Molecular Microbiology (2017) 00(00), 00–00 ■

doi:10.1111/mmi.13649
First published online 00 Month 2017

Evolved distal tail carbohydrate binding modules of *Lactobacillus* phage J-1: a novel type of anti-receptor widespread among lactic acid bacteria phages

AQ2 **Maria-Eugenia Dieterle,[1,2,3] Silvia Spinelli,[2,3] Irina Sadovskaya,[4] Mariana Piuri[1]* and Christian Cambillau[2,3]***

[1]*Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, IQUIBICEN-CONICET, Buenos Aires, Argentina.*

[2]*Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche Scientifique (CNRS), Campus de Luminy, Case 932, Marseille Cedex 09, 13288, France.*

[3]*Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université (AMU), Campus de Luminy, Case 932, Marseille Cedex 09, 13288, France.*

[4]*Université Lille Nord de France, F-59000 Lille, France, Université du Littoral-Côte d'Opale, LR2B/UMT 08, Bassin Napoléon, Boulogne-sur-Mer Cedex, BP 120, F-62327, France.*

Summary

Bacteriophage replication requires specific host-recognition. Some siphophages harbour a large complex, the baseplate, at the tip of their non-contractile tail. This baseplate holds receptor binding proteins (RBPs) that can recognize the host cell-wall polysaccharide (CWPS) and specifically attach the phage to its host. While most phages possess a dedicated RBP, the phage J-1 that infects *Lactobacillus casei* seemed to lack one. They have shown that the phage J-1 distal tail protein (Dit) plays a role in host recognition and that its sequence comprises two inserted modules compared with 'classical' Dits. The first insertion is similar to carbohydrate-binding modules (CBMs), whereas the second insertion remains undocumented. They determined the structure of the second insertion and found it also similar to several CBMs. They found that expressed insertion CBM2, but not CBM1, binds to *L. casei* cells and neutralize phage attachment to the bacterial cell wall and that the isolated and purified CWPS of *L. casei* BL23 prevents CBM2 attachment to the host. Electron microscopy single particle reconstruction of the J-1 virion baseplate revealed that CBM2 is projected at the periphery of Dit to optimally bind the CWPS receptor. Taken together, these results identify J-1 evolved Dit as the phage RBP.

Accepted 12 February, 2017. *For correspondence. E-mail AQ1 ccambillau@gmail.com. OR E-mail marianapiuri@gmail.com.

Introduction

Replication of viruses requires a prerequisite specific host recognition and infection. The large majority of bacteriophages, viruses that infect bacteria, utilize a tail to recognize their host and to inject DNA (tailed phages). Among them, *Siphoviridae* are made of a capsid, containing the DNA, and a long non-contractile tail attached to the capsid by the connector. A recent interest has occurred for the tail tip that holds the molecular components recognizing the host (Sciara *et al.*, 2010; Veesler *et al.*, 2012; Desmyter *et al.*, 2013; Legrand *et al.*, 2016). These components can form straight tips or large organelles named baseplates, by analogy to the corresponding part of *Myoviridae* (Kostyuchenko *et al.*, 2003; Taylor *et al.*, 2016). The tip or the baseplate contain a protein/s that can recognize a bacterial receptor, either a protein, carbohydrate or teichoic acid (TA) moiety on the cell surface (Mahony *et al.*, 2016). To date, all structural analysis of *Siphoviridae* tails by bioinformatics tools, such as HHpred (Soding *et al.*, 2005), or experimentally using X-ray diffraction, have reported that an hexameric protein named Dit (distal tail protein) is attached to the last tail hexamer MTP (major tail protein) (Sciara *et al.*, 2010; Veesler *et al.*, 2010; Veesler and Cambillau, 2011; Veesler *et al.*, 2012; Flayhan *et al.*, 2014), while, on its other side, Dit binds to a trimeric protein named Tal (tail associated lysin) (Sciara *et al.*, 2010; Goulet *et al.*, 2011; Veesler and Cambillau, 2011), homologous to phage T4 gp27 (Kanamaru *et al.*, 2002; Kostyuchenko *et al.*, 2003).

In *Lactococcus* phages belonging to the 936 and P335 groups peripheral antireceptors or receptor binding proteins (RBPs) are present in the baseplate, in addition to Dit and Tal (Spinelli *et al*., 2006; Veesler and Cambillau, 2011; Spinelli *et al*., 2014). Regardless of sequence dissimilarity, these structural modules are generally conserved. In recent years, evidence has accumulated to indicate that the bacterial receptor for phages infecting *Lactococcus lactis* is a cell-wall-polysaccharide (CWPs) carbohydrate moiety (Mahony *et al*., 2013; Ainsworth *et al*., 2014; Farenc *et al*., 2014).

Less information is available about the host interaction proteins of lactic acid bacteria phages infecting *Lactobacillus* spp. Bacteriophage J-1 was isolated in 1965 during a failed fermentation of the Japanese beverage Yakult using the strain *Lactobacillus casei* 'Shirota' (Hino Mai, 1965). Phage J-1 belongs to *Siphoviridae* and can also infect other *Lactobacillus casei*/*paracasei* spp. including several commercial strains (Capra *et al*., 2006). As early as 1971, it was shown that L-rhamnose was probably part of the phage receptor because this sugar blocked J-1 adsorption to its host *L. casei*-S1 (*L. casei* ATCC 27139) (Yokokura, 1971). In 2014, after the sequencing of the J-1 genome, (Dieterle *et al*., 2014b) we annotated gp16 and gp17 as the baseplate proteins Dit and Tal respectively. The structure of Dit was modelled on the crystal structure of other *Siphoviridae* (Dieterle *et al*., 2014a). However, two regions of the C-terminus of Dit could not be modelled using this template, and one showed sequence similarity to carbohydrate-binding modules (CBMs). GFP-Dit fusions specifically bound to *Lactobacillus casei/paracasei* cells, and the addition of L-rhamnose inhibited binding. Recombinant Dit efficiently inhibits J-1 adsorption to its host, implicating that this 'evolved' Dit protein is involved in host recognition (Dieterle *et al*., 2014a).

Because of its location within the baseplate, only a structural role of inert hub has been assigned to Dit (Veesler *et al*., 2010). However, recent reports indicate the abundance of Dits of different sizes in phage genomes, from classical-length Dits (~ 260–300 amino-acids), resembling those from phages SPP1 (Veesler *et al*., 2010), TP901-1 (Veesler *et al*., 2012), p2 (Sciara *et al*., 2010) or T5 (Flayhan *et al*., 2014), to those of evolved Dits composed of ~ 450–700 amino-acids (Dieterle *et al*., 2014a; Dieterle *et al*., 2016; Murphy *et al*., 2016). Furthermore, a HHpred scan among various siphophage genomes infecting different species convinced us that evolved Dits are not a curiosity but are as abundant as the classical Dits. In most cases, HHpred could identify in the evolved Dits the insertions as putative CBMs (manuscript in preparation). However, besides our recent contributions, there is no firm indication to date on the functionality of these modules. Considering these observations and our interest for the model phage J-1, we embarked in the structural and functional analysis of phage J-1 Dit insertions.

Results

HHpred (Soding *et al*., 2005) analysis has proved to be a powerful tool to identify protein folds in bacteriophages, as they assemble modules that span the phage world. The phage J-1 Dit protein (gp16) is 679 residues long, much longer than the classical Dit type illustrated in structures from phages SPP1 (Veesler *et al*., 2010) and TP901-1 (Bebeacua *et al*., 2010; Veesler *et al*., 2012). Dit proteins are assembled in a hexamer that forms a ring with a 40 Å internal diameter and projects a C-terminal extension out of the ring. This extension has been reported to be a galectin-like module (Sciara *et al*., 2010; Veesler *et al*., 2010; Veesler *et al*., 2012) or an OB-fold domain (Flayhan *et al*., 2014). In contrast, HHpred analysis indicates that J-1 Dit exhibits two insertions: one just after the N-terminal ring domain and the second within a loop of the galectin-like domain (Supporting Information Fig. S1), a feature also observed in phage PLE3 Dit (Dieterle *et al*., 2016). We defined this insertion-containing Dit as evolved Dit (Dieterle *et al*., 2016). The first insertion in phage J-1 Dit (residues 129–322) was unambiguously assigned by HHpred to a carbohydrate-binding domain (CBM1), with a probability of ~ 98% (2XOM, 98.3%; 1GUI, 98.1%) (Supporting Information Fig. S1). However, HHpred did not retrieve a significant hit in the PDB for the second insertion (residues 368–612). Based on this analysis, J-1 evolved Dit differs from classical Dits by two insertions containing putative CBM domains (Fig. 1). We therefore identified two domains to clone and express: Dit CBM1 between residues 129–322 and the putative domain of the second insertion (putative CBM2) between residues 368–614 (Table 1). Finally, to evaluate binding to *Lactobacillus* cells, we constructed and expressed GFP-CBM1 and mCherry-CBM2 fusions. F1 T1

