A Topological Model of the Baseplate of Lactococcal Phage Tuc2009*

Giuliano Sciara†, Stéphanie Blangy†, Marina Siponen‡, Stephen McGrath§, Douwe van Sinderen‡, Mariella Tegoni†, Christian Cambillau‡, and Valérie Campanacci‡

From the †Architecture et Fonction des Macromolécules Biologiques, UMR 6098 CNRS and Universités Aix-Marseille I & II, 13288 Marseille, France, the ‡Department of Microbiology, National University of Ireland, Cork, Ireland, and the §Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland

Phages infecting Lactococcus lactis, a Gram-positive bacterium, are a recurrent problem in the dairy industry. Despite their economical importance, the knowledge on these phages, belonging mostly to Siphoviridae, lags behind that accumulated for members of Myoviridae. The three-dimensional structures of the receptor-binding proteins (RBP) of three lactococcal phages have been determined recently, illustrating their modular assembly and assigning the nature of their bacterial receptor. These RBPs are attached to the baseplate, a large phage organelle, located at the tip of the tail. Tuc2009 baseplate is formed by the products of 6 open read frames, including the RBP. Because phage binding to its receptor induces DNA release, it has been postulated that the baseplate might be the trigger for DNA injection. We embarked on a structural study of the lactococcal phages baseplate, ultimately to gain insight into the triggering mechanism following receptor binding. Structural features of the Tuc2009 baseplate were established using size exclusion chromatography coupled to on-line UV-visible absorbance, light scattering, and refractive index detection (MALS/UV/RI). Combining the results of this approach with literature data led us to propose a “low resolution” model of Tuc2009 baseplate. This model will serve as a knowledge base to submit relevant complexes to crystallization trials.

The very first event in viral infection is the attachment of the virion particle to its host. This occurs through the recognition of a host cell wall receptor, which can be a polysaccharide or a protein (or both), by the viral receptor-binding protein (RBP)3.

§ This work was supported by the Marseille-Nice Génopole, the CNRS, and through Science Foundation Ireland Investigator Grant 02/IN.1/8198 (to D. v. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 To whom correspondence may be addressed: AFMB, 163 Av. de Luminy Case 932, 13288 Marseille Cedex 09, France. E-mail: cambillau@afmb.univ-mrs.fr.

2 To whom correspondence may be addressed: AFMB, 163 Av. de Luminy Case 932, 13288 Marseille Cedex 09, France. Fax: 33-491-266-720; E-mail: valerie.campanacci@afmb.univ-mrs.fr.

3 The abbreviations used are: RBP, receptor binding protein; SEC, size exclusion chromatography; SLS, static light scattering; MALS/UV/RI, on-line multi-angle laser light scattering, absorbance, and refractive index; MTP, major tail protein; TMP, tape measure protein; Tal, tail-associated lysozyme; Dit, Distal tail protein; IC, initiation complex; BP, baseplate; MS, mass spectrometry; ORF, open reading frame; BppU, upper baseplate protein; BppL, lower baseplate protein; BppA, baseplate-associated protein; HPLC, high pressure liquid chromatography; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol.
It has been shown by transmission EM on TP901-1 that BppU forms the upper disk of the baseplate and is required for the fixation of BppL, which represents the lower disk (5, 6). Therefore BppU and BppL should be expected to interact directly. BppL is the RBP that specifically recognizes the host, as demonstrated by experiments in which BppL from TP901-1 was swapped with that of Tuc2009, resulting in host specificity demonstrated by experiments in which BppL from TP901-1 was swapped with that of Tuc2009, resulting in host specificity.

Determination of the molecular weights of proteins or their complexes can be achieved by using different techniques such as size exclusion chromatography (SEC), native polyacrylamide gel electrophoresis, mass spectrometry, light scattering, and analytical centrifugation. Among these, SEC is a simple, fast and non-destructive method for estimating the molecular weight of a protein in its native form, based on a calibration curve obtained with protein standards. However, the elution position depends not only on the molecular weight but also on the shape of the protein, and its tendency to interact with the column matrix. To overcome these limitations, SEC can be used in combination with on-line multangle laser light scattering and refractometry. The molecular weight from this measurement is then independent of the elution order, permitting determination of the molecular weight of simple proteins, protein complexes, and their stoichiometry.

The on-line SEC-MALS/UV/RI used in this study is a system provided by Wyatt Technology (Santa Barbara, CA). It consists of a combination of three detectors placed in series after an SEC column: a static light scattering detector (SLS), a dynamic light scattering detector, and a refractometry detector. On top of this, multwavelength on-line UV-visible absorbance is provided by the HPLC system. SLS measures the intensity of the scattered light as a function of angle. SLS measurements can yield the molar mass and root mean squared radius (Rg).

