<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>The atomic structure of the phage Tuc2009 baseplate tripod suggests that host recognition involves two different carbohydrate binding modules</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Legrand, Pierre; Collins, Barry; Blangy, Stéphanie; Murphy, James; Spinelli, Silvia; Gutierrez, Carlos; Richet, Nicolas; Kellenberger, Christine; Desmyter, Aline; Mahony, Jennifer; van Sinderen, Douwe; Cambillau, Christian</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2016-01-26</td>
</tr>
<tr>
<td><strong>Type of publication</strong></td>
<td>Article (peer-reviewed)</td>
</tr>
<tr>
<td><strong>Link to publisher's version</strong></td>
<td><a href="http://dx.doi.org/10.1128/mBio.01781-15">http://dx.doi.org/10.1128/mBio.01781-15</a></td>
</tr>
<tr>
<td><strong>Access to the full text of the published version may require a subscription.</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2016 Legrand et al. This is an open-access article distributed under the terms of the Creative Commons Attribution-Noncommercial-ShareAlike 3.0 Unported license, which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited. <a href="https://creativecommons.org/licenses/by-nc-sa/3.0/">https://creativecommons.org/licenses/by-nc-sa/3.0/</a></td>
</tr>
<tr>
<td><strong>Item downloaded from</strong></td>
<td><a href="http://hdl.handle.net/10468/3664">http://hdl.handle.net/10468/3664</a></td>
</tr>
<tr>
<td><strong>Downloaded on 2018-12-27T14:50:09Z</strong></td>
<td></td>
</tr>
</tbody>
</table>
The Atomic Structure of the Phage Tuc2009 Baseplate Tripod Suggests that Host Recognition Involves Two Different Carbohydrate Binding Modules

Pierre Legrand, Barry Collins, Stéphanie Blangy, James Murphy, Silvia Spinelli, Carlos Gutierrez, Nicolas Richet, Christine Kellenberger, Aline Desmyter, Jennifer Mahony, Douwe van Sinderen, Christian Cambillau

ABSTRACT  The Gram-positive bacterium Lactococcus lactis, used for the production of cheeses and other fermented dairy products, falls victim frequently to fortuitous infection by tailed phages. The accompanying risk of dairy fermentation failures in industrial facilities has prompted in-depth investigations of these phages. Lactococcal phage Tuc2009 possesses extensive genomic homology to phage TP901-1. However, striking differences in the baseplate-encoding genes stimulated our interest in solving the structure of this host’s adhesion device. We report here the X-ray structures of phage Tuc2009 receptor binding protein (RBP) and of a “tripod” assembly of three baseplate components, BppU, BppA, and BppL (the RBP). These structures made it possible to generate a realistic atomic model of the complete Tuc2009 baseplate that consists of an 84-protein complex: 18 protein (RBP) and of a “tripod” assembly of three baseplate components, BppU, BppA, and BppL (the RBP). The RBP head domain possesses a different fold than those of phages p2, TP901-1, and 1358, while the so-called “stem” and “neck” domains share structural features with their equivalents in phage TP901-1. The BppA module interacts strongly with the BppU N-terminal domain. Unlike other characterized lactococcal phages, Tuc2009 baseplate harbors two different carbohydrate recognition sites: one in the bona fide RBP head domain and the other in BppA. These findings represent a major step forward in deciphering the molecular mechanism by which Tuc2009 recognizes its saccharidic receptor(s) on its host.

IMPORTANCE  Understanding how siphophages infect Lactococcus lactis is of commercial importance as they cause milk fermentation failures in the dairy industry. In addition, such knowledge is crucial in a general sense in order to understand how viruses recognize their host through protein-glycan interactions. We report here the lactococcal phage Tuc2009 receptor binding protein (RBP) structure as well as that of its baseplate. The RBP head domain has a different fold than those of phages p2, TP901-1, and 1358, while the so-called “stem” and “neck” share the fold characteristics also found in the equivalent baseplate proteins of phage TP901-1. The baseplate structure contains, in contrast to other characterized lactococcal phages, two different carbohydrate binding modules that may bind different motifs of the host’s surface polysaccharide.

Most viruses that infect bacteria—commonly termed bacteriophages (or phages)—belong to the Caudovirales order: they possess a tail for host recognition, adhesion, and genome delivery—three steps essential for host infection. Phages typically employ adhesion modules that are located on the capsid, tail tube, and/or tail tip to facilitate host scanning and reversible attachment to cell-wall-associated saccharides (1–6), as recently evidenced (7, 8). Following this reversible attachment, host-specific and irreversible adhesion is the exclusive responsibility of proteins located on the tail tip (9, 10). In Myoviridae, genome injection follows contraction of its long tail, thus leading to host cell piercing (11–13), while in Podoviridae (phages with short tails) and Siphoviridae (phages with long noncontractile tails), other mechanisms of nucleic acid delivery are presumed to occur (6, 8, 11).