Structure of CBM2

To confirm the nature of the second insertion, we subjected it to crystallization assays. The CBM2 structure was determined to 1.28 Å resolution (Table 2, Supporting Information Fig. S2), and the amino-acid chain was traced between residues 10 and 239 (Fig. 2a). T2 F2

The structure of CBM2 consists of 18 β−strands and several long stretches and loops (Fig. 2a). The core of the structure is a β-sandwich of 2 β-sheets gathering 12 β-strands: β-strands 2, 17, 8, 13, 14 and 15 for the first β-sheet, and β-strands 7, 16, 9, 10, 11 and 12 for the

COLOR IN ONLINE AND PRINT

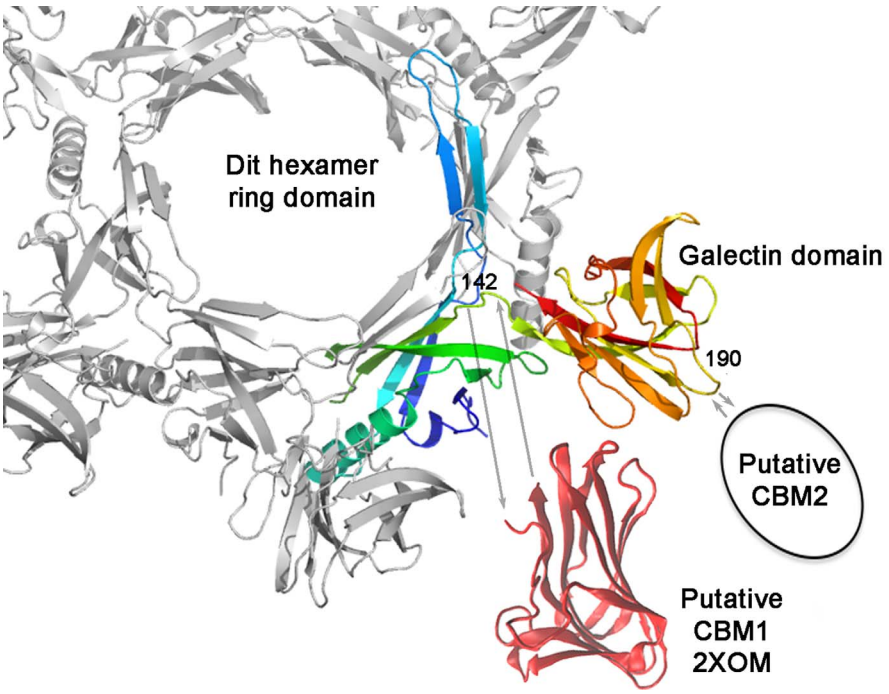


Fig. 1. Topological analysis of J-1 Distal Tail Protein (Dit). Topological model of J-1 Dit showing the insertion location of putative CBM1 and CBM2. The hexameric Dit is displayed and the SPP1 Dit insertions numbering is given (J-1 numbering is 142 and 367 respectively). In one Dit of the hexamer, insertions are shown and rainbow coloured.

second. This second β-sheet is concave and is covered by other segments of the module (see below).

The N- and C-termini form two antiparallel β-strands (strands 1 and 18) that join the module to the rest of the evolved-Dit. Following β-strand 2, an elongated stretch of ~ 20 residues crosses the concave face of β-sheet 2, turns at the level of β-strand 12, and continues by two short β-strands 3 and 4. After a turn, β-strand 5 extends antiparallel to β-strand 4 and is followed by ~ 10 residues that form a large unstructured turn. This turn is followed by β-strand 6, antiparallel to β-strand 3, that abuts to β-strand 7 of β-sheet 2. Long loops connect β-strands 7–8 and 15–16 and β-strands 10–11, 12–13 and 16–17 (Fig. 2a).

We submitted the crystal structure of this module to the DALI server to retrieve possible similar folds from the PDB (Holm *et al.*, 2008). Several dozen hits were returned with Z values above 15 and r.m.s.d. values between 2.5 and 3.2 Å (Supporting Information Fig. S3). All hits were carbohydrate-binding modules, mostly lectins. Of particular interest among the most similar folds are an ERGIC-53 lectin ($Z = 16.8$; r. m. s. d. 2.5 Å; 3wht, 3whu, 3wnx) and an acidic lectin ($Z = 16.3$; r. m. s. d. 2.8 Å; 1fay), as they have been co-crystallized with saccharides (mannotriose and methyl-alpha-D-galactose respectively) (Supporting Information Fig. S4a,b). These results confirmed our hypothesis that the second insertion module could belong to a CBM fold. After the superposition of CBM2 and the acidic lectin (PDB entry 1fay), the galactose bound to the acidic lectin was imported into the CBM2 structure, allowing the identification of a putative carbohydrate-binding domain (Fig. 2c). The saccharide fits nicely into a cavity opposite the N- and C-termini, at the extreme end of the β-sheet 2 concave surface. This cavity is delineated in part by β-sheet 2 and also by the extended stretch of residues and the turn crossing the concave surface (Fig. 2b and c). A cavity extension is visible near the galactose moiety, that accommodates divalent cations in classical lectins; in acidic lectins these cations are Ca^{++} and Mg^{++}. However, the residues that bind divalent cations in the acidic lectin (mainly acidic residues or main-chain carbonyls) are absent in CBM2; this portion of the cavity is mainly bordered by aromatic or aliphatic residues, often present in carbohydrate binding-sites and stacking against saccharide's apolar face. Notably, even though J-1 infection is strictly dependent on the presence of Ca^{++} (10 mM), CBM2 binding to cells is independent of Ca^{++} (see below).

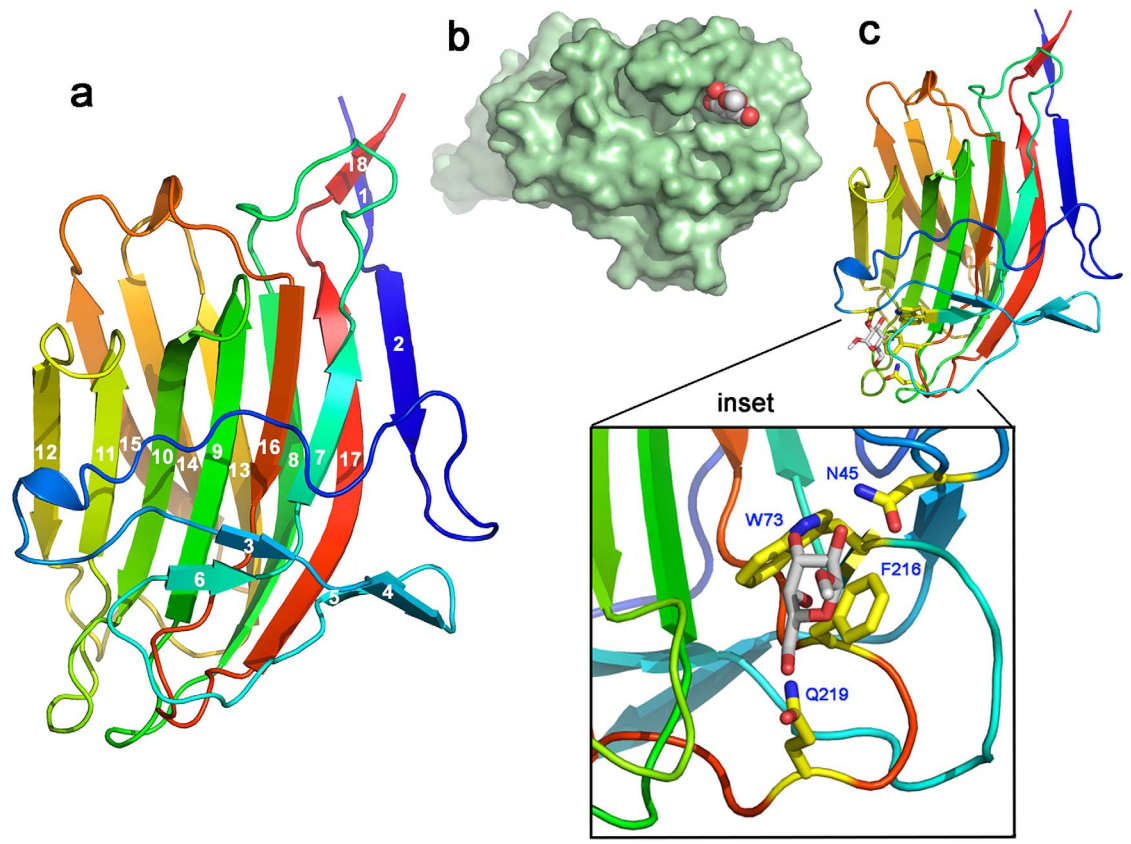
Looking more closely at the putative carbohydrate binding site, we noticed features remarkably consistent with a *bona fide* carbohydrate-binding site (Hudson *et al.*, 2015). The galactose stacks ring-to-ring against Phe 216, perpendicular to Trp 73, and is at a good hydrogen-bonding distance to Gln 219 NH_2, Trp 73 NH and the Asn 45 side-chain (Fig. 2c, inset). However, it would be hazardous to deduce the exact saccharide specificity of CBM2 from this model, as the carbohydrate-binding site may change conformation upon ligand binding. Finally, we performed crystallization assays with the well-expressed CBM1. However, despite our efforts, diffraction did not extend beyond 7 Å, and the project was abandoned.