Refractive index detector measures the concentration of protein in solution, the dn/dc (specific refractive index increment), and the absolute refractive index of the solution. Measurement of pro-
tein concentration is essential for the absolute characterization of the molar mass (14).

Using this HPLC/SEC-MALS/UV/RI system, we determined the self-association and the inter-association of the baseplate components, whereas we also measured their hydrodynamic radii. These data, together with some input from literature, made it possible to propose a topological model of the baseplate of Tuc2009, as well as an evaluation of the stoichiometry of the interacting components. We propose this model as a paradigm for the P335 phage species baseplate, with possible extension to other phages infecting Gram-positive bacteria.

**EXPERIMENTAL PROCEDURES**

Cloning, Expression, and Purification—Cloning of full-length ORF49, ORF52, and ORF53 in pQE60 were previously described by Mc Grath et al. (7). The DNA regions that represent the N-terminal portion of ORF50 (residues 1–590), the head-only fragment of ORF53 (residues 58–172) were amplified using PCR from the Tuc2009 bacteriophage genome using specific Gateway™ primers containing attB sequences at both ends, a ribosome-binding site, and a C-terminal His6 tag coding sequence. A bicistronic ORF51-ORF53 encoding construct was obtained using a two-step PCR: first, the coding regions of ORF51 and ORF53 were separately amplified using primers upstream from each ORF and allowing the presence of a sequence overlapping the two constructs; second, these two PCR products were used as DNA template for a second amplification with primers P1/INT-R and INT-F/P2, respectively (primer sequences are detailed in Table 2), to introduce a ribosome-binding site upstream from each ORF and allowing the presence of a sequence overlapping the two constructs; second, these two PCR products were used as DNA template for a second amplification with primers P1 and P2. The final construct consisted of DNA specifying a C-terminal His$_6$-tagged ORF51 upstream of DNA encoding a C-terminal Streptag-tagged ORF53. The amplicons representing ORF50-(1–590), ORF53-(17–172), ORF53-(58–172) and the bicistronic ORF51-ORF53 were cloned by recombination in a Gateway™ pDEST14 vector (Invitrogen). Protein expression experiments were carried out using M15pREP4 (Qiagen) and Rosetta(DE3)pLysS (Novagen) strains for pQE60 and pDEST14 constructs, respectively. After an overnight induction with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside at 25 or 17 °C depending on the target, cells were harvested by centrifugation for 10 min at 4000 g. Material pellets were resuspended in 40 ml/liter of culture of lysis buffer (Tris 50 mM, NaCl 300 mM, imidazole 10 mM, pH 8.0) supplemented with 0.25 mg/ml lysozyme, 1 μg/ml DNase, 20 mM MgSO$_4$, and antiproteases (Complete EDTA-free antiproteases, Roche) and frozen at −80 °C. After thawing and sonication, lysates were cleared by a 30-min centrifugation at 12,000 × g. Overexpressed proteins were purified on a Pharma- cia Äkta FPLC by nickel affinity chromatography (His-Trap 5-ml column, GE Healthcare) using a step-gradient of imidazole followed by a preparative Superdex 200 HR26/60 gel filtration in 10 mM Hepes, 50 mM NaCl, pH 7.5. An additional Strept-Tac affinity purification (IBA BioTAGnology) was used for the purification of the ORF51-ORF53 dual protein complex. Purified proteins were concentrated using Amicon Ultra-15 ml (Millipore) and characterized by SDS-PAGE, matrix-assisted laser desorption ionization time-of-flight mass spectrometry (Bruker Autoflex), trypsin peptide mass fingerprint, far-UV circular dichroism (Jasco J-810), and dynamic light scattering (Zetasizer Nano-S, Malvern).

**Blue Native PAGE**—Blue native PAGE (15, 16) was performed using a NativePAGE™ Novex® 4–16% BisTris gel (Invitrogen). Gels were run following the supplier’s instructions loading 5 μg of each protein. Native gels were stained with Coomassie® R-250. The NativeMark™ Unstained Protein Standard (Invitrogen) was used to estimate the molecular weight of proteins after native electrophoresis by plotting $R_f$ (retardation factor) values versus log$_{10}$ molecular weight.

