nature
structural &
molecular biology

Lactococcal bacteriophage p2 receptor-binding protein structure suggests a common ancestor gene with bacterial and mammalian viruses

Silvia Spinelli[1], Aline Desmyter[1], C Theo Verrips[2,3], Hans J W de Haard[2,4], Sylvain Moineau[5] & Christian Cambillau[1]

Lactococcus lactis **is a Gram-positive bacterium used extensively by the dairy industry for the manufacture of fermented milk products. The double-stranded DNA bacteriophage p2 infects specific *L. lactis* strains using a receptor-binding protein (RBP) located at the tip of its noncontractile tail. We have solved the crystal structure of phage p2 RBP, a homotrimeric protein composed of three domains: the shoulders, a β-sandwich attached to the phage; the neck, an interlaced β-prism; and the receptor-recognition head, a seven-stranded β-barrel. We used the complex of RBP with a neutralizing llama VHH domain to identify the receptor-binding site. Structural similarity between the recognition-head domain of phage p2 and those of adenoviruses and reoviruses, which invade mammalian cells, suggests that these viruses, despite evolutionary distant targets, lack of sequence similarity and the different chemical nature of their genomes (DNA versus RNA), might have a common ancestral gene.**

Lactococcus lactis is extensively used in starter cultures for the manufacture of fermented milk products such as cheeses, buttermilk and sour cream. Infection of these bacterial cells by phages impairs the fermentation process, and the dairy industry has a long history of dealing with contamination by phages, as these organisms are ubiquitous in their processing environments as well as in pasteurized milk[1].

Several hundred *L. lactis* phages have been characterized worldwide. They are all members of the order Caudovirales (phages with tails). A vast majority of these phages belong to the family Siphoviridae[1], characterized by a double-stranded DNA genome and a long non-contractile tail. Lactococcal phages are further divided into several groups[2], and those most frequently found in dairy factories belong to the 936, P335 and c2 species[3]. Phages from the 936 species, including the subject of this study, phage p2, and from the P335 species have an isometric capsid (morphotype B1; **Fig. 1a**), whereas those from the c2 species have a prolate capsid (morphotype B2).

It has been proposed that the first step of lactococcal phage infection involves interactions between the receptor-binding protein, RBP, and specific carbohydrate receptors evenly distributed on the surface of the cell wall[4–7]. For many phages, this binding step is reversible, and c2-like phages require a second, irreversible step of binding to the membrane-attached phage infection protein (PIP). However, phages from the 936 and P335 species do not use PIP as a secondary receptor[4,5]. A detailed understanding of the recognition between the phage and Gram-positive bacteria should help in designing tools to prevent costly phage infections.

Recently the use of a phage-neutralizing heavy-chain antibody fragment (VHH5) obtained from llama has been proposed as a useful strategy to prevent lactococcal phage infection. Camelid antibodies have a single immunoglobulin domain that functionally replaces the Fab fragment of classical antibodies[8]. VHHs contain only three complementarity-determining domains (CDRs) instead of six but often bind antigens with nanomolar affinities[8]. It has previously been shown that recombinant VHH5 produced from *Saccharomyces cerevisiae* prevents the infection of an *L. lactis* culture by phage p2 (ref. 9,10). VHH5 recognizes a 30-kDa protein (the phage ORF18 product) with a high affinity ($K_d = 1.4$ nM). This phage protein is located at the tip of the tail of phage p2 (**Fig. 1a**)[10]. Furthermore, it has recently been shown that the ORF18 of phage sk1 (936 species) is an RBP[5]. Because the amino acid sequence of ORF18 of phage p2 is 97% identical to that of ORF18 of phage sk1, ORF18 is the RBP of phage p2.

To better understand the initial molecular event underlying lactococcal phage infection, we have determined the crystal structure of phage p2 RBP at 2.3-Å resolution as well as the structure of the RBP in complex with a neutralizing llama VHH antibody domain[8]. This

[1]Architecture et Fonction des Macromolécules Biologiques, UMR 6098 CNRS and Universités d'Aix-Marseille I & II, Campus de Luminy, 163 Av. de Luminy 13288 Marseille Cedex 9, France. [2]Department of Biotechnology, Unilever Research Vlaardingen, Oliver van Noortlaan 120, 3133 AT Vlaardigen, The Netherlands. [3]Department of Molecular Cell Biology, EMSA, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands. [4]Ablynx, Technologiepark 4, 9052 Zwijnaarde, Belgium. [5]Groupe de Recherche en Écologie Buccale and the Félix d'Hérelle Reference Center for Bacterial Viruses, Faculté de Médecine Dentaire and Département de Biochimie et de Microbiologie, Faculté des Sciences et de Génie, Université Laval, Québec City, Québec, Canada G1K 7P4. Correspondence should be addressed to C.C. (christian.cambillau@afmb.univ-mrs.fr).

