No | Hit | Prob | E-value | P-value | Score | SS | Cols | Query HMM | Template HMM |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 4div_A ORF48; distal tail prot | 99.9 | 6.7E-22 | 1.9E-26 | 189.7 | 17.5 | 192 | 8-209 | 1-218 (299) |
| 2 | 3uh8_A ORF48; all beta protein | 99.8 | 7.9E-18 | 2.3E-22 | 144.0 | 14.7 | 121 | 8-133 | 1-128 (128) |
| 3 | 2f0c_A Phage TP901-1 ORF49 (BP | 97.4 | 0.00021 | 6E-09 | 64.1 | 5.0 | 142 | 158-341 | 30-179 (191) |
| 4 | 3d8m_A BPP, baseplate protein, | 96.7 | 0.0081 | 2.3E-07 | 52.9 | 9.1 | 147 | 157-343 | 1-153 (165) |

| No | Hit | Prob | E-value | P-value | Score | SS | Cols | Query HMM | Template HMM |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1gq8_A Pectinesterase; hydrola | 100.0 | 2.4E-38 | 6.9E-43 | 305.4 | 18.0 | 205 | 116-339 | 1-224 (319) |
| 2 | 1xg2_A Pectinesterase 1; prote | 100.0 | 2.5E-37 | 7.1E-42 | 298.2 | 17.4 | 201 | 120-339 | 1-220 (317) |
| 3 | 2nsp_A Pectinesterase A; micha | 100.0 | 3.4E-33 | 9.8E-38 | 272.0 | 17.0 | 186 | 119-323 | 1-232 (342) |
| 4 | 3grh_A YBHC, acyl-COA thioeste | 100.0 | 9.5E-33 | 2.7E-37 | 275.0 | 16.3 | 186 | 119-320 | 72-330 (422) |
| 5 | 4pmh_A Pectinesterase; beta-ba | 100.0 | 2.2E-32 | 6.3E-37 | 268.4 | 13.9 | 196 | 119-340 | 37-290 (366) |
| 6 | 3uw0_A Pectinesterase; right-h | 100.0 | 5.8E-32 | 1.7E-36 | 265.4 | 16.2 | 189 | 118-325 | 27-259 (364) |

**Fig. 1.** Analysis of phage A118 baseplate proteins with HHpred. HHpred (Hildebrand et al., 2009) outputs display (i) a bar graph summarizing the positions of the database matches with more than 95% probability and the PDB identifier inside the bar and (ii) a tabular hit list with probabilities, *E*-values, scores, and match regions in query and templates (for more details see http://toolkit.tuebingen.mpg.de/hhpred/help_results). (A) Results of the gp17 analysis. The two hits are the Dit proteins of phages SPP1 and TP901-1, respectively. (B) Results of the gp18 analysis. The first hit is the structure of a gp18 variant obtained by a structural genomics consortium (no publication). (C) Results of the gp20 analysis. All the HHpred hits returned are from the TP901-1 phage proteins BppU and BppL (Veesler et al., 2012). The blue box identifies the N-terminal β-barrel; the magenta box identifies the BppU short horizontal helix and helical turn. The gray box identifies the vertical helix of BppU and also the helical segment of BppL. The yellow box identifies the interlaced β-helix of BppL, and the green box identifies the β-barrel "head" domain of BppL (the receptor binding domain). (D) Results of the gp19 analysis. The HHpred hits returned esterase templates and cover the ~116–339 segment. The 115 first residues are of unknown structure and may represent the phage/enzyme interface.