Phages recognize and bind either a membrane protein receptor or a carbohydrate moiety located in or on the cell envelope of a specific host (6, 14, 15). Examples of bacteriophages that recognize proteinaceous receptors are the coliphages lambda and T5, which recognize LamB and FhuA, respectively, both located on the surface of the Escherichia coli cell envelope (16–19), and Bacillus phage SPP1 or lactococcal phage c2, which recognize the YueB or PIP membrane proteins emerging at the host’s cell surface (20, 21). The tail tip of phages that recognize a membrane protein receptor exhibits an elongated morphology as it extends out of the...
distal tail end as a fiber (6). In contrast, characterized myo- or siphophages that employ carbohydrate-dependent host recogni-
tion possess a large heteropolymERIC proteinaceous organelle, also
called the baseplate, with a multitude of individual binding sites to
ensure efficient and specific host adsorption (22–25). In the case
of myophage T4, host binding follows a two-step mechanism
whereby the baseplate-associated long tail fibers first reversibly
bind lipopolysaccharide (LPS) or OmpC (26–29), causing a con-
formational change in the baseplate that then allows irreversible
binding of the short tail fibers to LPS (30).

Siphophage baseplates have been thoroughly investigated in
the case of the Gram-positive bacterium Lactococcus lactis. L. lactis
is commonly employed as a starter culture in various dairy fer-
mentation processes, and its intensive industrial exploitation has
facilitated the isolation of a very large number of phages infecting
various strains of this bacterium (31, 32). Lactococcal phages have
been grouped into 10 different species based on DNA homology
and morphology (33). Among these, representatives of the P335
species) is maintained in a structural configuration that is Ca2+
dependent conformational change as a prerequisite for host infec-
tion (22, 23, 25, 43). The baseplate of P32 (936 species) undergoes a
dramatic, Ca2+-dependent conformational change as a prerequisite for host infec-
tion (10, 23, 44). In contrast, the baseplate of TP901-1 (P335 species)
is maintained in a structural configuration that is Ca2+-independent and “ready to use” (25, 45). In the case of TP901-1,
each of the six receptor-binding subunits, or so-called tripods that are
suspended from the hexameric Dtr protein of the baseplate, is
composed of three BppU and nine BppL protein monomers (25).

Lactococcal phage Tuc2009 (P335 species) is a close relative of
phage TP901-1, although they infect different hosts: L. lactis
UCS09.9 (or SMQ380) and 3107 (or SMQ415), respectively (39,
46). The Tuc2009 RBP and baseplate components have previously
been expressed as soluble modules, making it possible to obtain a
low-resolution “in solution” model (47), followed by the determin-
ation of the low-resolution structures of the virion’s baseplate and of the exposed tripod (composed of a BppUct/BppA/BppL
heteropolymer [see below]) by negative-staining electron micros-
copy (48). Tuc2009 structural proteins share 95 to 99% sequence
identity with those of TP901-1, with the exception of the RBP (also
called BppL). Furthermore, the C-terminus of Tuc2009-encoded
BppU contains a 25-residue extension (BppUct), and the base-
plate contains an accessory protein, BppA, whose precise role in
adsorption has remained elusive (48). Despite the availability of
soluble proteins and its similarity to equivalent TP901-1 proteins,
the Tuc2009 RBP and the “tripod” have stubbornly resisted ex-
haustive crystallization attempts.

In this report, we employed a combination of limited proteol-
ysis and the use of so-called “nanobodies” (49) to crystallize
Tuc2009 RBP and determined its structure, together with that of the
“tripod,” making it possible to determine the overall structure of the
Tuc2009 baseplate. The atomic structures demonstrate that the
RBP domain involved in host recognition (the so-called “head”) of Tuc2009, although a β-domain, exhibits a fold differ-
ent from that of its TP901-1 equivalent. Surprisingly, we discov-
ered that BppA possesses a carbohydrate binding domain with an
extended saccharidic binding site, suggesting that this protein par-
ticipates in receptor binding together with the bona fide RBP.

RESULTS
The nanobody L06 structure. Despite extensive and exhaustive
efforts, all of our attempts to crystallize Tuc2009 RBP (open reading
frame 53 [ORF53] or BppL) had been unsuccessful. We there-
fore decided to immunize a camel with this protein, hoping that
 generation of and subsequent association with specific nanobod-
ies would promote crystallization, as had previously been ob-
served (49), while it was also hoped that it would render effective
neutralizers of phage infection. Following an established proce-
dure (described in reference 49), we isolated nano-L06 that could
be expressed at reasonable yields (5 mg/liter). Nano-L06 crystal-
ized readily, and a data set was collected at 1.1-Å resolution at
Soleil Proxima 1 (Table 1). Nano-06 exhibits the classical
9-stranded β-sandwich fold of nanobodies with a long CDR3 fre-
quently observed in nanobodies from dromedaries, in contrast
with those originating from llamas (Fig. 1). The 17-residue-long
CDR3 forms a long loop followed by a 2.5 turn α-helix. A disulfide
bridge is observed between the CDR3 Cys 100h and the frame-
work Cys 45.

The nanobody L06/RBP complex structure. We cocrystallized the
RBP with nano-L06, although the RBP had unfortunately been
cleaved by trypsin digestion prior to complex purification. A data
set was collected at a 2.70-Å resolution at beamline Soleil Proxima
1 (Table 1), and the corresponding structure of the cocrystallized
complex was solved by molecular replacement using the obtained
nano-L06 structure (described above). The complex contains
a trimeric RBP domain and three nano-L06 moieties (Fig. 2A and
B). The N-terminal part of the RBP is not visible in the electron
density map, as the model could be built from residue Asp 47 to
residue Asn 173 (Fig. 2A to C).