**SEC with On-line Multitangle Laser Light Scattering, Absorbance, and Reflective Index (MALS/UV/RI) Detectors**—SEC was carried out on an Alliance 2695 HPLC system (Waters) using a Silica Gel KW804 column (Shodex) eluted with 10 mM Hepes and 50 mM NaCl at pH 7.5 at a flow of 0.5 ml/min. Detection was performed using a triple-angle light scattering detector (MiniDAWN™ TREOS, Wyatt Technology), a quasi-elastic light scattering instrument (Dynapro™, Wyatt Technology), and a differential refractometer (Optilab® REX, Wyatt Technology). Molecular weight and hydrodynamic radius determination was performed by the ASTRA V software (Wyatt Technology) using a dn/dc value of 0.185 ml/g. Proteins were loaded at a final concentration of 0.02 mM.

**RESULTS**

Production of the Baseplate Proteins—We undertook the production of seven tail and baseplate proteins from Tuc2009: ORF45 (MTP), ORF48 (TMP), ORF49 (Dit), ORF50 (Tal), ORF51 (BppU), ORF52 (BppA), and ORF53 (BppL, representing the RBP). All these proteins could successfully be overproduced either as full-length products, or as individual domains, using our previously published screening procedure (17, 18). Except in the case of TMP, all proteins are soluble up to several milligrams/ml, making it possible to perform biophysical characterization and interaction studies on them.

ORF45 (MTP) was expressed as a soluble protein and could be concentrated up to 10 mg/ml. However, no further exper-
ments have been performed on it. ORF48 (TMP) is a 1025-residue long protein, which has been demonstrated to control the length of the phage tail (11). Secondary structure predictions suggest that TMP consists of two domains: a long, almost continuous, $\alpha$-helix, spanning residues 1 to 960, and a putative mixed $\alpha/\beta$ domain from residue 961 to the C terminus. It has been suggested that the long $\alpha$-helix of TMP drives the MTP assembly and interacts at its N terminus with the phage portal, whereas the C-terminal domain, possibly more globular, may interact with the other components of the IC, Dit, and/or Tal (6, 7, 12). We cloned two DNA fragments that encompass different sections of TMP: the full-length protein and the 961–1025 C-terminal domain. The former construct was not expressed in any of our screening conditions, whereas the latter was expressed as inclusion bodies, which could not be refolded using our standard procedures (18). ORF49 (Dit) has been proposed to orchestrate the IC formation once a certain concentration of Dit was reached (7). Dit was expressed as a soluble protein at concentrations up to 6 mg/ml. However, this protein appeared to be susceptible to protease cleavage, because several time-dependent degradation products appeared on SDS gels. ORF50 (Tal) incorporates a self-cleavage site at a consensus sequence GGS$$GG$$, between the two last Gly (7, 13). The N terminus, encompassing residues 1–590, was proposed to act as a structural domain, whereas the C-terminal domain of this tail fiber (represented by amino acid residues 591–906), bears the catalytic activity: the last 120 residues of its C terminus domain share ~40% sequence identity with the catalytic domain of a zinc-containing peptidase from Vibrio cholerae, the structure of which has been determined (PDB 2GU1).4 We therefore cloned the two domains of Tal using the Gateway™ system. Both N- and C-terminal Tal domains are well expressed and soluble. The N-terminal domain could be concentrated up to 7 mg/ml, whereas the C-terminal domain was not subjected to further studies. These three components, TMP C-term, Dit, and Tal N-term, are believed to form the conical structure of the IC. In TP901-1, two proteins have been shown to assemble stepwise around the IC: first BppU, followed by BppL (5, 6). In Tuc2009, it was assumed that the assembly of these proteins (represented by ORF51 and -53, respectively) may follow the same order (7). However, the situation in Tuc2009 is more complex because ORF52, which is absent in TP901-1, is probably associated with the IC structure. The orf52 gene is located between orf51 and orf53 on the phage genome. Its protein product, ORF52, has been identified among the phage structural proteins, but its exact localization could not be assigned using immuno-gold EM (7). Its involvement in the peripheral baseplate is therefore hypothetical. We were able to express ORF51 (BppU) as a soluble protein at a very high concentration of 50 mg/ml. ORF52 (BppA) is also well expressed and is soluble up to 10 mg/ml. These two proteins are stable and no time-dependent cleavage/degradation was detected. ORF53 or BppL acts as the RBP, conferring host specificity (11). We cloned the gene encoding ORF53 (BppL) and expressed it as a soluble protein. It was impossible, however, to obtain the intact protein, due to rapid proteolytic deterioration, probably at the N terminus, yielding shorter species. Protein cleavage has been observed for BppL of TP901-1, because its crystal structure lacks the first 15 residues (8).