Table 1. Plasmids and primers used in this study.

Plasmid	Region (amino-acids)	Primers used (5′ to 3′)
Plic07–CBM1	129–322 of Dit	ED104 CGAGAACCTGTACTTCCAATCAATGAAGACGGCTGACAACATGCC
		ED105 GGATCCGTATCCACCTTTACTGTTATTAATCCGCTGGGTTAGGTGACCA
Plic07–CBM2	368–614 of Dit	ED106 CGAGAACCTGTACTTCCAATCAATGGGCTCTCCCGATGAGACTG
		ED107 GGATCCGTATCCACCTTTACTGTTATTATTTGAAGCGGTTAGGAATATCAT
GFP–CBM1	129–322 of Dit	ED86 TAGCAGAATTCACGGCTGACAACATGCCATA
		ED87 TGAACGAGCTCCTAAGGTACCGCTGGGTTAGGTGACC
mCherry–CBM2	368–614 of Dit	ED96 TAGCAGCTAGCATGGTTTCAAAAGGGGAGG CHERRY
		ED97 TGAACGGATCCTTTATATAATTCGTCCATGCCAC
		ED98 TAGCAGGATCCCCCGATGAGACTGATGGTT CBM2
		ED99 TGAACGAGCTCTTAGTCTTTGAAGCGGTTAGG
GFP–Dit Cterminal	129–614 of Dit	ED86 TAGCAGAATTCACGGCTGACAACATGCCATA
		ED99 TGAACGAGCTCTTAGTCTTTGAAGCGGTTAGG
GFP–Dit	1–679 of Dit	Previous work (23)

CBM1 and CBM2 binding assays to the host

To complement our previous results regarding Dit binding (Dieterle *et al.*, 2014a), we performed flow cytometry assays and observed a shift between cells alone (Fig. 3a, upper graph) and cells in the presence of GFP-Dit (below graph and Fig. 3d). Furthermore, anti-Dit antibodies inhibited J-1 adsorption to host cells in a concentration-dependent manner (Fig. 3b) and completely inhibited phage infection (Fig. 3c),



COLOR IN ONLINE AND PRINT

Fig. 2. Crystal structure of CBM2 and comparison to other CBMs. **a.** Ribbon representation of the CBM2 X-ray structure, coloured in rainbow mode, from blue (N-terminus) to red (C-terminus)**. b.** Surface representation of the model of a complex between CBM2 (green) with a saccharide in its cavity. The saccharide moiety (α-methyl-d-Galactose) belongs to 1fay that has been superimposed onto CBM2. **c.** Ribbon representation of the model of a complex between CBM2 (rainbow mode) and a α-methyl-d-Galactose (stick representation) belonging to 1fay that has been superimposed onto CBM2. The side-chains of CBM2, putatively contacting the saccharide, are represented as sticks in the main view and in the **inset**. Their numbering corresponds to that of the deposited PDB. The numbering corresponding to the full-length Dit is N412 (N45), W440 (W73), F583 (F216) and Q586 (Q219).

Table 2. Data collection and refinement statistics.

Data collection	CBM2 CsI/NaI	CBM2 native
PDB	–	5LY8
Source	ESRF BM14	Soleil PX1
Wavelength (Å)	1.7712	0.953
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Cell (Å), angles (˚)	48.6, 64.4, 74.5/3 x 90.0	49.0, 63.7, 74.5/3 x 90.0
Resolution limits[a] (Å)	50.0–1.97 (2.04–1.97)	34.4–1.28 (1.31–1.28)
R_{merge}[a]	0.061 (0.212)	0.034 (0.432)
CC1/2	99.9 (98.6)	100.0 (93.2)
Unique reflections[a]	15081 (768)	60896 (4359)
N.ano	13028 (664)	
Mean((I)/sd(I))[a]	31.0 (8.8)	28.7 (3.8)
Completeness[a] (%)	88.2 (46.5)	99.8 (97.6)
Multiplicity[a]	13.3 (7.5)	7.0 (5.9)
SigAno[a]	1.52 (1.3 at 2.3 Å)	–
CCano	65 (56 at 2.3 Å)	–
REFINEMENT		
Resolution[a] (Å)		34.4–1.28 (1.31–1.28)
Number of reflections[a]		60732 (4217)
Number of protein/water/ligand atoms		1923/317/2
Number of test set reflections		3037
R_{work}/R_{free}[a] (%)		0.184/0.198 (0.203/0.217)
r.m.s.d. bonds (Å)/angles (˚)		0.012/1.21
B-wilson/B-average		14.0/20.4
Ramachandran: preferred/allowed/outliers (%)		97.5/2.5/0

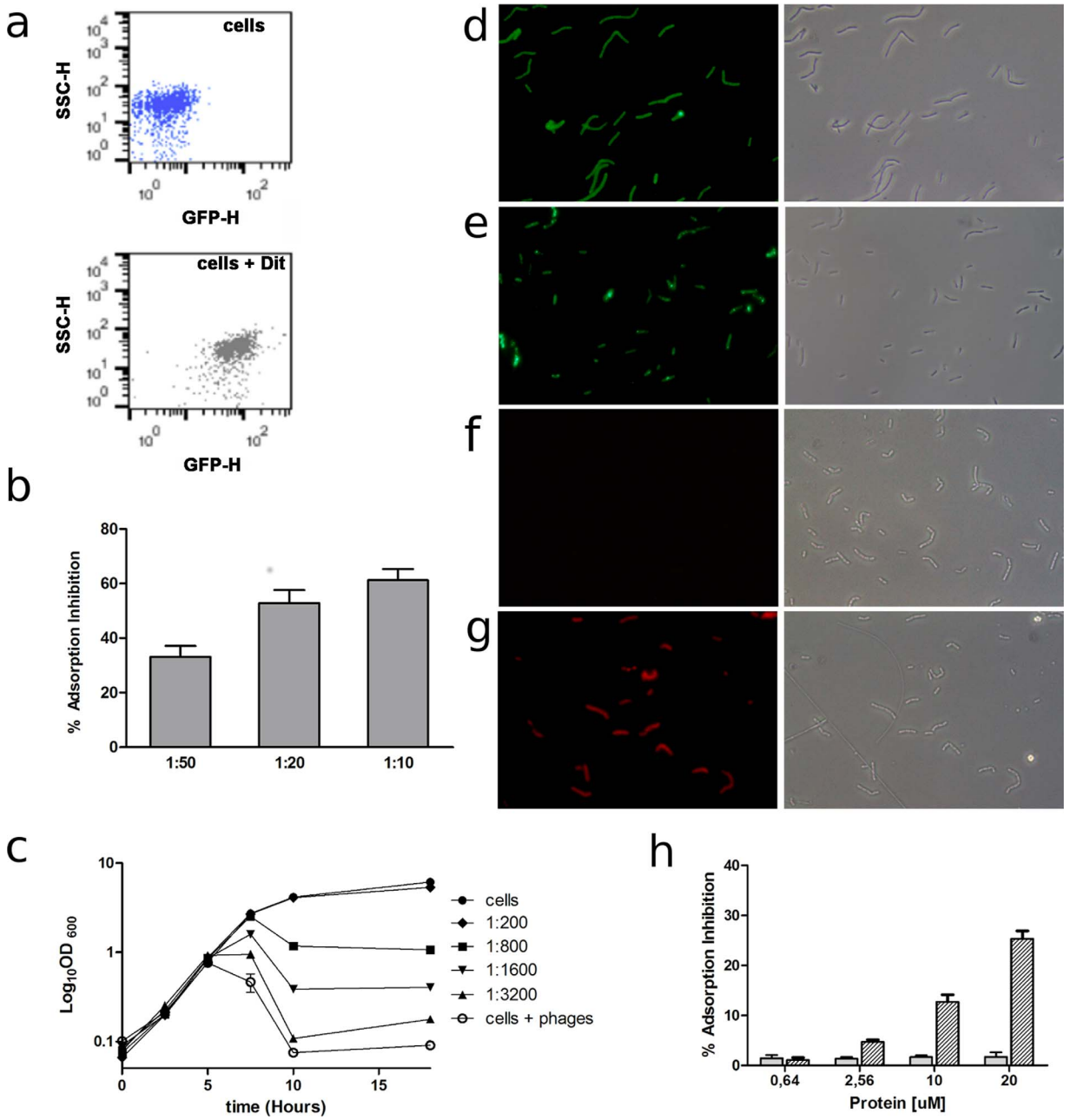
a. Numbers in brackets refer to the highest resolution bin.