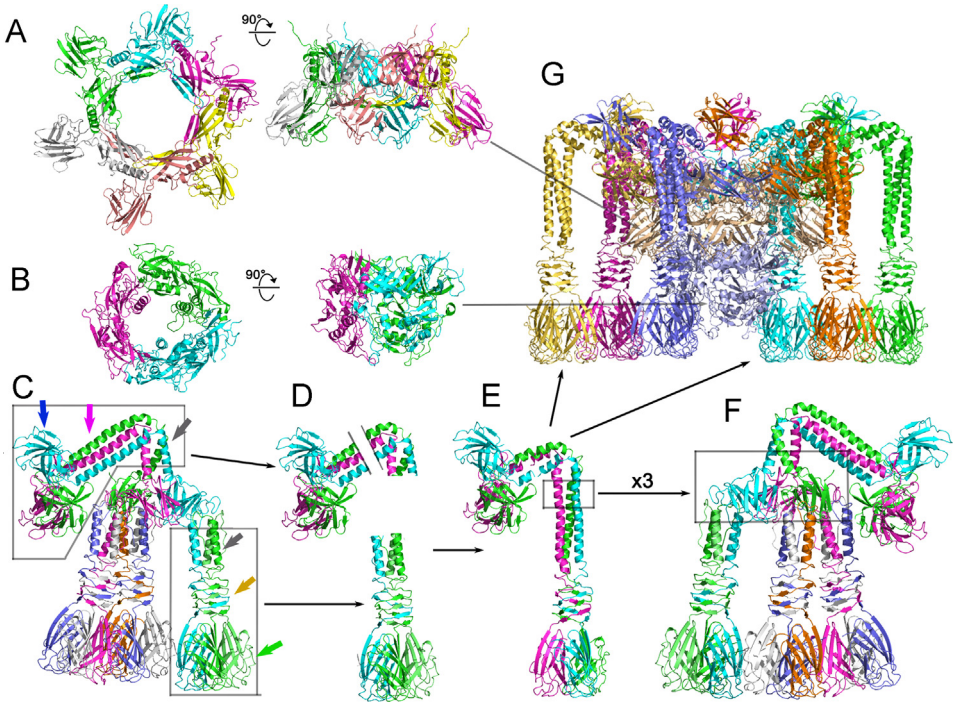


**Fig. 2.** Model building of phage A118 baseplate. (A) The model of gp17 (Dit) hexamer (4DIV). (B) The model of gp18 (Tal) trimer (3GS9). (C) The 3xBppU/9xBppL tripod from phage TP901-1; the parts used to model gp20 are boxed; the colour of the arrows refer to the boxes in Fig. 1c. (D) The extracted parts of the tripod used to model gp20. (E) The gp20 trimer. (F) The TP901-1 3xBppU/9xBppL tripod (rotated 180° relative to C) compared to the gp20 trimer, emphasizing the amplified module of the BppU C-terminal domain (boxes). (G) The phage A118 baseplate core, still lacking gp19 and gp21. Figure made with Pymol.