The RBP structure starts with a 12-residue-long N-terminal
α-helix (residues 49 to 61) followed by a 12-strand β-barrel, con-
stituting the so-called “head” domain (Fig. 2C and D). At the
“bottom” of the RBP head monomer, a first antiparallel β-sheet
assembles β1–β12–β3–β4–β′–β″ and stacks against strands β3–β9
(Fig. 2D). At the “top” of the domain, strands β11–β4 face the
β7-β8-β10 β-sheet (Fig. 2C and D).

Each nanobody interacts with two RBP monomers (see Ta-
TABLE 1 Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Nano-L06 (S-SAD)</th>
<th>Nano-L06</th>
<th>RBP/nano-L06</th>
<th>Tripod</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDB no.</td>
<td>5E7B</td>
<td>5E7F</td>
<td>5E7T</td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>Soleil PX 1</td>
<td>Soleil PX 1</td>
<td>Soleil PX 1</td>
<td></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.7712</td>
<td>0.8856</td>
<td>0.9786</td>
<td>1.0087</td>
</tr>
<tr>
<td>Space group</td>
<td>P6_22</td>
<td>P6_22</td>
<td>P2_1_2_2</td>
<td>P2_3</td>
</tr>
<tr>
<td>Cell dimension (Å)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a = b = 52.1, c = 162.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angles (°)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a = β = θ = 120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution limits (Å)</td>
<td>45.1–1.9 (2.01–1.9)</td>
<td>45.1–1.1 (1.13–1.1)</td>
<td>47.1–2.7 (2.77–2.7)</td>
<td>34.9–2.9 (3.0–2.9)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.055 (0.086)</td>
<td>0.068 (1.55)</td>
<td>0.146 (1.60)</td>
<td>0.187 (1.91)</td>
</tr>
<tr>
<td>CC1/2</td>
<td>100 (99.7)</td>
<td>100 (61.7)</td>
<td>99.7 (74.8)</td>
<td>99.7 (50.6)</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>16,724 (561)</td>
<td>52,014 (3,596)</td>
<td>31,023 (2,264)</td>
<td>70,164 (6,732)</td>
</tr>
<tr>
<td>Mean I/σ(I)</td>
<td>61.92 (20.74)</td>
<td>18.4 (1.5)</td>
<td>11.0 (1.2)</td>
<td>10.7 (1.25)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>85.9 (39.1)</td>
<td>96.6 (91.7)</td>
<td>99.9 (99.9)</td>
<td>99.8 (99.3)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>37.2 (14.6)</td>
<td>19.5 (19.2)</td>
<td>7.2 (7.1)</td>
<td>13.7 (13.7)</td>
</tr>
<tr>
<td>SigAno</td>
<td>1.534 (0.841)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCnano/CCall/CCweak</td>
<td>37.6/22.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOM</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>15.0–1.10 (1.125–1.10)</td>
<td>34.5–2.7 (2.79–2.7)</td>
<td>34.9–2.9 (2.98–2.9)</td>
<td></td>
</tr>
<tr>
<td>No. of reflections</td>
<td>51,977 (2,905)</td>
<td>31,009 (2,812)</td>
<td>70,164 (5,103)</td>
<td></td>
</tr>
<tr>
<td>No. of protein/water/ligand atoms</td>
<td>1,968/188</td>
<td>5,848/40</td>
<td>8,126/61/47</td>
<td></td>
</tr>
<tr>
<td>No. of test set reflections</td>
<td>2,359</td>
<td>1,550</td>
<td>3,508</td>
<td></td>
</tr>
<tr>
<td>Rwork/Rfree (%)</td>
<td>15.8/18.9 (31.6/30.6)</td>
<td>19.4/20.8 (28.0/35.6)</td>
<td>21.4/23.7 (28.2/31.7)</td>
<td></td>
</tr>
<tr>
<td>RMSD bonds (Å)/angles (°)</td>
<td>0.0096/1.37</td>
<td>0.008/1.02</td>
<td>0.011/1.18</td>
<td></td>
</tr>
<tr>
<td>B Wilson/B avg</td>
<td>14.5/23.0</td>
<td>75.1/77.8</td>
<td>104.0/104.5</td>
<td></td>
</tr>
<tr>
<td>Ramachandran preferred/allowed/outliers (%)</td>
<td>98.6/1.4/0.0</td>
<td>98.2/0.2/0.0</td>
<td>95.1/8.0/0.1</td>
<td></td>
</tr>
</tbody>
</table>

"Numbers in parentheses refer to the highest-resolution bin.

It is noteworthy that the interactions reported by the PISA server (50) are quite different at the three interfaces, as the buried surface area (BSA) values of the nanobodies are 638, 609, and 840 Å² for monomers A, B, and C, respectively. Similarly, the BSA values for RBP monomers G, H, and I are 764, 636, and 787 Å² (see Table S1). The interaction involves the three CDRs but is dominated by CDR3 (see Table S1). On the RBP, most of the contacts of the first nanobody are established with the N-terminal helix (638 Å²), and the remaining interactions involve loops 68 to 71 (185 Å²), while the second nanobody interacts with loops 113 to 116 (147 Å²) and 139 to 142 (45 Å²), all of these loops being localized at the RBP “bottom.” A DALI server (51) search retrieved several significant hits: the best reported hit was a virulence plasmid protein pgp3-d (4JDO, 2.6 Å on 100 residues), and the second was a lectin, the N-terminal domain of bc2l-c lectin from Burkholderia cenocepacia (52) (2WQ4, 7.9 Å on 95 residues). It is noteworthy that these domains bear no fold similarity to the head domains of other structurally characterized lactococcal phage RBPs—i.e., those of phages p2, TP901-1, and 1358 (40, 42, 53).