corroborating the involvement of Dit in host recognition. Because a fold is not always related to a function, we analysed the role of CBM1 and CBM2 in J-1 adhesion and infection. To this end, we used GFP-CBM1 and mCherry-CBM2 and GFP-Dit C-terminal protein fusions in binding assays (Table 1). Fluorescent cells can be visualized using fluorescence microscopy when the fusion protein binds to the bacterial envelope. When *L. casei* subsp. *casei* ATCC 27139 cells were incubated with GFP-Dit C-terminal, cells appeared bright green (Fig. 3e). Similarly, when incubated with mCherry-CBM2, cells were decorated and appeared bright red (Fig. 3g). Conversely, the GFP-CBM1 fusion did not bind to the bacterial surface (Fig. 3f). To ensure that the failure to bind was not due an incorrect fold of CBM1 in the GFP-CBM1 construct, we performed additional binding experiments with CBM1, which is presumably folded properly, because of its ability to crystallize. The cells were first incubated with recombinant CBM1, and detection was performed using an anti-Dit polyclonal antibody and a secondary antibody conjugated with a red fluorophore. The binding of CBM1 to the bacterial surface was not detected, but fluorescent cells were easily visualized when performing a control assay of the technique with CBM2 (Fig. 4). In agreement with these results, CBM2, but not CBM1, was able to prevent J-1 adsorption to its host in a competition assay (Fig. 3h). It is notable that CBM2 could not exert the level of inhibition of the entire Dit protein as previously reported, (Dieterle *et al.*, 2014a) suggesting that, indeed, the avidity of Dit with six CBM2 far surpasses single-domain affinity.

CBM1 and CBM2 binding assays to the purified cell wall polysaccharide

Phage J-1 can infect its host strain *L. casei* subsp. *casei* ATCC 27139 and several other *Lactobacillus casei/paracasei* strains. Since, we and others have previously reported that rhamnose inhibits J-1 adsorption and Dit binding (Yokokura, 1971; Dieterle *et al.*, 2014a), we tested CMB2 binding in the presence of this sugar. CBM2 could not decorate *L. casei* subsp. *casei* ATCC 27139 cells in the presence of rhamnose (Supporting Information Fig. S5). Recently, the presence and structure of a cell wall polysaccharide fraction in the model *L. casei* BL23 strain was described and all constituent polysaccharides were rich in rhamnose (Vinogradov *et al.*, 2016). J-1 can infect *L. casei* BL23, and its efficiency of plating is similar to its host strain *L. casei* subsp. *casei* ATCC 27139 (data not shown). We repeated the described binding assays with Dit and CBM fusions using *L. casei* BL23 cells in the absence and presence of isolated CWPS from the same strain. All tested fusion proteins, with the exception of GFP-CBM1 (data not shown), decorated these bacteria (Fig. 5a and b; left panel); binding was

J_ID: MMI Customer A_ID: MMI13649 Cadmus Art: MMI13649 Ed. Ref. No.: MMI-2016-16308.R2 Date: 16-February-17 Stage: Page: 6



COLOR IN ONLINE AND PRINT

Fig. 3. Binding assays of phage J-1 Dit (gp16), CBM1 and CBM2 to the host *L. casei* subsp. *casei* ATCC 27139. **a.** Flow cytometry assays showing GFP-Dit binding to the host. The panel shows the shift between cells alone (upper graph) and cells in the presence of Dit protein (below graph), the intensity of fluorescence is plotted against light side scatter**. b.** Inhibition of the adsorption of phage J-1 to its host using increasing concentration of anti-Dit antibodies. **c.** Growth kinetics of the host strain determined by OD_{600nm} for a culture containing 1×10^3 PFU ml^{-1} and the indicated amounts of anti-Dit antibodies over 18 hours at 37˚C. **d–g.** Recombinant proteins GFP-Dit **(d),** GFP-C term **(e),** GFP-CBM1 **(f)** and mCherry-CBM2 **(g)** were incubated with *Lactobacillus casei* subsp. *casei* ATCC 27139. Cells were visualized using phase-contrast microscopy (right image) and fluorescence microscopy (left image). Magnification 1000×. **h.** Adsorption inhibition was determined when *L. casei* subsp. *casei* ATCC 27139 cell walls were incubated with increasing amounts of CBM1 domain (grey bars) or CBM2 domain (striped bars), followed by adsorption assays using phage J-1. The error bars represent standard deviations from experiments performed in triplicate.

inhibited in the presence of the CWPS (Fig. 5a and b; right panel). In addition, the CWPS of *L. casei* BL23 inhibited Dit and CBM2 binding to *L. casei* subsp. *casei* ATCC 27139 cells (Fig. 5c and d; right panel).

Electron microscopy structure of the baseplate

Considering the location of Dit in the genome, its homology to other phage proteins and our previous results, (Dieterle *et al.*, 2014a) we expected it to localize to the

ID: rajae Time: 21:40 I Path: //chenas03/Cenpro/ApplicationFiles/Journals/Wiley/MMI#/Vol00000/170032/Comp/APPFile/JW-MMI#170032

COLOR IN ONLINE AND PRINT

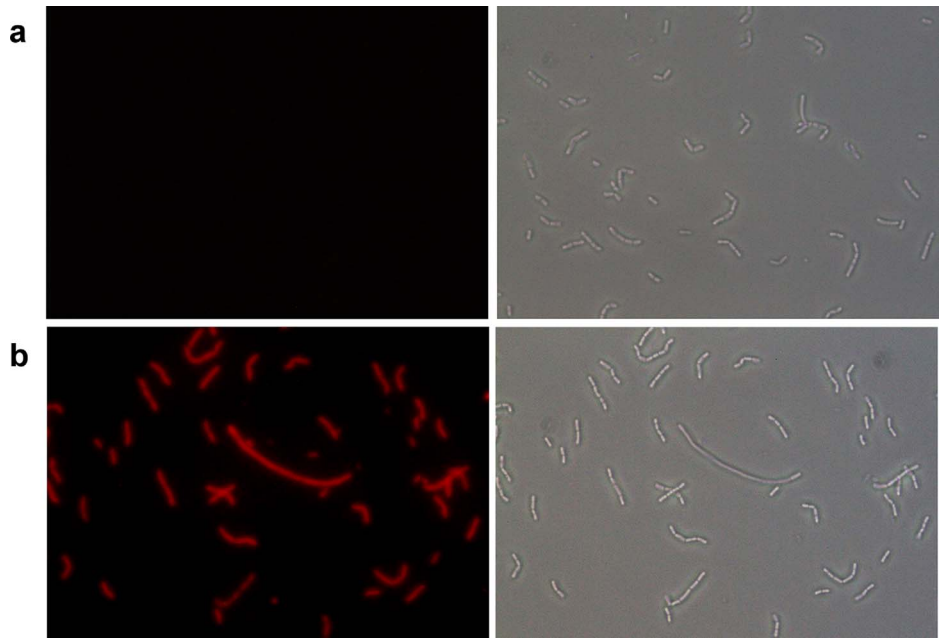


Fig. 4. Immunofluorescence binding assay. Recombinant CBM1 (upper panels) and CBM2 protein (lower panels) was challenged with *Lactobacillus casei* subsp. *casei* ATCC 27139, and binding was detected using a polyclonal antibody against Dit and a secondary antibody conjugated with a red fluorophore. Cells were visualized using phase-contrast microscopy (right image) and fluorescence microscopy (left image).

baseplate at the tip of the phage tail. Transmission electron microscopy (TEM) of the phage incubated with specific primary antibodies was performed. A secondary gold-labelled antibody allowed determination of the location of Dit at the very tip of the phage tail, with more than one copy per phage particle (Fig. 6a and b).