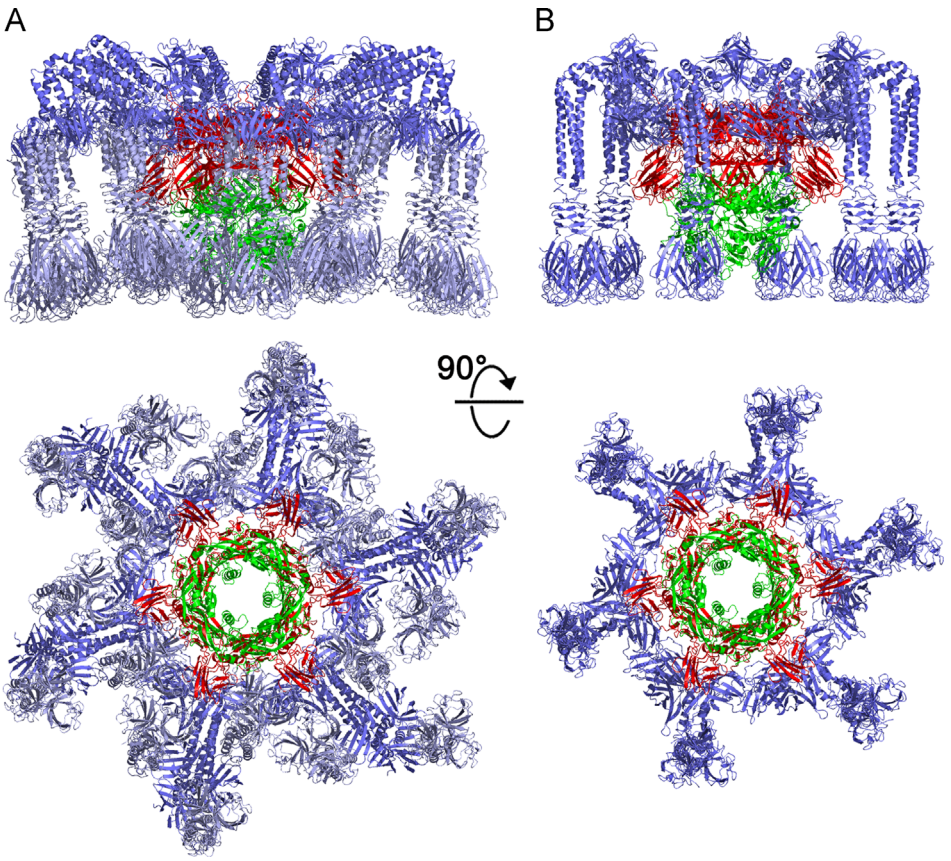


**Fig. 3.** Comparison of phages T0901-1 and A118 baseplates. (A) The phage TP901-1 baseplate (PDB 4D96) and Tal (3GS9). The size is 280 Å wide, 150 Å high. Colors are: Dit, red, Tal, green, BppU, blue; BppL (RBP), light blue. (B) The model phage A118 baseplate core. The size is 200 Å wide, 150 Å high. Colors are: gp17 (Dit), red, gp18 (Tal), green; gp20 (RBP), blue. Figure made with Pymol.