The Tuc2009 tripod structure. Following a strategy similar to that employed for the baseplate structure of phage TP901-1 (25, 54), we expressed the so-called “tripod” of Tuc2009 (48). In TP901-1, the tripod is an assembly of a trimer of BppU (involving the 100 C-terminal residues of BppU) and three trimers of BppL (25). The situation is more complex with Tuc2009: we know from negatively stained electron microscopy (EM) images that such a tripod can be observed (48) and that it comprises the 100 C-terminal residues of BppU, followed by a C-terminal extension of ~25 residues specific to Tuc2009 and implicated in BppA association, the supplementary BppA protein, and BppL. This complex (in association with nano-L06) was crystallized, and a data set was collected at a 2.9-Å resolution at Soleil Proxima 1 (Table 1).

The tripod structure was solved by molecular replacement using similar, previously determined structures: the trimeric structure of ~25 residues specific to Tuc2009 and implicated in BppA association, the supplementary BppA protein, and BppL. This complex (in association with nano-L06) was crystallized, and a data set was collected at a 2.9-Å resolution at Soleil Proxima 1 (Table 1).
FIG 2 Structure of the RBP head domain in complex with nano-L06. (A) Ribbon view of the RBP head domain trimer (monomers colored blue, green, and pink) bound to three nano-L06 proteins (colored gray). (B) Ninety degree rotated view. Note the contact of nano-L06 with the helix of one monomer and the loops of a second one. (C) Ribbon view of the RBP head domain trimer (monomers colored blue, green, and pink) bound to three nano-L06 proteins (colored gray). (D) Ribbon view of the RBP head domain trimer rotated view. Note the contact of nano-L06 with the helix of one monomer and one of its adjacent domains (domain 1, residues 30 to 185) and a junction segment (residues 49 to 61) was found in an α-helical conformation in the cleaved RBP/head/nano-L06 complex (Fig. 2C and D). (E) Superposition between RBP monomers from the nano-L06/RBP complex (green) and the serendipitous domain found in the tripod structure (beige). The RMSD value is 2.4 Å.

RBP was available from a previous study, we used Chimera (57) to fit the complete tripod into the EM map (emd_2343.map at EMDB). The observed fit (ml = 0.83; 2,855/20,952 atoms outside contour) is good for the central part of the tripod, BppUct and BppL, but is not satisfactory for BppA (Fig. 3D). It is clear that BppA in the X-ray structure is rotated toward the tripod bottom (direction of RBP head), while it is situated higher in the EM map. However, the BppA volume of the X-ray structure is compatible with that of the EM map.

The trimeric RBP chain could be identified in the electron density map from Ala 2 to Asn 173 (Fig. 3A). The N-terminal segment of the RBP or so-called “stem” domain (residues 2 to 28) closely resembles that of TP901-1 as it starts with an elongated stretch (residues 2 to 10) followed by a turn and an α-helix (12–27) (Fig. 3A). An interlaced triple β-helix encompasses residues 29 to 59. This is followed by a loosely packed segment (residues 58 to 62) joining the triple β-helix to the head domain. However, the head domain, between residues 63 and 173, exhibits a very weak electron density map and consequently elicits very high B-factors. Noteworthy and surprisingly, a part of the triple β-helix and of the junction segment (residues 49 to 61) was found in an α-helical conformation in the cleaved RBP/head/nano-L06 complex (Fig. 2C and D).

The RBP three N-terminal stem stretches of residues 10 to 15 form a cup that harbors a loop of BppU (residues 217 to 235), ensuring a stable association between BppU and the RBP in a manner identical to that observed for the corresponding components of phage TP901-1 (25) (Fig. 3A). The BppU domain from Pro 195 (the N terminus of the construct) to Val 295 is quasi-identical to its corresponding domain in TP901-1. In Tuc2009, however, it is followed by a 25-residue extension (residues 296 to 321) forming a β-hairpin (Fig. 3A and F). This hairpin interacts strongly with the C-terminal domain of BppA since it completes one of its β-sheets.

BppA can be traced from Ala 3 to Lys 286; it begins with a loosely structured segment (residues 3 to 29) followed by a compact domain (domain 1, residues 30 to 185) and a junction segment (residues 186 to 229) interacting with the N-terminal segment, and it terminates in a β-sandwich fold (domain 2, residues 230 to 286) harboring the extension that links it to BppU as mentioned above (Fig. 3A and F). A DALI server-mediated search for domain similarity was performed for BppA domain 1, which retrieved similarity to carbohydrate binding domains 1GUI (Z = 10.1, RMSD = 2.6 Å on 120 residues, [58]) and 1GU3 (Z = 10.1, RMSD = 2.7 Å on 121 residues [58]), with a polysaccharide occupying an extended crevice (Fig. 3E). We performed a similar search for BppA domain 2, to which the β-hairpin extension of BppU (residues 296 to 321) was added. This composite domain was found to be very similar to a titin domain, 2NZI (Z = 5.9, RMSD = 1 Å on 56 residues [59]) (Fig. 3F).