Knowing that CBM2 binds to host cells, we wondered about its exact location in the baseplate. To answer this question, we investigated the structure of the phage J-1 virion using negative-staining transmission electron microscopy. The baseplate structure was easily recognizable at the end of the tail, followed by a long extension of the Tal protein (Fig. 6a). We collected 200 images and boxed 865 particles of the baseplate; we then determined the baseplate structure at 20 Å resolution. (Fig. 6c–e). Based on our HHpred analysis, we fit as a block with Chimera (Pettersen *et al*., 2004) the X-ray structure of Dit and Tal of phage p2 (Sciara *et al*., 2010), in which the galectin domain extension ('arm and hand') was deleted. This ensemble provides the 'classical' part of J-1 Dit and the N-terminal (T4 gp27-like) domain of its Tal (Fig. 6c–e, blue). We then calculated the difference in volume between the experimental EM map (without the tail part) and the p2 Dit/Tal model. We identified two series of 6 EM density bulbs that we assigned to both CBM domains. The first group of EM densities, located at the extremity of the galectin domains, could be unequivocally ascribed to CBM2 based on its position in the Dit sequence (Fig. 6c–e; green). Notably, this EM density is faint compared with the size of CBM2, most likely due to its mobility. The second group of EM densities is located between the galectin domains against the ring domain of Dit (Veesler *et al*., 2010). We assigned this density to CBM1 due to its position in the Dit sequence at the junction between the Dit ring and the galectin domains (Fig 6c–e; red). We placed the two series of CBM1 and CBM2 structures manually into the difference EM density using Chimera (Pettersen *et al*., 2004). Based on the above described hybrid model, we measured the baseplate as ~ 125–210 Å wide and ~ 110 Å high. Notably, the Tal spike extension beyond the 400 first residues (Fig. 6a, right) was averaged out due to its flexibility and is, therefore, not visible in the structure aside from a small bulb of density at its base.

Discussion

We reported earlier that the expressed Dit of phage J-1, infecting *Lactobacillus casei* BL23, was able to bind to the host and to prevent infection. Two insertions in two positions of the Dit sequence suggested that extra domains could be present in J-1 Dit, as compared with classical Dits. Sequence comparison and HHpred revealed that the first one was related to a carbohydrate binding domain (CBM1), while the nature of the second one remained unknown. This allowed identification of the borders of both domains that were cloned and expressed in *Escherichia coli*. Both were expressed in a soluble form and could be isolated as a unique form by gel filtration. We then confirmed the nature of the second insertion. To this end, we subjected it to crystallization assays, collected a X-ray data set and determined

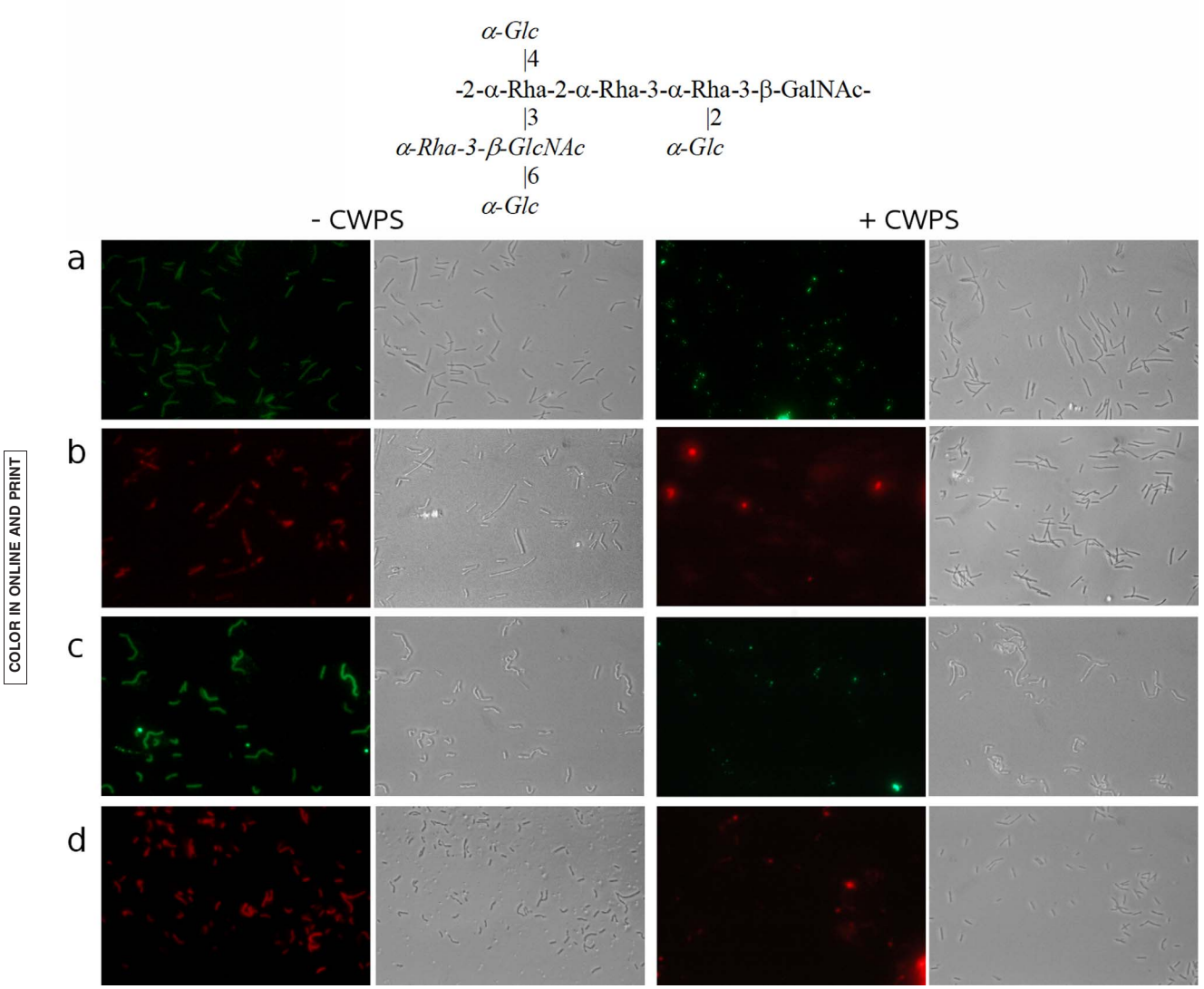


Fig. 5. Influence of purified *L. casei* CWPS on CBMs and phage binding. Repeating unit of the *L. casei* BL23 cell-wall polysaccharide (Vinogradov *et al.*, 2016). Components in italic are non-stoichiometrical (above). **a–d.** Binding of J-1 GFP-Dit (**a**, **c**) or mCherry-CBM2 (**b**, **d**) to *L casei* BL23 (**a, b**) or *L. casei* subsp. *casei* ATCC 27139 (**c, d**) was evaluated in the absence (left panels) or presence (right panels) of CWPS from *L. casei* BL23.

the structure to 1.28 Å resolution. This crystal structure allowed to identify the second insertion as a carbohydrate binding domain, CBM2. Comparison with similar structures in the PDB made it possible to locate the saccharide binding site.

We then constructed and expressed GFP-CBM1 and mCherry-CBM2 fusions of CBM1 and CBM2 domains in view to evaluate their binding to *Lactobacillus* cells. Rather surprisingly, only CBM2 but not CBM1 was able to bind to the host cells, and prevent J-1 adsorption to its host in a competition assay. The good in vitro behaviour of CBM1, its ability to crystallize (although diffracting at low resolution) and immunofluorescence assay, all suggest that CBM1 is properly folded, but is unable to bind to the host. As compared with Dit, CBM2 could not exert the same level of inhibition of phage adsorption to its host. Indeed, the Dit hexamer with six CBM2 has by far a larger affinity than a single-domain (avidity). Our previous analysis (Dieterle *et al.*, 2014a) indicates that CBM1 is variable among the *Lactobacillus* phages so we cannot discard that this fold could be functional in other phage-host systems.

Interestingly, the structure of the cell wall polysaccharide (CWPS) covering the *Lactobacillus casei* BL23 host was recently isolated and purified, and its structure determined (Vinogradov *et al.*, 2016). We used this purified complex polysaccharide to check if the only active module, CBM2, could bind to it. To this end, we