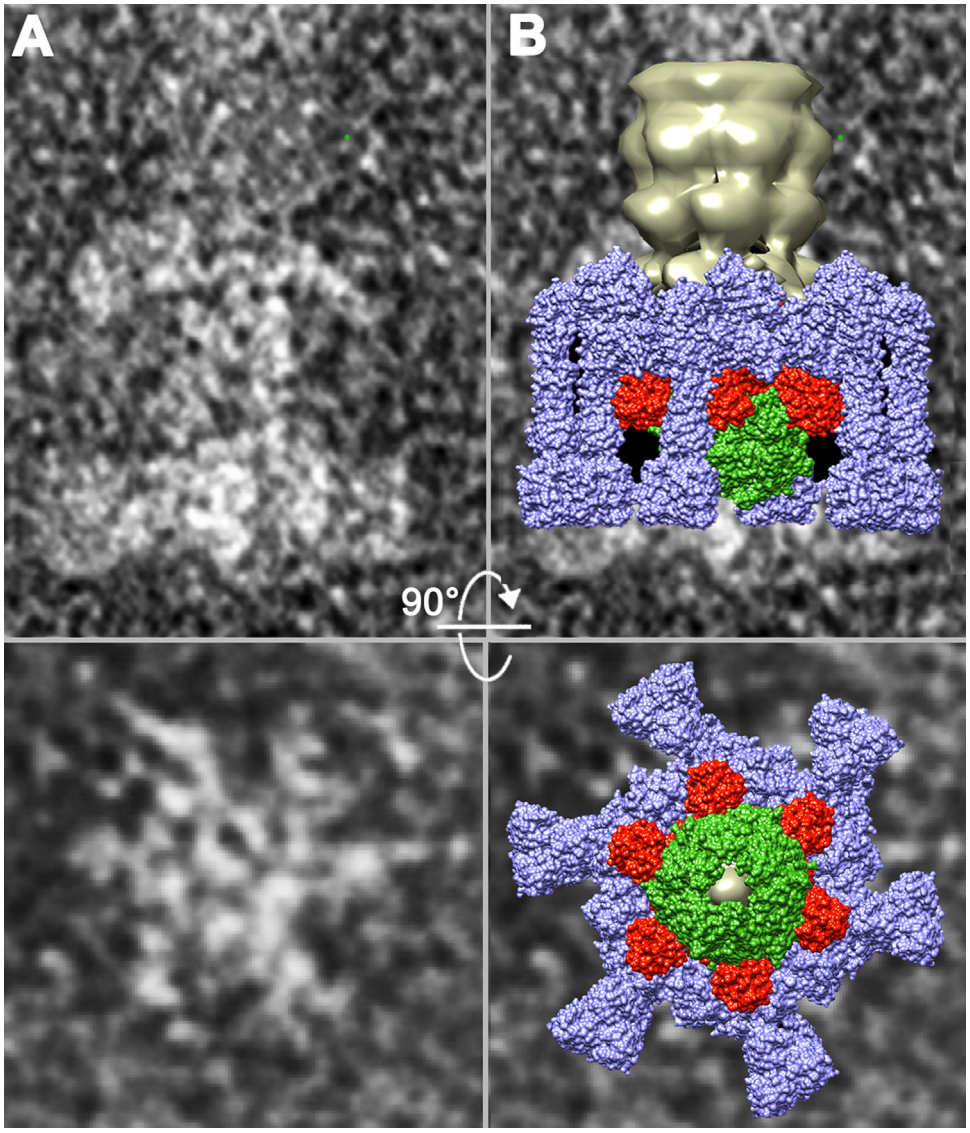


**Fig. 4.** Fitting of the phage A118 baseplate core into an EM micrograph of its tail tip (Bielmann et al., 2015). (A) An EM micrograph of phage A118 tail tip. (B) The model of phage A118 tail and baseplate core superposed to the previous EM picture. Colors are: gp17 (Dit), red, gp18 (Tal), green; gp20 (RBP), blue. Figure made with Chimera (Pettersen et al., 2004).

baseplate, the second being at its lower end (Bielmann et al., 2015). The lower position would imply that gp19 is bound to the tip of Tal (gp18) and/or between the gp20 head domains (C-termini), consistent with its putative role of cell wall-degrading enzyme. The gp21 component has no structural relatives as reported by HHpred. Secondary structure prediction with PsiPred indicated that it essentially consists of a long helix. EM immuno labeling indicated that gp21 is located at the periphery of the A118 baseplate, suggesting that the long helix might run parallel to the gp20 helix bundle. With the current knowledge, the precise positions of gp19 and gp21 cannot be assigned, and would require structural dermination of the baseplate structure.

## Discussion

A topological model of listerial phage A118 has been obtained, guided mainly by structural similarity detection based on HHpred and EM observations reported by Bielmann et al. (Bielmann et al., 2015). The similarities reported by Hhpred were significant enough to attempt integration of molecular data. Indeed, the limits of this model should be considered carefully: it does not provide information (i) on the inter-domain contacts, (ii) on the exact length of the gp20 helices, nor (iii) on the local structure of the receptor binding site. What can be safely inferred from this model is the overall three-dimensional structure of the baseplate and its components. We can confidently assert that A118 gp20 originates from a genetic reshuffling of the individual TP901-1 BppU and BppL-encoding genes, and together with the gp18 (Dit) defines the overall shape of the baseplate.

Indeed, it has been well established previously that molecular modules are widely shared between different phages and even viruses (Veesler and Cambillau, 2011). The most striking example is their Major Capsid Proteins (MCP), whose fold is conserved in most capsid-containing phages (Veesler and Cambillau, 2011) and by other viral capsids, pointing towards an ancient common ancestor (Bamford, 2003). The Major Tail Proteins (MTP) of most siphophages share also a common fold (Pell et al., 2013), as well as the tail tube protein (Hcp) of T6SS (Mougous et al., 2006). The folds of portal proteins, head-to-tail connectors, tail terminators are also shared by a large number of phages (Veesler and Cambillau, 2011).

The host-recognition devices, RBPs and baseplates, have a distinct functional purpose as each phage needs to recognize its specific host(s). This characteristic is pushed to an extreme in lactococcal phages: hundreds of different phages have been isolated worldwide, yet surprisingly each replicate within a specific and very limited set of *L. lactis* strains (Mahony et al., 2013; Mahony and van Sinderen, 2012). Despite this functional diversity, the jelly-roll fold forming the RBP head domain is conserved in many different phages, lactococcal phages (such as phages p2, bIL170, TP901-1, 1358, etc, reviewed in (Spinelli et al., 2014b)), coliphages (such as phages T4 (Riede, 1987) and PRD1 (Xu et al., 2003)), and even in human adenovirus (Spinelli et al., 2006a; Veesler and Cambillau, 2011). Other modules have been shown to be exchanged among phages: the p2 RBP "shoulder" domain has counterparts in the RBP of lactococcal phage 1358 (Farenc et al., 2014) and in gp22 from SPP1 *B. subtilis* phage (Veesler et al., 2010a).