Characterization of the Oligomerization State of the Tuc2009 Baseplate Proteins—The purified Tuc2009 proteins were subjected to weight and size analysis using the MALS/UV/RI Wyatt instrument (Table 1). ORF49 (Dit) has a theoretical mass of 30,152 Da, and an MS measured mass of 30,100 Da. The MALS/UV/RI analysis generated a measured mass of 29,000 Da, indicating that the molecule is monomeric under the conditions tested, with an $R_n$ value of 3.3 ± 1.0 nm. The ORF50 (Tal) N-terminal domain (1–590) has a theoretical mass of 67,100 Da, and an MS measured mass of 66,800 Da. The MALS/UV/RI measured mass of 190 kDa, which is close to the theoretical value for a monomer (200 kDa), with an $R_n$ value of 4.0 ± 1.0 nm. For ORF51 (BppU), a unique peak was observed in the chromatogram, at a mass of 109 kDa. The theoretical mass of ORF51 being 37,350 Da, ORF51 thus appears to act as a trimer in solution (theoretical value of 112 kDa). The $R_n$ value measured for this peak is 4.5 ± 0.3 nm. ORF52 (BppA), instead, appears as a monomer of 29,000 Da by MALS/UV/RI, a value in line with its theoretical mass of 32,920 Da (MS: 32,330 Da). An $R_n$ of 2.8 nm was measured for this protein. ORF53 (RBP, BppL) has a theoretical mass of 20,214 Da (MS: 20,219 Da), and appears as a protein with a mass of 60 kDa by MALS/UV/RI, a value pointing to a trimeric state. Because this protein is prone to proteolytic cleavage, the quality of the chromatographic data deteriorated within a few days, ultimately becoming uninterpretable. We expressed a variant of ORF53 that comprises the head domain only. This head domain is measured at 42,700 Da, compared with a theoretical mass of the monomer of 13,838 Da. These data clearly indicate that the trimeric state of the full-length protein is maintained in these constructs, as seen in phase p2 (9). The $R_n$ value of the head trimeric domain is 2.8 ± 0.3 nm.

Interaction Studies Using MALS/UV/RI—Following these experiments on the individual proteins, we performed pairwise interaction analysis. When ORF51 and ORF52 were mixed, a new peak appeared with a mass of 215 (+15) kDa, which is in agreement with the theoretical mass of a heterohexamer of 3xORF51 + 3xORF52 (210,810 Da) (Fig. 2). In contrast, mixing ORF51 and ORF53 did not yield any change in the chromatograms, as compared with the individual components. Indeed, this was very surprising, because we anticipated an interaction between ORF51 and ORF53, based on the results obtained for TP901-1 (5, 6). This result can be explained by assuming that the observed cleavage of the N-terminal trimeric $\alpha$-helix bundle, which anchors the RBP in the BP, prevents interaction with ORF51. Finally, mixing ORF52 and ORF53 also did not yield any change in the chromatograms.

To overcome the presumed cleavage problem occurring with ORF53, which might have caused the lack of interaction with ORF51, we co-expressed both proteins from a bicistronic construct (see “Experimental Procedures”). In the latter construct, the ORF51 protein was expressed to include a His$_6$ C-term tag, whereas ORF53 included a STREP C-term tag, making it possible to achieve rapid purification of the putative complex. This

---

4 R. Sugadev, D. Kumaran, S. Swaminathan, and S. K. Burley, unpublished data.
FIGURE 1. Gels of Tuc2009 baseplate proteins. The numbers on top of the gels refer to the ORFs. The ruler occupies the left lanes. A, SDS gels after affinity and gel filtration. B, native blue gels of complexes of ORFs belonging to the peripheral baseplate.

complex was indeed isolated and analyzed using the MALS/UV/RI. The chromatogram displayed the expected peak at 170 kDa of a protein complex consisting of 3xORF51 + 3xORF53, and another peak at 370 kDa, probably comprising a dimer of this aforementioned complex, i.e., $2 \times (3x\text{ORF51} + 3x\text{ORF53})$ (Fig. 2). In this case, most of ORF53 remains intact, where its helix bundle is apparently protected through enclosure within ORF51 (Fig. 1a). However, the main peak lacks symmetry, indicating the presence of unbound ORF51, a possible trimer of ORF51 and a minor peak of possibly three shoulder + neck domains of ORF53 attached to the ORF51 trimer. When we added ORF52 to this ORF51-ORF53 complex, we observed a major band of 255 kDa representing the complex of the three components, with a 3:3:3 ratio. We also observed a lower molecular weight shoulder that we interpret as a 3:3:3 complex in which the head domain of ORF53 has been cleaved, as well as a smaller band of higher molecular mass ($\sim 520$ kDa) corresponding to a dimer of the nonamer complex, $2 \times (3x\text{ORF51} + 3x\text{ORF52} + 3x\text{ORF53})$ (Fig. 2). When we tried mixing proteins from the initiation complex, ORF49 (Dit) and ORF50N (Tal, amino residues 1–590), we did not observe a new peak, indicating that no interaction occurs, which may be explained by the absence of the TMP component of the assumed initiation complex.