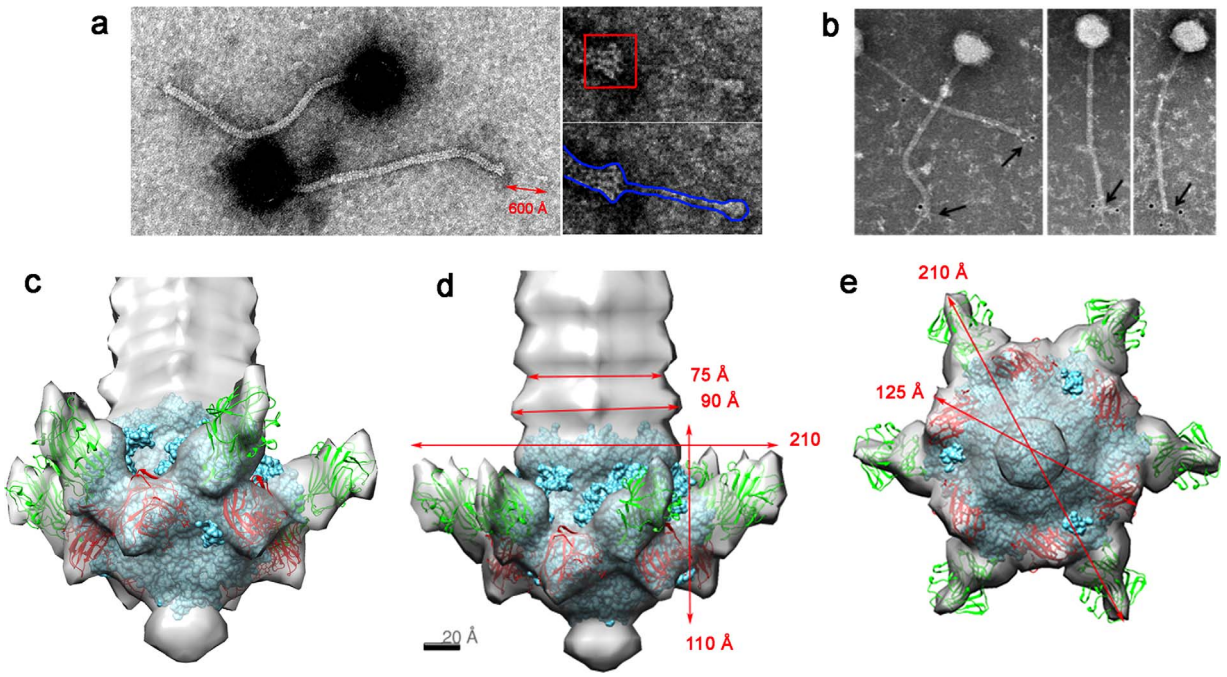


Fig. 6. Negative staining electron microscopy single particle reconstruction of J-1 virion baseplate. **a.** View of phage J-1 virion (left) and its tail tip with the baseplate boxed in red (right, above) and the baseplate and Tal tip evidenced in blue (right, below). **b.** Immunolabelling of virion Dit protein. **c.** View of tail segments with the baseplate comprising Dit and Tal N-terminus. **d.** Lateral view of the baseplate with dimensions. **e.** Bottom view of the baseplate with dimensions. (**c–e**) Dit core and Tal are blue; CBM2 is green; and 2xom, a topological model of CBM1, is red.

decorated the host cells with mCherry-CBM2 fusion, in the presence or absence of CWPS. CBM2 (and Dit) binding to the *L. casei* subsp. *casei* BL23 and *L. casei* subsp. *casei* ATCC 27139 cells was inhibited in the presence of CWPS.

Finally, we determined the negatively stained electron microscopy structure of the tail tip of phage J-1, assembling the MTP, Dit and Tal modules. Using homologous modules of known structures and CBM2, we propose a complete model of the phage J-1 baseplate fit into its EM map. We determined that CBM2 is pointing out of the J-1 baseplate, fully available for efficient interaction with the host CWPS.

All in all, our results suggest that Dit CBM2 is a *bona fide* receptor binding protein of phage J-1. We suggest that such Dit inserted CBM domains – putative receptor binding protein – exist in other species of the siphophage world. To support this hypothesis, we recently reported the presence of putative CBMs in *Lactobacillus* phage PLE3 (Dieterle *et al.*, 2016). Furthermore, in a recent study involving 38 lactococcal phage genomes from the 936 group, 16 Dits showed a length compatible with an evolved Dit type (~ 450 residues or more) (Murphy *et al.*, 2016). As an example, we analysed the phage PhiM1127 Dit that has a sequence of 492 residues. This exhibits an insertion between residues 132 and 241, and HHpred identifies it to BppA, an accessory CBM of lactococcal phage Tuc2009 (Legrand *et al.*, 2016) (Fig. 7a). Other examples of insertions in Dit have been reported in the lactococcal phage group P335 (Mahony *et al.*, 2017) (e.g., phage C4143) (Fig. 7b).

Taken together, the present results of our analysis of phage J-1 infecting *Lactobacillus casei* BL23 clearly show that its evolved Dit CBM2 is a functional *bona fide* receptor-binding protein. Considering the abundance of long evolved Dits in the siphophage world, this result introduces a drastic paradigm shift concerning the Dit proteins: evolved Dits should be considered (and investigated) as 'active' proteins toward the host, and not passive hubs. They may serve as unique RBPs or as associated recognition modules involved in the host adhesion process.

Experimental procedures

Strains, bacteriophages and growth conditions

Lactobacillus paracasei subsp. *paracasei* ATCC 27139 was grown in MRS medium (Difco, USA) at 37˚C under static conditions. *Escherichia coli* DH5α was used for cloning, and *E. coli* BL21(DE3) pLysS (Invitrogen, USA) was used for protein expression. *E. coli* strains were grown in LB broth or Terrific Broth (Difco, USA) at 37˚C under moderate shaking. When appropriate, antibiotics were added at the

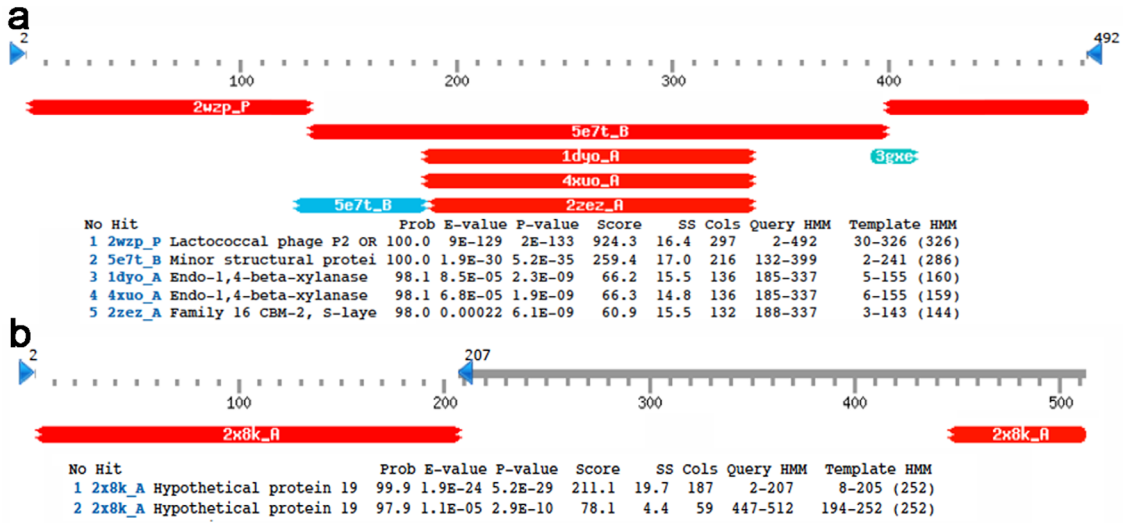


Fig. 7. HHpred plots of lactococcal Distal Tail Proteins. **a.** Dit protein from phage PhiM1127 belonging to lactococcal group 936 (Murphy *et al.*, 2016). **b.** Dit protein from phage C4143 belonging to lactococcal group P335 (Mahony *et al.*, 2017).

following concentrations: kanamycin, 30 μg ml^{-1} (Sigma, USA); chloramphenicol, 34 μg ml^{-1} (Sigma, USA).

Lactobacillus phage J-1 was propagated on *L. casei* subsp. *casei* ATCC 27139. Bacteriophage stocks were stored at 4˚C in phage buffer (20mM Tris-HCl, 100 mM NaCl and 10 mM $MgSO_4$).

Cloning procedures

All primers used were salt-free and purchased from Eurofins MWG Operon. Sequences were amplified from *Lactobacillus* phage J-1 genome using primers detailed in Table 1. Sequences were cloned into the pLIC07 vector (kindly provided by BioXtal; unpublished work) or pET28-GFP (Dieterle *et al.*, 2014a) or pET28-mCherry (This work). The pLIC07 vector was designed for ligation-independent cloning (Aslanidis and de Jong, 1990) and is a derivative of the pET-28a+ expression vector (Novagen) in which a cassette coding for a 6× His tag, Trx gene and a Tobacco etch virus (TEV) protease-cleavage site followed by the suicide gene sacB. A restriction free protocol (Unger *et al.*, 2010) was used to construct plasmids plic07 CBM1, plic07 CBM2 (Table 1).

For fluorescence microscopy assays, CBM1, CBM2 and Dit-Cterminal PCR products were digested with EcoRI/SacI (CBM1), BamHI/SacI (CBM2), EcoRI/SacI (Dit Cterminal) and cloned into pET28b GFP or pET28b mCherry. Plasmids were named GFP-CBM1, mCherry-CBM2, GFP-Dit CTerminal.