Besides or together with module exchange and reshuffling, phage host exchange may occur in certain cases, as genome-wide similarities are observed among phages infecting hosts from different species. In this context, the lactococcal phage 1358 structural proteins, including those of the baseplate, were shown to share high sequence similarity with their counterparts in listerial phages P35 and P40 (Dupuis and Moineau, 2010). In particular, while the RBPs of 1358 and P40 share 28% identity within their 200 N-terminal amino acids, which represent the part that interacts with the rest of the phage base plate, their C-terminal ~100 amino acids, which encompass the receptor binding site, are different, each probably adapted to recognize their specific host. Noteworthy, significant similarity has also been reported between phages P35 and p2 baseplate components (Bielmann et al., 2015). The striking structural similarity between TP901-1 and A118 baseplates also suggests that one phage might originate from the other. Since TP901-1 and A118 recognize and attach to cell wall-associated saccharidic structures, represented by the pellicle and teichoic acids, respectively (Bielmann et al., 2015; Farenc et al., 2014; McCabe et al., 2015; Spinelli et al., 2014a), it may be envisaged that the receptor binding site of a phage from one species has mutated to adapt to recognize the cell wall surface saccharide of another species. Therefore, the case reported here may illustrate a case of host swapping from *Lactococcus* to *Listeria*, whereas phages 1358 and P40 may represent an example of an inverse host species exchange from *Listeria* to *Lactococcus*. Noteworthy, interspecies exchanges may be facilitated as *L. lactis* and *Listeria monocytogenes* share in some instances the same environmental niche, i.e. the human gut (Ramaswamy et al., 2007). As a significant outcome, the present model of the phage A118 baseplate suggests that this phage, as phage TP901-1, does not need activation and conformational change of the baseplate for infection, in contrast with phages p2 and 1358.

## Methods

Molecular modelling of phage A118 baseplate was performed using as a starting model the crystal structure of the TP901-1 baseplate ((Veesler et al., 2012); PDB identifier 4D96, previously 4DIV, 4DIW). Model building was performed with Coot (Emsley et al., 2010). Phage A118 gp17 structure was represented directly from the structures of the TP901-1 Dit (extracted from 4D96). Phage A118 gp18 sequence is 95.5% identical to that of 3GS9. Hence the 3GS9 structure was used to represent gp18 (Tal) structure. The 3GS9 trimer was docked "manually" (by rotation-translation of the whole molecule) against the gp17. Phage A118 gp20 structure was built by fusing parts of TP901-1 BppU and BppL. More precisely, the C-terminal domain of BppU was removed (after residue 190) and the horizontal arm was shortened by 28 residues (deletion, Δ146–163). The horizontal helix of BppU was then extended by 15 helical residues. The BppL N-terminus non-helical stretch wad deleted (Δ2–15) and BppL helix 16–30 was placed manually in continuation of the BppU extended vertical helix. During this process, attention was paid to obtain a gp20 model with the proper length and a baseplate model fitting the dimensions of the EM views. Fig. 1 is composed from the HHpred program outputs. Figs. 2 and 3 are made with Pymol. Fig. 4 uses as background an EM view of phage A118 presented in Figure 5 of Bielmann et al. (2015); the EM tail representation (taken from the structure of phage TP901-1 tail (Bebeacua et al., 2013)) and the A118 baseplate view were produced with Chimera (Pettersen et al., 2004).

## Acknowledgments

This work was supported by grants from the Agence Nationale de la Recherche (Grants ANR-11-BSV8-004-01 "Lactophages". Prof. Jochen Klumpp and Prof. Douwe van Sideren are warmly acknowledged for useful discussions.