**Interaction Analysis Using Blue Native Gels**—We performed blue native gradient gel (polyacrylamide 4–16%) analysis of the components of the peripheral baseplate (Fig. 1b). In lane 51, ORF51 exhibits three bands, whereas on SDS gels (Fig. 1a) ORF51 appears as a unique band. Furthermore, the apparent mass of the first band, calculated using the calibration curve with standards (see “Experimental Procedures”), is much larger than expected, 140 versus 100 kDa. We suspect that ORF51 does not migrate at its expected size. Using the calibration curve, the molecular masses for the second and third bands are 295 and 437 kDa, respectively, i.e., those of a hexamer (280 kDa expected) and a nonamer (440 kDa expected). This autoassociation suggests that ORF51 is responsible for building the framework for the peripheral baseplate.

The co-expressed complex of ORF51 and ORF53 (lane 51/53) presents a more complex pattern of several bands. The 51/52) exhibits, starting from the bottom, a band at 322 kDa, versus the expected 210 kDa. The next band is observed at 620 kDa and would represent a dimer of the former 322-kDa complex. The complex between ORFs 51, 52, and 53 (lane 51/52/53) exhibits a major band of apparent molecular mass of 400 kDa instead of 279 kDa. The next band, at 740 kDa would represent the dimer of this complex.

To correct for the anomalous behavior of ORF51 on these gels, we established an internal calibration curve based on the weights of ORF51 trimer, hexamer, and nonamer. In this case, the calculated molecular masses for complexes ORF51 + 53, ORF51 + 52, and ORF51 + 52 + 53 are 188, 240, and 300 kDa, respectively, which is much closer to the theoretical values of 170, 210, and 270 kDa.

**DISCUSSION**

**The Peripheral Baseplate**—Most EM pictures of TP901-1 or Tuc2009 baseplate display lateral views of the baseplate, in a direction perpendicular to the tail axis (5–7). In EM views of TP901-1, the baseplate appears to be crushed and its 6 RBPs (BppL) bulbs and peduncles can be identified in the EM image plane (5, 6). The x-ray structure of TP901-1 RBP fits perfectly well into these low-resolution images once adjusted to the scale of the EM picture (8). Interestingly, in pictures in which the baseplate was rotated by 90°, the baseplate presents a clear 6-fold symmetry. Although the available pictures are less clear for Tuc2009, the same conclusions should remain valid because the BP proteins of the two phages display 78–97% identity, if we exclude their BppL. Analysis of Tuc2009 EM images made it possible to deduce the gross dimensions of the baseplate (7). The upper disk, formed by BppL and possibly BppA, has a diameter of 27 nm and a height of 7.5 nm, whereas the lower disk, the petticoat structure, is 21 nm wide and 7.5 nm high. In contrast, the TP901-1 upper disk is 23 nm wide, and the lower disk is wider (28 nm). In this latter phage, the IC conical structure is 7 nm high and the fiber extends 16 nm out of the conical structure (6).
We have shown that the three proteins constituting the peripheric baseplate, ORF51, ORF52, and ORF53, form a homotrimeric complex of $3 \times (\text{ORF51 + ORF52 + ORF53})$. This ensemble has a $R_h$ of 6.2 nm, hence a diameter of 12.4 nm, a value much lower than the diameter of 27 nm measured on the EM images of Tuc2009 (7). This ensemble cannot, therefore, constitute alone the baseplate disks. Because contrast EM pictures have revealed a 6-fold symmetry for the baseplate, it is likely that the baseplate disks consist of a hexamer of ORF51 $\times H1001$ ORF52 $\times H1001$ ORF53. Straightforward geometrical calculations from the measured $R_h$ of this ensemble yield a baseplate diameter value comprised between 32 and 36 nm, a value higher but compatible with the EM measurement (27 nm). This slight difference between solution data and EM might be explained in two ways: (i) the proteins interpenetrate, or (ii) they are more elongated in the direction perpendicular to the baseplate disk. This latter explanation seems the most probable because ORF53 points downwards from the baseplate and therefore does not increment the $R_h$ horizontally (Fig. 3). A $R_h$ of 4.8 nm, slightly larger than the $R_h$ of ORF51 alone (4.5 nm) would lead to a baseplate dimension of 26.4 nm, close to the EM value, indicating thus that ORF52 participates little in the horizontal component of the $R_h$.