Protein expression

Recombinant plasmids were transformed in *E. coli* BL21 (DE3)pLysS (Novagen) or T7 Express Iq pLysS (New England Biolabs) strains. Cells were grown at 37˚C in Terrific Broth or minimal medium until the OD600nm reached 0.6, after which protein expression was induced with 1 mM IPTG overnight at 17˚C. The cells were harvested by centrifugation (4000*g* for 10 min) and the pellet was homogenized and frozen in lysis buffer [50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.1mg ml^{-1} lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF)]. After thawing, DNAse I (20 μg ml^{-1}) and $MgSO_4$ (1 mM) were added and the cells were lysed by sonication. The pellet and soluble fractions were separated by centrifugation (16,000*g* for 30 min). Purification was performed on an ÄKTA FPLC system following an immobilized metal ion-affinity chromatography using a 5 ml HisTrap Crude (GE Healthcare) Ni^{2+}-chelating column equilibrated in buffer A (50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole). The proteins were eluted with buffer B (buffer A supplemented with 250 mM imidazole). After desalting and when it was necessary, a tev protease cleavage was performed (1:20 tev protease: protein) and a second Ni^{2+}–NTA column chromatography was done followed by a preparative HiLoad 16/60 Superdex 75 pg or Superose 6 HR16/60 equilibrated in 10 mM HEPES pH 7.5, 100 mM NaCl or 10 mM HEPES pH 7.5 300 mM NaCl. For crystallization trials, the purified domains were concentrated by centrifugation in the same buffer as used for size-exclusion chromatography using an Amicon 10–100 kDa cut-off.

Immunolabelling

Purified phage preparation was dialyzed for 20 minutes against TGB buffer [200 mM Tris, 500 mM glycine, 1% (vol/vol) butanol pH 7.5] using Novagen D-tube TM dialyzer MWCO 3.5 kDa. Dialyzed phages were put over a Ni glow discharged grid for 20 seconds and then were incubated overnight (20˚C) with primary antibody solution (anti-GFP Dit), which were diluted 1/2000 in TGB buffer. Afterward, the grid was washed in TGB and incubated for 1 hour in a dilution 1:40 of the secondary antibody solution (goat anti-mouse immunoglobulin G 5 nm Gold conjugate solution-Sigma-Aldrich). Then fixation was done using 0.25%

(vol/vol) glutaraldehyde for 20 min in phosphate-buffered saline buffer at RT. The grid was washed and blotted five times in filtered dialyzed water. Finally, the samples were stained with 2% phosphotungstic acid (pH 7.0) for about 30 sec and were observed using a Tecnai Spirit electron microscope operated at 120 kV and a 2000- by 2000-pixel CCD camera. Antibodies against the recombinant purified GFP-Dit were raised in mice by Hebe Perez in Argentina.

Inhibition of adsorption assays and anti-Dit antibody neutralization capacity

The inhibition of adsorption assay was done as described before (Dieterle *et al*., 2014a). In brief, 50 μl of cell walls (100 μg) were incubated with 50 μl of buffer (control), CBM1 or CBM2 at different concentrations at room temperature for 20 min. Then, 50 μl of phage was added (1×10^5 PFU). The mixture was incubated at 37˚C for 1 h, and cell walls were removed by centrifugation at 3200*g* for 10 min. The unabsorbed phage in the supernatant was measured using the double agar method.

Bacteriophages (1×10^5 PFU) were challenged against anti-Dit antibodies for 20 minutes. Then, 100 μl of *L. casei* subsp. *casei* ATCC 27139 (1×10^8CFU) was added. Same protocol as described above was followed. For the liquid assay, bacteriophages (1×10^3 UFP ml^{-1}) were incubated with anti-Dit antibodies at different concentrations (dilutions 1:200 to 1:3200) for 20 minutes in 20 ml of MRS broth. Then, 1% of an ON cell culture and 10 mM of $CaCl_2$ were added to each sample. Controls without phages and without antibodies were run in parallel.

Fluorescence binding assays

Cell binding assays using purified CBMs (GFP-CBM1, mCherry-CBM2, GFP-Dit Cterminal) domains and GFP-Dit were carried out as described before (Dieterle *et al*., 2014a) with some modifications. Briefly, 0.3 ml of exponentially growing bacterial cells were centrifuged and resuspended in 100 μl of modified phage buffer (MPB) (50 mM Tris-HCl, 100 mM NaCl, 0.1% Tween 20, 10 mM $CaCl_2$) and incubated with 2 μg of protein fusions for 20 min at room temperature. Cells were washed three times with phosphate-buffered saline (PBS) buffer, and binding to the bacterial cells was detected by fluorescence microscopy (Axiostar Plus; Carl Zeiss) with a 100×× objective with oil immersion and phase contrast. When binding was tested in the presence of CWPS, GFP-Dit or mCherry-CBM2 was pre-incubated with 300 μg of CWPS for 30 min at room temperature, and the protocol was followed as described above. CWPS were purified as described in Vinogradov *et al*. (2016). For sugar binding inhibition, same protocol as described in Dieterle *et al*. (2014a) was followed.

Immunofluorescence assays

About 0.5 ml of exponentially growing bacterial cells culture of *L. casei* subsp. *casei* ATCC 27139 was washed three times with PBS. Cells were resuspended on 50 μl of PBS and 3 μg of protein was added (CBM1, CBM2 or negative control) and it was incubated for 30 min at room temperature. Cells were washed three times with PBS and anti Dit antibody (1/200) was added for an hour. Cells were washed three times with PBS and a dilution of 1/200 goat anti mouse IgG Alexa Fluor-647 (Life Technologies) was added for an hour. Cells were washed three times with PBS and resuspended in 30 μl of PBS before inspection by fluorescence microscopy.

Flow cytometry

Same protocol as described above in fluorescence binding assay was used with some modification. After washing the cells with MPB, 50 μl of cells were incubated with 2 μg of GFP-Dit or without it for 30 min at room temperature. The mixture was washed three times with PBS and resuspended in 250 μl of PBS with approximately 1×10^6 cells. Cells were analysed by BD FACSAria II flow cytometer to a 488-nm light source. The bacterial populations first were gated based on their SSC and FSC profiles and were recognized by sampling a blank PBS solution in parallel and 10,000 events were recorded in each experiment. Fluorescent and nonfluorescent cells within the gated population were discriminated based on fluorescent intensity (GFP-H). BD FACSAria software (BD FACSDiva, firmware version 6.1.3) was used for data acquisition, and FlowJo v10 software was used for subsequent analysis.

Crystallization and structure determination of the CBM modules

Crystallization of CBM2 was achieved using protein at 8 mg ml^{-1} in 0.8–1.2 M sodium citrate as precipitant with 10 mM sodium borate at pH 8.9–9.7. A crystal of CBM2 was soaked with CsI/NaI and a data set was collected at ESRF BM14 beamline (ESRF, Grenoble, France) on a MAR 225 CCD detector at 100K. Diffraction images were processed and scaled with the HKL2000 (Denzo and Scalepack) package (Otwinowski, 1993). A dataset of 360˚ was collected on a single CsI/NaI crystal derivative at 1.7712 Å X-ray wavelength up to 1.7 Å resolution (Table 2). The heavy atom substructure was obtained with the SHELXC/D/E (Sheldrick, 2008)software suite using the HKL2MAP graphical interface (Pape and Schneider, 2004). This substructure was subsequently refined and used for phases calculation with Phenix AutoSol package (Adams *et al*., 2010). Phases improvement and extension by density modification, produced readily interpretable maps that allowed building the whole CBM2 model. Inspection of electron density maps and model adjustment and rebuilding were performed using COOT (Emsley *et al*., 2010) alternating with autoBUSTER model refinements (Blanc *et al*., 2004). A 1.28 Å resolution native dataset was then collected at Proxima 1 beamline (Soleil, France) on a DECTRIS Quantum 6 detector at 100K. Diffraction images were processed and scaled with the XDS/XSCALE package (Kabsch, 2010) (Table 2). The CsI/NaI derivative model was used for molecular replacement with Molrep (Vagin and Teplyakov, 2010). The model was corrected manually using COOT (Emsley *et al*., 2010)

alternating with autoBUSTER refinements (Blanc *et al.*, 2004) (Table 2). The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 5LY8).