The serendipitously cleaved BppL head domain (residues 52 to 173) is docked against BppA domain 1 (Fig. 3A). A monomer is observed in the crystallographic asymmetric unit, and the trimer is reconstituted by the crystallographic 3-fold axis. It exhibits an N-terminal α-helix followed by the β-domain of the RBP head as observed in the cleaved RBP/head/nano-L06 complex, but with a significant conformational change of some β-strand traces, with an RMSD value of 2.4 Å, while the RMSD value between the RBP head domains in the tripod and in the nano-L06 complex is 0.4 Å. In fact, the upper part of the RBP comprising the short upper
β-strands (as seen in Fig. 2C) has changed conformation, while the bottom part remained similar (Fig. 2E).

**The Tuc2009 baseplate model.** Knowing the atomic structure of Tuc2009 tripod facilitated the construction of a model of the complete baseplate of this phage based on the TP901-1 baseplate structure, since the Dit, Tal, and BppU C-terminal domains of the Tuc2009 baseplate are 99% identical to those of TP901-1. To build the Tuc2009 baseplate model, the BppU C-terminal domain of the TP901-1 baseplate was used to fit the Tuc2009 tripod onto the TP901-1 BppU/Dit complex. As demonstrated in our previous study, one of the BppA molecules (i.e., the most internal one) has to be removed from the tripod as it clashes with full-length BppU (48). Therefore, the Tuc2009 baseplate model assembles 93 proteins: those already found in TP901-1 baseplate, 6 Dit, 18 BppU, 54 BppL, and 3 Tal proteins (not included in the model), plus the 12 BppA proteins specific to Tuc2009 (Fig. 4A and B). This Tuc2009 baseplate model has a significantly larger diameter than that of TP901-1 (~320 Å versus ~270 Å) due to the BppA position at the baseplate periphery. We then superimposed the thus-obtained baseplate model on the low-resolution (34-Å) model based on a negatively stained EM map (emd-2340 at EMDB). We fitted it in the map using Chimera, which reported that 33,190 out of 154,323 (20%) atoms were situated outside the map (Fig. 4C). This rather mediocre result is due to the fact that the BppA molecules in the atomic structure are placed at a good radial position but are at a much lower position in the EM map. Most of the carbohydrate binding module (CBM) domain is therefore outside the map, thus explaining the poor fitting statistics. However, the remainder of the baseplate proteins (i.e., the Dit, BppU, and BppL subunits) all fit very well in the EM map (10% of atoms outside boundaries). The apparently lower position of BppA may be explained by the fact that BppA is joined to the baseplate core by a flexible linker and that the position of BppA may consequently be influenced by its surroundings, being different in the tripod compared to the virion’s baseplate, or by differences generated by the two techniques employed (EM versus X-ray). This flexibility is likely to have functional implications in allowing improved accessibility to cell wall polysaccharides.

**The putative saccharide binding site of RBP and BppA.** We have seen that the RBP head domain has a fold similar to that of a
trimeric lectin domain (2WQ4). The galactopyranoside Se derivative of the 2WQA structure binds in a crevice between two monomers in a typical saccharide binding site formed by an aromatic component (Tyr 48) and a strong network of hydrogen bonds involving Arg 85 and Arg 111, the carbonyl moiety of Thr 83, and the hydroxyl moiety of Thr 74. When positioned in Tuc2009 RBP, a typical saccharide binding is also observed involving Tyr 121, and hydrogen bond donors/acceptors Lys 162, His 163, Glu 101, and Asn 123 (Fig. 5) (data not shown). It is noteworthy that the saccharide binding site of phages TP901-1 and p2 is also located in a crevice between two monomers.

BppA was found to be similar to CBMs 1GUI and 1GU3. In the former structure, the bound saccharide exhibits a V-shape, while it is linear in the latter. When superimposing 1GUI and 1GU3 onto BppA-CBM (Fig. 5A), it is clear that the nature of the putative BppA-CBM saccharide binding site is not compatible with that of 1GU3. In contrast, it is very similar to the saccharide binding site of 1GUI—in both cases involving three tryptophan residues and several hydrogen bond donors/acceptors (Fig. 5B).

**Nanobody binding and neutralization studies.** We assayed the binding of nano-L06 to RBP and the tripod using biolayer interferometry (BLI). The nano-L06 was attached to the chip, and the RBP and tripod were dispensed over it at increasing concentrations (Fig. 6). A curve fitting using a 2:1 model (bivalent analyte) gave a $K_d$ (dissociation constant) value of $9.1 \pm 0.7 \text{nM}$ ($k_{on1} = 6.25 \times 10^4 \text{M}^{-1} \text{s}^{-1}$, $k_{off1} = 5.77 \times 10^{-4}$, and $k_{off2} = 4.3 \text{s}^{-1}$). The binding data of the tripod were analyzed using two methods. A curve fitting using a 2:1 model gave a $K_d$ value of $9.1 \pm 0.6 \text{nM}$ ($k_{on1} = 4.7 \times 10^4$ to $5.1 \times 10^4 \text{M}^{-1} \text{s}^{-1}$, $k_{off1} = 2.7 \times 10^{-4}$ to $2.8 \times 10^{-4}$, and

---

**FIG 4** Structure of the Tuc2009 baseplate. (A) Ribbon view of the baseplate tilted by 30°. (B) Ribbon top view of the baseplate. The Dit and BppU proteins are colored beige; the RBP's have the same color within a tripod. (C) Superposition of the EM map of the baseplate (yellow) and of the baseplate map calculated at a 25-Å resolution, indicating the different positions of BppA at the periphery (red arrows).