## References

Ackermann, H.W., 2009. Phage classification and characterization. Methods Mol. Biol. 501, 127–140.
Bamford, D.H., 2003. Do viruses form lineages across different domains of life? Res. Microbiol. 154, 231–236.
Bebeacua, C., Lai, L., Vegge, C.S., Brondsted, L., van Heel, M., Veesler, D., Cambillau, C., 2013. Visualizing a complete siphoviridae member by single-particle electron microscopy: the structure of lactococcal phage TP901-1. J. Virol. 87, 1061–1068.
Bielmann, R., Habann, M., Eugster, M.R., Lurz, R., Calendar, R., Klumpp, J., Loessner, M.J., 2015. Receptor binding proteins of Listeria monocytogenes bacteriophages A118 and P35 recognize serovar-specific teichoic acids. Virology 477, 110–118.
Boulanger, P., Jacquot, P., Plancon, L., Chami, M., Engel, A., Parquet, C., Herbeuval, C., Letellier, L., 2008. Phage T5 straight tail fiber is a multifunctional protein acting as a tape measure and carrying fusogenic and muralytic activities. J. Biol. Chem. 283, 13556–13564.
Casey, E., Mahony, J., O'Connell-Motherway, M., Bottacini, F., Cornelissen, A., Neve, H., Heller, K.J., Noben, J.P., Dal Bello, F., van Sinderen, D., 2014. Molecular characterization of three *Lactobacillus delbrueckii* subsp. bulgaricus phages. Appl. Environ. Microbiol. 80, 5623–5635.
Desmyter, A., Farenc, C., Mahony, J., Spinelli, S., Bebeacua, C., Blangy, S., Veesler, D., van Sinderen, D., Cambillau, C., 2013. Viral infection modulation and neutralization by camelid nanobodies. Proc. Natl. Acad. Sci. USA 110, E1371–1379.
Dupuis, M.E., Moineau, S., 2010. Genome organization and characterization of the virulent lactococcal phage 1358 and its similarities to Listeria phages. Appl. Environ. Microbiol. 76, 1623–1632.
Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501.
Farenc, C., Spinelli, S., Vinogradov, E., Tremblay, D., Blangy, S., Sadovskaya, I., Moineau, S., Cambillau, C., 2014. Molecular insights on the recognition of a *Lactococcus lactis* cell wall pellicle by the phage 1358 receptor binding protein. J. Virol. 88, 7005–7015.
Flayhan, A., Vellieux, F.M., Lurz, R., Maury, O., Contreras-Martel, C., Girard, E., Boulanger, P., Breyton, C., 2014. Crystal Structure of pb9, the distal tail protein of bacteriophage T5: a conserved structural motif among all siphophages. J. Virol. 88, 820–828.
Flayhan, A., Wien, F., Paternostre, M., Boulanger, P., Breyton, C., 2012. New insights into pb5, the receptor binding protein of bacteriophage T5, and its interaction with its *Escherichia coli* receptor FhuA. Biochimie 94, 1982–1989.
Goulet, A., Lai-Kee-Him, J., Veesler, D., Auzat, I., Robin, G., Shepherd, D.A., Ashcroft, A.E., Richard, E., Lichiere, J., Tavares, P., Cambillau, C., Bron, P., 2011. The opening of the SPP1 bacteriophage tail, a prevalent mechanism in gram-positive-infecting siphophages. J. Biol. Chem. 286, 25397–25405.
Hildebrand, A., Remmert, M., Biegert, A., Soding, J., 2009. Fast and accurate automatic structure prediction with HHpred. Proteins 77 (Suppl. 9), 128–132.
Hu, B., Margolin, W., Molineux, I.J., Liu, J., 2013. The bacteriophage t7 virion undergoes extensive structural remodeling during infection. Science 339, 576–579.
Kanamaru, S., Leiman, P.G., Kostyuchenko, V.A., Chipman, P.R., Mesyanzhinov, V.V., Arisaka, F., Rossmann, M.G., 2002. Structure of the cell-puncturing device of bacteriophage T4. Nature 415, 553–557.
Klumpp, J., Loessner, M.J., 2013. Phages: genomes, evolution, and application. Bacteriophage 3, e26861.

Kostyuchenko, V.A., Chipman, P.R., Leiman, P.G., Arisaka, F., Mesyanzhinov, V.V., Rossmann, M.G., 2005. The tail structure of bacteriophage T4 and its mechanism of contraction. Nat. Struct. Mol. Biol. 12, 810–813.