Taking these considerations into account, as well as the observed $R_h$ of ORF53 head trimer (2.8 nm), we reach a lower disk diameter of 19–21 nm (Fig. 3), in agreement with the EM data (21 nm) (7). With this model, however, the upper disk is thicker than observed by EM, whereas the lower disk thickness fits with EM data. Therefore, we have to assume that the ORF51 trimer is not spherical, but is more likely to resemble a flatter ellipsoid 7.5 nm thick, and somewhat larger than 24–27 nm. Because ORF52 is not present in TP901-1, its presence in Tuc2009 is intriguing. ORF52 clearly participates in the upper disk of the baseplate, because we have shown that it forms a strong 3:3 complex with ORF51, but does not interact with ORF53 alone. ORF52 orthologs are present in several bacteriophages, whereas domains of other phage proteins involved in host recognition also display significant similarity with this ORF.

The Initiation Complex—The recent structure of the tail tip of SPP1 determined at 9-Å resolution by cryoEM, provides a wealth of information on interactions in the initiation complex, considering the significant homology between its proteins and those of Tuc2009 (12). In SPP1, the cap (TMP C-term), Dit, and Tal form, in this order, an elongated structure of 35 nm with a width that ranges between 3 and 14 nm (12). Both Dit and Tal were assigned to be present as trimers, whereas TMP might preferably be a hexamer. The SPP1 Tal N-terminal domain exhibits structural resemblance with a...
protein of known structure, the trimeric P22 tail spike (1TSP) (19). This protein is formed by three /H9252 -helical mono-
mers that are assembled tightly. Our observation of a Tal trimer in solution is consistent with the cryoEM data on
SPP1. DitSPP1 is also interpreted to be present as a trimer in the cryoEM map. However, under our experimental condi-
tions, Dittuc2009 is a monomer, indicating that the interaction is less tight. It seems therefore that Tal-(1–590) promotes
the trimeric assembly of Dit, as for ORF51-ORF52 and also described in the case of gp4 of phage P22, which exists as a
monomer in solution and will only oligomerize upon binding to the portal protein (20).

Immuno-gold labeling experiments have failed to label Taltuc2009, suggesting that this trimer is masked from the sol-
vent by the peripheral BP components. We measured a \( R_g \) of 4.0 nm for Tal. The space within the ORF51 18-mer upper disk
structure is estimated to be \( \sim 9–10 \) nm in width, which would therefore be large enough to accommodate the Tal trimer (Fig. 3, top). A side view of the resulting assemblage suggests that Tal is fully buried, being inaccessible to large IgGs (Fig. 3). In con-
trast, both Dit and TMP can be visualized by immuno-gold labeling, but phages lacking the peripheral components of the
BP exhibited stronger labeling. This suggests that both TMP and Dit are left partially accessible to solvent by the peripheral
baseplate.

The Mass of the Tail and Baseplate—Knowing the composition of the tail and the baseplate, from literature or the present
data, we can evaluate the mass of Tuc2009 organelles. The tail has the largest mass (3.4 MDa) counting only the MTP, but 4.0
MDa if we assign the TMP to the tail. The baseplate weight is 2.6 MDa (Table 3).

Evolutionary Considerations—The role of Tal illustrates interesting aspects of the mosaic assembly of phage proteins. B. subtilis
phage SPP1 has been shown to recognize its host through an interaction of its tail tip, probably involving Tal,
with YueB, a large membrane protein (21). Tal-(1–590), its structural domain, is similar between Tuc2009 and SPP1 (12).
In contrast, the Taltuc2009 C terminus differs from that of SPP1 in several ways: (i) it is capable of self-cleavage at a consensus
site, and both forms, cleaved and uncleaved are observed in the phage particle, (ii) it has a clearly identified zinc-dependent
protease domain at its tip, and (iii) it has not been shown to interact with any protein. These facts indicate that, besides a
common scaffold function of the Tal N terminus in each of those phages, the C terminus might fulfill very different func-
tions, adhesion and DNA injection initiation for one group, cell lysis for the other.