**FIG 5** The Tuc2009 tripod domains and their putative saccharide binding sites. (A) Right, fragment of the Tuc2009 ribbon structure showing BppA and RBP domains. Left top, the BppA domain 1 and the polysaccharide from the related structure of the CBD in the PDB entry 1GUI. Left below, the RBP head domain and the saccharide from the related structure of the lectin in PDB entry 2WQ4. (B) The polysaccharide/protein contacts in the 1GUI PDB entry structure (left) and in the BPPA domain 1 in which the saccharide from 1GUI has been modeled.
$k_{on}^2 = 1.94$ to 3.7 s$^{-1}$, with satisfactory statistics. The improvement of the fitting upon using a 2:1 model reveals that besides the low-affinity site, nonspecific weak binding may also occur.

In order to discern if nano-L06 interfered with the infectious capability of Tuc2009, infection neutralization assays were performed (see Materials and Methods). This nanobody was not observed to inhibit infection of the lactococcal host strain UC509.9. This result was not completely unexpected for nano-L06 since the crystal structure revealed that it binds at the base of the RBP head domain, far from the putative receptor binding site (Fig. 2A).

**DISCUSSION**

Recently, structural knowledge on the lactococcal phages’ RBPs and their host-encoded polysaccharidic cell wall receptors has provided some insights into lactococcal phage-host interactions at an atomic level (60). The complexity of lactococcal cell wall polysaccharides explains the exquisite specificity presented by these phages for their hosts. In this context, two previous studies pertaining to the Tuc2009 baseplate made it possible to determine its overall topology and low-resolution structure (47, 48). Here, the atomic structure of the Tuc2009 tripod reveals that this host-adhesion device harbors two different carbohydrate-binding modules (CBMs), the first being a classical bona fide RBP (i.e., BppL) resembling the N-terminal domain of bc21-c lectin from *Burkholderia cenocepacia* (52) (2WQ4), and the other, represented by the accessory protein BppA (which is absent in phage TP901-1), resembling the CBM domain of the *Thermotoga maritima* laminarinase 16a (58) (1GUI). The RBP head domain of Tuc2009 exhibits a striking flexibility, since the serendipitous head domain upper part conformation differs largely from those found in the bona fide tripod RBP or in the nano-L6/RBP complex. This flexibility may explain the difficulties encountered in the crystallization assays of Tuc2009 RBP. Noteworthy, the RBP head domain has a different fold from those of phages p2, TP901-1, and 1358, while the stem and neck share the structure of the equivalent baseplate proteins present in phage TP901-1 (Fig. 7). Nonetheless, the putative receptor binding site of Tuc2009 RBP is located in a crevice between two monomers, as is also the case for the RBPs of phages p2 and TP901-1. Analysis of the putative receptor binding sites indicates that they are very likely to be functional, in contrast to the Dit galectin domains of phages p2, TP901-1, and SPP1 in which the galectin carbohydrate-binding site is absent (23).

Indeed, phages are known to accommodate several carbohy-
mRNA from the acquired RNA using a reverse-PCR protocol. A
cloned into the Gateway destination vector pETG-20A for protein pro-
protein complex production in cloned into the nisin-inducible expression vector pTX8049 for BppUctAL
Camelus dromedarius
was then purified by Ni affinity and gel filtration chromatography follow-
was synthesized from the acquired RNA using a reverse-PCR protocol. A
lymphocytes were isolated from blood samples, and cDNA
and incomplete for the subsequent boosters) weekly for 6 weeks, and
combinant Tuc2009 tripod in phosphate-buffered saline (PBS) buffer was
sible to test the strength of its interaction with the CBM modules
questions will only be answered upon deciphering the host's pel-
Our findings raise intriguing questions about Tuc2009-host inter-
host's (L. lactis
phage Tuc2009 was
orf51
orf53
of phage Tuc2009 was
orf53
of phage Tuc2009 was
termed BppUct),
Protein production, purification, and crystallization.

MATERIALS AND METHODS

Protein production, purification, and crystallization. The orf53 gene of
phage Tuc2009 (encoding RBP or BppL [see the Results section]) was
cloned into the Gateway destination vector pETG-20A for protein pro-
duction in E. coli BL21, purified by Ni affinity and gel filtration chroma-
tography according to standard procedures (64, 65).

Tuc2009 tripod preparation was published previously (48). Briefly, a DNA
fragment encompassing the 3’ end of orf51 (encoding the C-terminal portion of BppU), encompassing residues 194 to 322, and
tered BppUct1, orf52 (encoding BppA), and orf53 of phage Tuc2009 was
cloned into the nisin-inducible expression vector pTX8049 for BppUctAL
protein complex production in L. lactis. The BppUctAL protein complex was
then purified by Ni affinity and gel filtration chromatography follow-
ing thioredoxin removal with tobacco etch virus (TEV) protease, accord-
ing to standard procedures.