Kostyuchenko, V.A., Leiman, P.G., Chipman, P.R., Kanamaru, S., van Raaij, M.J., Arisaka, F., Mesyanzhinov, V.V., Rossmann, M.G., 2003. Three-dimensional structure of bacteriophage T4 baseplate. Nat. Struct. Biol. 10, 688–693.

Lander, G.C., Tang, L., Casjens, S.R., Gilcrease, E.B., Prevelige, P., Poliakov, A., Potter, C.S., Carragher, B., Johnson, J.E., 2006. The structure of an infectious P22 virion shows the signal for headful DNA packaging. Science 312, 1791–1795.

Leiman, P.G., Basler, M., Ramagopal, U.A., Bonanno, J.B., Sauder, J.M., Pukatzki, S., Burley, S.K., Almo, S.C., Mekalanos, J.J., 2009. Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. Proc. Natl. Acad. Sci. USA 106, 4154–4159.

Lhuillier, S., Gallopin, M., Gilquin, B., Brasiles, S., Lancelot, N., Letellier, G., Gilles, M., Dethan, G., Orlova, E.V., Couprie, J., Tavares, P., Zinn-Justin, S., 2009. Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating. Proc. Natl. Acad. Sci. USA 106, 8507–8512.

Mahony, J., Ainsworth, S., Stockdale, S., van Sinderen, D., 2012. Phages of lactic acid bacteria: the role of genetics in understanding phage-host interactions and their co-evolutionary processes. Virology 434, 143–150.

Mahony, J., Kot, W., Murphy, J., Ainsworth, S., Neve, H., Hansen, L.H., Heller, K.J., Sorensen, S.J., Hammer, K., Cambillau, C., Vogensen, F.K., van Sinderen, D., 2013. Investigation of the relationship between lactococcal host cell wall polysaccharide genotype and 936 phage receptor binding protein phylogeny. Appl. Environ. Microbiol. 79, 4385–4392.

Mahony, J., van Sinderen, D., 2012. Structural aspects of the interaction of dairy phages with their host bacteria. Viruses 4, 1410–1424.

McCabe, O., Spinelli, S., Farenc, C., Labbe, M., Tremblay, D., Blangy, S., Oscarson, S., Moineau, S., Cambillau, C., 2015. The targeted recognition of *Lactococcus lactis* phages to their polysaccharide receptors. Mol. Microbiol..

Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., Goodman, A.L., Joachimiak, G., Ordonez, C.L., Lory, S., Walz, T., Joachimiak, A., Mekalanos, J.J., 2006. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. Science 312, 1526–1530.

Pell, L.G., Cumby, N., Clark, T.E., Tuite, A., Battaile, K.P., Edwards, A.M., Chirgadze, N.Y., Davidson, A.R., Maxwell, K.L., 2013. A conserved spiral structure for highly diverged phage tail assembly chaperones. J. Mol. Biol. 425, 2436–2449.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Plisson, C., White, H.E., Auzat, I., Zafarani, A., Sao-Jose, C., Lhuillier, S., Tavares, P., Orlova, E.V., 2007. Structure of bacteriophage SPP1 tail reveals trigger for DNA ejection. EMBO J. 26, 3720–3728.

Pymol, The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

Ramaswamy, V., Cresence, V.M., Rejitha, J.S., Lekshmi, M.U., Dharsana, K.S., Prasad, S.P., Vijila, H.M., 2007. Listeria—review of epidemiology and pathogenesis. J. Microbiol. Immunol. Infect. 40, 4–13.

Ricagno, S., Campanacci, V., Blangy, S., Spinelli, S., Tremblay, D., Moineau, S., Tegoni, M., Cambillau, C., 2006. Crystal structure of the receptor-binding protein head domain from *Lactococcus lactis* phage bIL170. J. Virol. 80, 9331–9335.

Riede, I., 1987. Receptor specificity of the short tail fibres (gp12) of T-even type *Escherichia coli* phages. Mol. Gen. Genet. 206, 110–115.

Sao-Jose, C., Lhuillier, S., Lurz, R., Melki, R., Lepault, J., Santos, M.A., Tavares, P., 2006. The ectodomain of the viral receptor YueB forms a fiber that triggers ejection of bacteriophage SPP1 DNA. J. Biol. Chem. 281, 11464–11470.