The tail tip constitutes the base of the phage tail and is the minimal structure observed on Siphoviridae such as SPP1
and T5, whereas in other Siphoviridae much larger append-
geages decorate this scaffold. An intriguing question is why some phages have developed these appendages, whereas oth-
ers function satisfactorily with the minimal tail tip. It seems that although similar, Siphoviridae have developed different
strategies to attach specifically to their hosts. In phage T5, the antireceptor, the monomeric gp5, associates tightly (sub-
nanomolar) with a siderophore transport porin, FhuA (22). This mechanism is observed with two other phages, BF23,
whose antireceptor binds to BtuB (23), a B12 vitamin trans-
porter, and H8, which attaches to FepA, another iron trans-
porter (24). The lactococcal phage c2 has been shown as well

### Table 3

| Tuc-ORF/Source | Subunits | Total \( M_r \) | Total \( M_s \) |
|---------------|----------|----------------|----------------|
| MTP | 34 × 6 | 7, 12 | 204 | 16.8 | 3427.2 |
| BP-IC | | | | | |
| TMP | 6 | 12 | 6 | 110.26 | 661.5 |
| Dit | 3 | 12 | 29.12 | 87.4 |
| Tal | 3 | 12 and this work | 3 | 100.9 | 302.7 |
| Total | | | | 1051.6 |
| BP disks | | | | | |
| BppU | 6 × 3 | This work | 18 | 36.4 | 655.2 |
| BppA | 6 × 3 | This work | 18 | 31.85 | 573.3 |
| BppL | 6 × 3 | This work | 18 | 18.88 | 339.8 |
| Total | | | | 1568.3 |
| Total BP | | | | | |
| Total | | | | 2619.9 |
| Total | | | | 6047.1 |

**FIGURE 3.** Topological model of Tuc2009 tail-tip and baseplate. Top, view perpendicular to the main tail axis. The 3:3:3 complex formed between ORFs 51, 52, and 53 is assembled in a large hexagonal structure. The center of the structure is occupied by the initiation complex, and more probably interacts strongly with ORF50-(1–590), which is non-accessible to specific antibodies (7). Bottom, view parallel to the main tail axis. The geometry of the initiation complex (ORFs 48, 49, and 50) is derived from Plisson et al. (12) and from stoichiometries obtained here (Table 1). The topology of the peripheral baseplate is calculated based on our MALS/UV/RI analysis.
to recognize its host through an antireceptor/protein contact, very similar to the mechanism shown for SPP1 (21). Its receptor, named phage interacting protein, exhibits ~25% identity with YueB, the SPP1 receptor (25). Both proteins possess a hexahelical-membrane embedded domain, and a very large ektodomain of ~900 residues. However, their physiological function remains unknown.

In contrast to phages that bind membrane-attached proteins, it has been proposed that most lactococcal phages that belong to the 936 or P335 families exclusively attach to saccharidic receptors (3, 26, 27). This hypothesis has been corroborated by the recent elucidation of the three-dimensional structures of RBPs from representatives of both families (8–10). We have shown that phages p2, bII.190 (936), and phage TP901-1 (P335) bind glycerol molecules in a crevice between the monomers of the trimeric head domain. This glycerol binding site is reminiscent of typical saccharide binding sites (28, 29), with a floor constituted by an aromatic amino acid (Trp, Tyr, or Phe) and a dense network of H-bond binding sites (28, 29, 30). This is also the case with amphiphatic acids. In the case of TP901-1, the K_d values observed between the RBP and MurNac-dipeptide or phosphoglycerol are 4 and 50 nm, respectively. However, these affinities are weaker than those reported for gp5 of phage T5 for FhuA. This domain clearly has a fibronectin fold and might be involved in recognition of host polysaccharides, on top of the already described interaction between the tail tip and the host protein YueB. For lactococcal phages, this two-step mechanism is documented for c2, which seems to recognize its host first by means of a reversible step, which involves a cell-wall rhamnose moiety, followed by an irreversible step that requires phage interacting protein (25). In contrast, host recognition by phages such as Tuc2009 and TP901-1 on the one hand, and sk1 and bII.170 on the other, seems to involve a single step binding to saccharides (3, 26, 27). If this is true, the mechanism of high avidity displayed by Tuc2009 or TP901-1 would be a clever way to overcome the weakness of a single step host recognition mechanism involving rather weak interactions with saccharides.