With a view to obtaining nanobodies against Tuc2009 baseplate com-
ponents, a dromedary camel (Camelus dromedarius) from the
Canary Islands was immunized with Tuc2009 BppUctAL. About 300 µg of re-
combinant Tuc2009 tripod in phosphate-buffered saline (PBS) buffer was
jected in subtractively using Freund’s adjuvant (complete the first time
and incomplete for the subsequent boosters) weekly for 6 weeks, and
blood samples were collected aseptically in EDTA tubes 4 days after the
last booster. Lymphocytes were isolated from blood samples, and cDNA
was synthesized from the acquired RNA using a reverse-PCR protocol. A
nanobody phage display library of about 2 × 10^6 independent trans-
formants was generated using the phagemid vector pHEN4 (66). Phage dis-
play selection and screening of specific nanobodies were performed as
previously published (45). A clear enrichment of antigen-specific clones
was observed after three consecutive rounds of selection on solid-phase-
coated antigen. After the third round, nano-L06, specific for BppL,
was identified, after which the insert of the corresponding pHEN4-derived
plasmid was sequenced and the relevant sequence cloned into the PHEN6
evector (67). Nanobody (nano-L06) expression and purification were
performed as previously described (67).

The nano-L06 nanobody was subjected to crystallization screening
with a TTP Labtech Mosquito device in Greiner Bio-One CrystalQuick
plates. Crystals were obtained at 20°C by mixing 100 to 300 nl of 12
mg·ml^-1 protein (Na2HPO4, 10 mM; KH2PO4, 1.8 mM [pH 7.4];
NaCl, 137 mM; KCl, 2.7 mM) with 100 nl precipitant solution in 25
to 30% polyethylene glycol (PEG) 4000–0.2 M imidazole (pH 6.0). BppL
and nano-L06 were mixed at a ratio of 1:2 (mol/mol). The complex was
subjected to trypsin digestion (1,000:1 mol/mol) at room temperature for
2 h and immediately purified on Superdex 200 26/600 (GE Healthcare) in
PBS at pH 7.4. It was then concentrated to 9 mg·ml^{-1} with a 30-kDa
cutoff Amicon-Ultra for crystallization assays with a TTP Labtech Mos-
quito in Greiner Bio-One CrystalQuick plates. Crystals were obtained at
20°C by mixing 300 nl of protein (10 mM HEPES [pH 7.5], 150 mM NaCl
with 100 nl precipitant solution (25 to 30% PEG 4000, 0.2 M imidazole
[pH 6.0], or 0.1 M Tris [pH 8.0]).

The tripod complex of 280 kDa, as determined by size exclusion
chromatography-UV multimode light scattering (SEC-UV-MALS (48),
was concentrated to 10 mg/ml and subjected to crystallization screening
with a TTP Labtech Mosquito in Greiner Bio-One CrystalQuick plates.
Crystals were obtained at 20°C by mixing 300 nl of protein (Na2HPO4,
10 mM; KH2PO4, 1.8 mM [pH 7.4]; NaCl, 137 mM; KCl, 2.7 mM) with
100 nl precipitant solution (2 M ammonium sulfate, 0.1 M Na HEPES
[pH 7]).

Crystal structure determination. Crystals of the nano-L06 nanobody
were cryocooled without cryo-protectant, and data sets were collected at
Soleil Proxima 1 (Soleil synchrotron, Saint-Aubin, France) at wavelengths
of 1.7712 Å (1.90-Å resolution) and 0.86 Å (1.1-Å resolution). Data were
treated with XDS and XSCALE (68). The nano-L06 crystal was shown to
belong to space group P6_22 with cell dimensions a = b = 52.1 Å and c =
162.7 Å and to contain 44.6% solvent for a molecule per asymmetric unit.
The structure was determined by sulfur single-wavelength anomalous dis-
sersion (S-SAD) with the data set collected at the wavelength of 1.7712 Å.
Five initial sulfur sites (three individual methionines plus two cysteines
involved in a disulfide bond) were located with the SHELXC/D programs
(69). Using the program Phaser (70), the heavy-atom model was then
completed up to 7 sulfur sites, and phases were refined before being im-
proved by density modification using Parrot (71). The first steps of model
building were performed automatically using Buccaneer (72) and
completed manually with Coot (73). The model was finally refined using the
high-resolution data set collected at a 1.1-Å resolution with refmac
and Phenix (74, 75).

A crystal of the complex nano-L06/RBP was cryocooled in trimethyl-
amine N-oxide (TMAO) or 10% PEG 600, and a data set was collected at
Soleil Proxima 1 (Soleil synchrotron, Saint-Aubin, France) up to a
2.70-Å resolution. It was demonstrated to belong to space group P2_12_1 with cell
dimensions a = 84.7 Å, b = 88.04 Å, and c = 147.6 Å. Molecular
replacement was performed with Molrep using the refined structure of nano-L06.
It yielded three nano-L06 molecules related by a 3-fold axis and good
statistics and packing. Initial refinement was performed with Auto-
BUSTER (76) using dummy water molecules to complete the model, al-
though allowing further positioning of several β-strands. After density
modification using Parrot (71), the model was further automatically refined using
Buccaneer (72) and then completed manually with Coot (73). The result-
ing final structure, however, consisted only of the trimeric head domain

Legrand et al.
and an N-terminal β-helix, indicating that the stem had been removed during trypsin cleavage.