Sciara, G., Bebeacua, C., Bron, P., Tremblay, D., Ortiz-Lombardia, M., Lichiere, J., van Heel, M., Campanacci, V., Moineau, S., Cambillau, C., 2010. Structure of lactococcal phage p2 baseplate and its mechanism of activation. Proc. Natl. Acad. Sci. USA 107, 6852–6857.

Simpson, A.A., Tao, Y., Leiman, P.G., Badasso, M.O., He, Y., Jardine, P.J., Olson, N.H., Morais, M.C., Grimes, S., Anderson, D.L., Baker, T.S., Rossmann, M.G., 2000. Structure of the bacteriophage phi29 DNA packaging motor. Nature 408, 745–750.

Soding, J., Biegert, A., Lupas, A.N., 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 33, W244–248.

Spinelli, S., Bebeacua, C., Orlov, I., Tremblay, D., Klaholz, B.P., Moineau, S., Cambillau, C., 2014a. Cryo-electron microscopy structure of lactococcal siphophage 1358 virion. J. Virol. 88, 8900–8910.

Spinelli, S., Campanacci, V., Blangy, S., Moineau, S., Tegoni, M., Cambillau, C., 2006a. Modular structure of the receptor binding proteins of *Lactococcus lactis* phages. The RBP structure of the temperate phage TP901-1. J. Biol. Chem. 281, 14256–14262.

Spinelli, S., Desmyter, A., Verrips, C.T., de Haard, H.J., Moineau, S., Cambillau, C., 2006b. Lactococcal bacteriophage p2 receptor-binding protein structure suggests a common ancestor gene with bacterial and mammalian viruses. Nat. Struct. Mol. Biol. 13, 85–89.

Spinelli, S., Veesler, D., Bebeacua, C., Cambillau, C., 2014b. Structures and host-adhesion mechanisms of lactococcal siphophages. Front. Microbiol. 5, 3.

Stockdale, S.R., Mahony, J., Courtin, P., Chapot-Chartier, M.P., van Pijkeren, J.P., Britton, R.A., Neve, H., Heller, K.J., Aideh, B., Vogensen, F.K., van Sinderen, D., 2013. The lactococcal phages Tuc2009 and TP901-1 incorporate two alternate forms of their tail fiber into their virions for infection specialization. J. Biol. Chem. 288, 5581–5590.

Veesler, D., Blangy, S., Spinelli, S., Tavares, P., Campanacci, V., Cambillau, C., 2010a. Crystal structure of *Bacillus subtilis* SPP1 phage gp22 shares fold similarity with a domain of lactococcal phage p2 RBP. Protein Sci. 19, 1439–1443.

Veesler, D., Cambillau, C., 2011. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. Microbiol. Mol. Biol. Rev. 75, 423–433.

Veesler, D., Robin, G., Lichiere, J., Auzat, I., Tavares, P., Bron, P., Campanacci, V., Cambillau, C., 2010b. Crystal structure of bacteriophage SPP1 distal tail protein (gp19.1): a baseplate hub paradigm in gram-positive infecting phages. J. Biol. Chem. 285, 36666–36673.

Veesler, D., Spinelli, S., Mahony, J., Lichiere, J., Blangy, S., Bricogne, G., Legrand, P., Ortiz-Lombardia, M., Campanacci, V., van Sinderen, D., Cambillau, C., 2012. Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. Proc. Natl. Acad. Sci. USA. 109, 8954–8958.

Wang, J., Hofnung, M., Charbit, A., 2000. The C-terminal portion of the tail fiber protein of bacteriophage lambda is responsible for binding to LamB, its receptor at the surface of Escherichia coli K-12. J. Bacteriol. 182, 508–512.

Xia, G., Corrigan, R.M., Winstel, V., Goerke, C., Grundling, A., Peschel, A., 2011. Wall teichoic acid-dependent adsorption of staphylococcal siphovirus and myovirus. J. Bacteriol. 193, 4006–4009.

Xu, L., Benson, S.D., Butcher, S.J., Bamford, D.H., Burnett, R.M., 2003. The receptor binding protein P2 of PRD1, a virus targeting antibiotic-resistant bacteria, has a novel fold suggesting multiple functions. Structure 11, 309–322.