REFERENCES

1. Dupont, K., Janzen, T., Vogensen, F. K., Josephsen, J., and Stuer-Lauridsen, B. (2004) Appl. Environ. Microbiol. 70, 5825–5832
2. Duplessis, M., and Moineau, S. (2001) Mol. Microbiol. 41, 325–336
3. Dubert, C., Duplessis, M., Moineau, S., and Sinding, D. (2003) J. Bacteriol. 185, 2260–2270
4. Deveau, H., Labrie, S. J., Chopin, M. C., and Moineau, S. (2006) Appl. Environ. Microbiol. 72, 4338–4346
5. Vegge, C. S., Vogensen, F. K., Mc Grath, S., Neve, H., van Sinderen, D., and Brondsted, L. (2006) J. Bacteriol. 188, 55–63
6. Vegge, C. S., Brondsted, L., Neve, H., Mc Grath, S., van Sinderen, D., and Vogensen, F. K. (2005) J. Bacteriol. 187, 4187–4197
7. Mc Grath, S., Neve, H., Seegers, J. F., Eijlander, L., Vegge, C. S., Brondsted, L., Heller, K. J., Fitzgerald, G. F., Vogensen, F. K., and van Sinderen, D. (2006) J. Bacteriol. 188, 3972–3982
8. Spinelli, S., Campanacci, V., Blangy, S., Moineau, S., Tegoni, M., and Cambillau, C. (2006) J. Biol. Chem. 281, 14256–14262
9. Spinelli, S., Desmyter, A., Verrirs, C., de Haard, H. J., Moineau, S., and Cambillau, C. (2006) Nat. Struct. Mol. Biol. 13, 85–89
10. Pedersen, M., Ostergaard, S., Bresciani, J., and Vogensen, F. K. (2000) Virology 276, 315–328
11. Pedersen, M., Ostergaard, S., Bresciani, J., and Vogensen, F. K. (1999) Virology 276, 315–328
12. Plancon, L., Janmot, C., le Maire, M., Desmadril, M., Bonhivers, M., Lemoine, A. S., Bhardwaj, A., Joss, L., Casjens, S., and Cingolani, G. (2007) J. Bacteriol. 189, 4380–4391
13. Kenny, J. G., Mc Grath, S., Fitzgerald, G. F., and van Sinderen, D. (2004) J. Bacteriol. 186, 5818–5824
14. Ye, H. (2006) Anal. Biochem. 356, 76–85
15. Schagger, H., Cramer, W. A., and von Jagow, G. (1994) Anal. Biochem. 217, 220–230
16. Schagger, H., and von Jagow, G. (1991) Anal. Biochem. 199, 223–231
17. Vincentelli, R., Bignon, C., Guerez, A., Cansaan, S., Sunhonen, G., Tegoni, M., Campanacci, V., and Cambillau, C. (2003) Acc. Chem. Res. 36, 165–172
18. Vincentelli, R., Cansaan, S., Offant, J., Cambillau, C., and Bignon, C. (2005) Anal. Biochem. 346, 77–84
19. Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) Science 265, 383–386
20. Olia, A. S., Bhardwaj, A., Joss, L., Casjens, S., and Cingolani, G. (2007) Biochemistry 46, 8776–8784
21. Sao, C., Lhuillier, S., Lu, R., Melki, R., Lepaul, J., Santos, M. A., and Tavares, P. (2006) J. Biol. Chem. 281, 11464–11470
22. Plancon, L., Janmot, C., le Maire, M., Desmadril, M., Bonhivers, M., Letellier, L., and Boulangier, P. (2002) J. Mol. Biol. 318, 557–569
23. Heller, K. J. (1984) Virology 139, 11–21
24. Rabsch, W., Ma, L., Wiley, G., Najar, F. Z., Kaser, W., Schuerr, D. W., Klebba, J. E., Roe, B. A., Laverde, G. J., Schallmey, M., Newton, S. M., and Klebba, J. E. (2007) J. Bacteriol. 189, 5658–5674
25. Montville, M. R., Ardestani, B., and Geller, B. L. (1994) Appl. Environ. Microbiol. 60, 3204–3211
26. Geller, B. L., Ngo, H. T., Mooney, D. T., Su, P., and Dunn, N. (2005) J. Dairy Sci. 88, 900–907
27. Valyasevi, R., Sandine, W. E., and Geller, B. L. (1994) J. Dairy Sci. 77, 1–6
28. Bourne, Y., Mazarier, J., Legrand, D., Rouge, P., Montreuil, J., Spik, G., and Cambillau, C. (1994) Structure 2, 209–219
29. Bourne, Y., Roussel, A., Frey, M., Rouge, P., Fontecilla-Camps, J. C., and Cambillau, C. (1990) Proteins 8, 365–376
30. Leiman, P., Chipman, P. R., Kostychenkov, V. A., Mesyanzhinov, V. V., and Rossman, M. G. (2004) Cell 118, 419–429

P. Tavares, personal communication.