A crystal of the Tu2009 tripod BppUctAL complex was cryocooled in TMAO, and a data set was collected at Soleil Proxima 1 (Soleil synchrotron, Saint-Aubin, France) up to a 2.9-Å resolution. It was shown to belong to the cubic space group P2_13 with cell dimensions a = b = c = 2120 Å. Molecular replacement was performed with Phaser (70) using an ensemble1 comprising BppUct and the N-terminal part of the BppL head structure as determined above, yet deleted of its N-terminal helix. Phaser yielded a solution comprising a unique ensemble1, but forming a trimer reconstituted by the cubic 3-fold axis, and a trimer of ensemble2 representing the BppL head structure as determined above. Initial refinement was performed with AutoBUSTER (76) using dummy water molecules to complete the model, allowing further positioning of several β-strands between the BppL N terminus and head, where the expected BppL β-helix should be found, and at other positions that we presumed to be part of BppA. After density modification using Parrot (71), the model was further automatically refined using Buccaneer (72) and then manually completed with Coot (73), thereby yielding the BppL head domain and the complete BppA. At this stage, some extra density was still visible. After manual construction with Coot (73), we could identify the presence of a BppL head domain sitting on the cubic 3-fold axis that reconstituted a BppL head trimer.

Modeling of the phage Tu2009 baseplate. The BppU C-terminal domain of the TP901-1 baseplate, being highly similar to that of Tu2009, was used as a template to fit the Tu2009 tripod onto the TP901-1 BppU/Dit complex. Using Coot (73) option "SSM superpose," we used the Tu2009 BppUct domain of each tripod as the "source" structure and the BppUct domain of TP901-1 baseplate as the "target." Following this, the TP901-1 baseplate BppL and BppUct components were deleted. Furthermore, the most internal BppA modules of each Tu2009 tripod had to be removed from the tripod as they physically clash with full-length BppU (48).

BLL Prior to biosensor interferometry (BLI) assays, nano-L06 was biotinylated at a 1:1 ratio using the EZ-Link NHS-PEG4-biotin kit (Pierce Science, France). The reaction was stopped by removing the excess of biotin reagent using a Zeba Spin desalting column (Pierce Science, France). OctetRed96 (Fortebio, United States) was used for BLL studies. Assays were performed in black 96-well plates. The total working volume for samples or buffer was 0.2 ml, and the revolutions-per-minute setting was 1,000 rpm for baseline, loading, association, and dissociation steps. The experiments were performed at 25°C. Prior to each assay, streptavidin (SA) biosensor tips (Fortebio, USA) were hydrated in 0.2 ml kinetic buffer (KB; Fortebio, USA) for 20 min. The SA biosensor tips were then loaded with biotinylated nano-L06 at 5 μg/ml in KB, followed by a quenching step using biocytin. A baseline was recorded, and nano-L06 binding to RBP was monitored at concentrations of 0.22 to 140 nM. Association and dissociation were carried out for 400 s and 600 s, respectively. Complete dissociation of the complex was achieved by 3-fold regeneration (5 s in glycine 10 mM [pH 1.7]) and neutralization (5 s in KB). For the tripod, the same protocol was applied with a concentration range of 0.24 to 530 nM.

Neutralization studies. Neutralization assays were performed as described previously (45). Briefly, bromocresol purple (BCP) broth (3 ml) supplemented with 10 mM calcium chloride was the background medium in which the neutralization assays were performed. Nanobody nano-L06 was used for neutralization of infection by preincubation of 10^5 PFU Tu2009 and nanobody at 0.05, 0.1, 0.2, 5, 20, and 50 μg/ml for 1 h at 30°C before addition of 45 μl of a fresh overnight culture of the host, L. lactis UC509.9 (10^7 CFU), and further incubation for 7 h at 30°C. Controls were included to validate the assay, whereby the host alone (L. lactis UC509.9) or the host and its infecting phage (L. lactis UC509.9 plus Tu2009) were applied to show growth or lack of growth due to phage infection, respectively. Color change from purple to yellow indicated acidification and therefore growth of L. lactis UC509.9, whereas a purple color indicated lack of growth and thus phage infection.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl.do?doi=10.1128/mBio.01781-15/-DCSupplemental.

Table S1, DOCX file, 0.2 MB.

ACKNOWLEDGMENTS

We thank the Soleil synchrotron (France) for the beam time allocation (proposal 20140754).

This work was supported by grants from the Agence Nationale de la Recherche (grants ANR-11-BSV8-004-01 “Lactophages” and French Infrastructure for Integrated Structural Biology [FRISBI]). D. van Sinderen is supported by a Principal Investigator award (reference no. 13/IA/1953) through Science Foundation Ireland (SFI). J. Mahony is the recipient of a Technology Innovation Development Award (TIDA) (reference no. 14/TIDA/2287) funded by Science Foundation Ireland (SFI).

FUNDING INFORMATION

Science Foundation Ireland (SFI) provided funding to Douwe van Sinderen under grant number 13/IA/1953. Science Foundation Ireland (SFI) provided funding to Jennifer Mahony under grant number 14/TIDA/2287. Agence Nationale de la Recherche (ANR) provided funding to Christian Cambillau under grant number ANR-11-BSV8-004-01. French Infrastructure for Integrated Structural Biology (FRISBI) provided funding to C. Cambillau under grant number ANR-10-INSB-05-01.

REFERENCES


53. Spinelli S, Desmyter A, Verrips CT, de Haard HJ, Moineau S, Cambillau C. 2006. Lactococcal bacteriophage p2 receptor-binding protein structure suggests a common ancestor gene with bacterial and mamma-