Electron microscopy negative straining structures of the virion's baseplate

About 6 μl of a purified J-1 virus (10^{10} PFU ml^{-1}) were deposited onto a glow-discharged carbon-coated grid and incubated for one minute. Sample excess was blotted off, rinsed twice with water and stained with 10 μl of 1% uranyl acetate for 30 sec. Micrographs (200) were recorded on a 2K × 2K FEI Eagle CCD camera using a Tecnai Spirit electron microscope operated at 120 kV and a magnification of 110,000× (resulting in a pixel size of 4.83 Å/pixel). The three-dimensional reconstruction was produced using a single particle procedure and the XMIPP software package (Sorzano *et al.*, 2004). A total of 865 particles were manually boxed around the baseplate and subjected to maximum likelihood (ML) classification and alignment implemented in Xmipp (Scheres *et al.*, 2008), imposing a 6-fold symmetry. The initial volume was determined using a random sample consensus (RANSAC) approach (De la Rosa-Trevin *et al.*, 2013) with three 2D classes (Supporting Information Fig. S6a). The resolution of the final volume was estimated at 20 Å using the Fourier Shell Correlation 0.5 criterion (Supporting Information Fig. S6b). The EM map reported in this article has been deposited in the EMDataBank database, www.emdatabank.org (accession number EMD-4150).

Funding

The work was in part supported by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01 to CC and by UBACYT 2014-2017 GC 20020130100444BA to MP. M.E.D. is a doctoral fellow of Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET, Argentina) and was also supported by BecAr-Campus France fellowship.

Acknowledgements

We thank the ERSF and Soleil Synchrotron radiation facilities for beamline time allocation.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**: 213–221.

Ainsworth, S., Sadovskaya, I., Vinogradov, E., Courtin, P., Guerardel, Y., Mahony, J., *et al.* (2014) Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. *MBio* **5**: e00814–e00880.

Aslanidis, C., and de Jong, P.J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18**: 6069–6074.

Bebeacua, C., Bron, P., Lai, L., Vegge, C.S., Brondsted, L., Spinelli, S., *et al.* (2010) Structure and molecular assignment of lactococcal phage TP901-1 baseplate. *J Biol Chem* **285**: 39079–39086.

Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S.M., and Bricogne, G. (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr D Biol Crystallogr* **60**: 2210–2221.

Capra, M.L., Quiberoni, A., and Reinheimer, J. (2006) Phages of Lactobacillus casei/paracasei: response to environmental factors and interaction with collection and commercial strains. *J Appl Microbiol* **100**: 334–342.

De la Rosa-Trevin, J.M., Oton, J., Marabini, R., Zaldivar, A., Vargas, J., Carazo, J.M., and Sorzano, C.O. (2013) Xmipp 3.0: an improved software suite for image processing in electron microscopy. *J Struct Biol* **184**: 321–328.

Desmyter, A., Farenc, C., Mahony, J., Spinelli, S., Bebeacua, C., Blangy, S., *et al.* (2013) Viral infection modulation and neutralization by camelid nanobodies. *Proc Natl Acad Sci U S A* **110**: E1371–E1379.

Dieterle, M.E., Bowman, C., Batthyany, C., Lanzarotti, E., Turjanski, A., Hatfull, G., and Piuri, M. (2014a) Exposing the secrets of two well-known Lactobacillus casei phages, J-1 and PL-1, by genomic and structural analysis. *Appl Environ Microbiol* **80**: 7107–7121.

Dieterle, M.E., Jacobs-Sera, D., Russell, D., Hatfull, G., and Piuri, M. (2014b) Complete genome sequences of Lactobacillus Phages J-1 and PL-1. *Genome Announc* **2**: e00998.

Dieterle, M.E., Fina Martin, J., Duran, R., Nemirovsky, S.I., Sanchez Rivas, C., Bowman, C., *et al.* (2016) Characterization of prophages containing 'evolved' Dit/Tal modules in the genome of Lactobacillus casei BL23. *Appl Microbiol Biotechnol* **100**: 9201–9215.

Emsley, P., Lohkamp, B., Scott, W.G., and 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., and 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., and 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.

Goulet, A., Lai-Kee-Him, J., Veesler, D., Auzat, I., Robin, G., Shepherd, D.A., *et al.* (2011) The opening of the SPP1 bacteriophage tail, a prevalent mechanism in Gram-positive-infecting siphophages. *J Biol Chem* **286**: 25397–25405.

Hino Mai, N. (1965) Lactic Acid Bacteria employed for beverage production. II. Isolation and some properties of a

bacteriophage isolated during the fermentation of lactic acid beverage. *J Chem Soc Jpn* **39**: 472–476.

Holm, L., Kaariainen, S., Rosenstrom, P., and Schenkel, A. (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**: 2780–2781.

Hudson, K.L., Bartlett, G.J., Diehl, R.C., Agirre, J., Gallagher, T., Kiessling, L.L., and Woolfson, D.N. (2015) Carbohydrate-aromatic interactions in proteins. *J Am Chem Soc* **137**: 15152–15160.

Kabsch, W. (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**: 125–132.

Kanamaru, S., Leiman, P.G., Kostyuchenko, V.A., Chipman, P.R., Mesyanzhinov, V.V., Arisaka, F., and Rossmann, M.G. (2002) Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**: 553–557.

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

Legrand, P., Collins, B., Blangy, S., Murphy, J., Spinelli, S., Gutierrez, C., *et al*. (2016) The atomic structure of the phage tuc2009 baseplate tripod suggests that host recognition involves two different carbohydrate binding modules. *MBio* **7**: 15.

Mahony, J., Kot, W., Murphy, J., Ainsworth, S., Neve, H., Hansen, L.H., *et al*. (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., McDonnell, B., Casey, E., and van Sinderen, D. (2016) Phage-host interactions of cheese-making lactic acid bacteria. *Annu Rev Food Sci Technol* **7**: 267–285.

Mahony, J., Oliveira, J., Collins, B., Haanemaaijer, L., Lugli, G.A., Neve, H., *et al*. (2017) Genetic and functional characterisation of the lactococcal P335 phage-host interactions. *BMC Genom* **18**: 146.

Murphy, J., Bottacini, F., Mahony, J., Kelleher, P., Neve, H., Zomer, A., Nauta, A., and van Sinderen, D. (2016) Comparative genomics and functional analysis of the 936 group of lactococcal Siphoviridae phages. *Sci Rep* **6**: 21345.

Otwinowski, Z. (1993) DENZO: oscillation data and reducing program. In *Data Collection and Processing*. Sawyer, L., Isaacs, N.W., and Bailey, S. (eds). Warrington, UK: SERC Daresbury Laboratory.

Pape, T., and Schneider, T.R. (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J Appl Crystallogr* **37**: 853–844.

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

Scheres, S.H., Nunez-Ramirez, R., Sorzano, C.O., Carazo, J.M., and Marabini, R. (2008) Image processing for electron microscopy single-particle analysis using XMIPP. *Nat Protoc* **3**: 977–990.

Sciara, G., Bebeacua, C., Bron, P., Tremblay, D., Ortiz-Lombardia, M., Lichiere, J., *et al*. (2010) Structure of lactococcal phage p2 baseplate and its mechanism of activation. *Proc Natl Acad Sci U S A* **107**: 6852–6857.

Sheldrick, G.M. (2008) A short history of SHELX. *Acta Crystallogr A* **64**: 112–122.

Soding, J., Biegert, A., and Lupas, A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**: W244–W248.

Sorzano, C.O., Marabini, R., Velazquez-Muriel, J., Bilbao-Castro, J.R., Scheres, S.H., Carazo, J.M., and Pascual-Montano, A. (2004) XMIPP: a new generation of an open-source image processing package for electron microscopy. *J Struct Biol* **148**: 194–204.

Spinelli, S., Desmyter, A., Verrips, C.T., de Haard, H.J., Moineau, S., and Cambillau, C. (2006) 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., and Cambillau, C. (2014) Structures and host-adhesion mechanisms of lactococcal siphophages. *Front Microbiol* **5**: 3.

Taylor, N.M., Prokhorov, N.S., Guerrero-Ferreira, R.C., Shneider, M.M., Browning, C., Goldie, K.N., Stahlberg, H., and Leiman, P.G. (2016) Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature* **533**: 346–352.

Unger, T., Jacobovitch, Y., Dantes, A., Bernheim, R., and Peleg, Y. (2010) Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. *J Struct Biol* **172**: 34–44.

Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**: 22–25.

Veesler, D., and 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., and Cambillau, C. (2010) 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., *et al*. (2012) Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. *Proc Natl Acad Sci U S A* **109**: 8954–8958.

Vinogradov, E., Sadovskaya, I., Grard, T., and Chapot-Chartier, M.P. (2016) Structural studies of the rhamnose-rich cell wall polysaccharide of Lactobacillus casei BL23. *Carbohydr Res* **435**: 156–161.

Yokokura, T. (1971) Phage receptor material in Lactobacillus casei cell wall. I. Effect of L-rhamnose on phage adsorption to the cell wall. *Jpn J Microbiol* **15**: 457–463.

AUTHOR QUERY FORM

Dear Author,
During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof.
Many thanks for your assistance.

Query References	Query	Remarks
AQ1	Please provide the Fax and Tel numbers for the corresponding authors.	
AQ2	Please confirm that given names (red) and surnames/family names (green) have been identified correctly.	