Received 6 July; accepted 26 October; published online 4 December 2005; doi:10.1038/nsmb1029

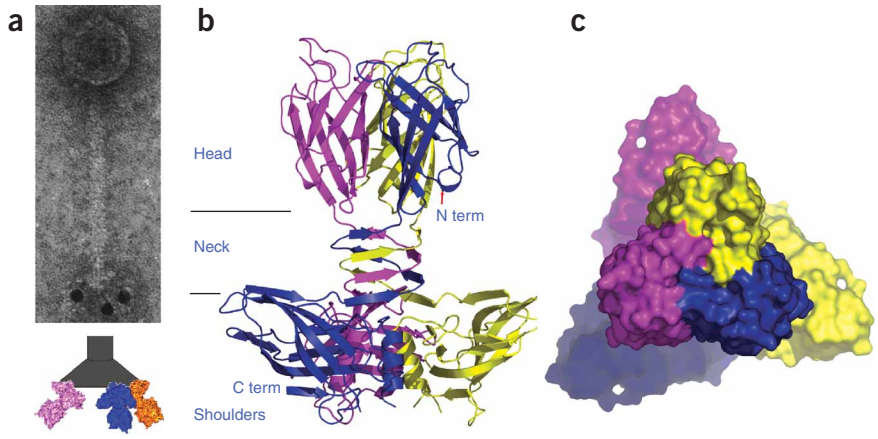

Figure 1 The receptor-binding protein from lactococcal phage p2. (**a**) Top, immuno–electron micrograph of phage p2 recognized by the gold-labeled neutralizing fragment VHH5 (ref. 10). Bottom, schematic view of the possible position of RBP on the phage baseplate. (**b**) The RBP trimer. The three-fold axis is vertical. The three subunits are colored pink, blue and yellow. (**c**) The RBP trimer molecular surface, rotated 90° from the view in **b**; the three-fold axis is perpendicular to the plane of the paper.

represents the first RBP structure from a phage infecting a Gram-positive bacterium. The RBP forms a homotrimer consisting of three domains: shoulders, neck and head. The neck and shoulder domains are structurally similar to those from other bacterial and mammalian viruses, suggesting that they may originate from a common source. The receptor-binding area is located on the head domain, as deduced from the structure of the RBP head domain in complex with VHH5.

RESULTS

Structure of the RBP

Phage p2 RBP consists of 264 amino acid residues and has sequence similarity only with other RBPs from closely related lactococcal phages. Three subunits form a homotrimer of overall dimensions 85 Å × 70 Å × 70 Å (**Fig. 1b**,**c**). A similar trimeric arrangement has also been observed in RBPs of mammalian adenoviruses and reoviruses[11–13] as well as for the protein gp12 of the *Escherichia coli* phage T4 (Myoviridae family, contractile tail)[14]. The three polypeptide chains in the complex have identical structures (r.m.s. deviations of 0.19–0.33 Å for all Cα atoms) and are related by a noncrystallographic three-fold axis parallel to the longest dimension of the trimer (vertical in **Fig. 1b**). As for the gp12 of phage T4 (ref. 14), the p2 RBP trimer is organized into three regions: shoulder, interlaced neck and head (**Fig. 1b**).

The shoulder domain, encompassing residues 1–141, has a β-sandwich fold (**Fig. 2a**). It also includes an α-helix between strands 1 and 2. A coil connects strands 2 and 3. A stretch of residues (positions 10–18) preceding the helix and facing the solvent is not visible in the electron density map. The long axis of each domain is approximately perpendicular to the three-fold axis of the complex (**Fig. 1b**). The helix and the second β-sheet (β-strands 5, 7 and 8) mediate tight interactions in the trimerization of the shoulder domain. A Dali search for similar structures to the shoulder domain resulted in no hits (best Z score = 1.2).

The neck is a triple-stranded β-helix that forms an equilateral triangular prism around the three-fold axis (**Fig. 1b**). Residues 142–163 from each subunit form four β-strands, and the strands interlace to form the β-prism. Thus, for each face of the prism, one subunit contributes two of the four strands, whereas the other two subunits each contribute one strand (**Fig. 2b**,**e**). This rigid neck structure has been observed in the gp12 (short tail fibers) of coliphage T4. In both cases, the neck connects a phage-linked N-terminal domain and a receptor-binding head[14,15]. The diameter of the gp12 β-prism is almost identical to that of the phage p2 neck. The phage p2 RBP neck is structurally different from the RBP tails of mammalian adenovirus and reovirus, which have a similar function[11,13].

The receptor head domain (residues 164–264), which immediately follows the neck (**Fig. 1b**), is a β-barrel made of seven anti-parallel β-strands (**Fig. 2c**,**d**,**e**). A Dali[16] search for this domain returned hits with low but significant similarity to RBPs of other viruses: the reovirus attachment protein s1 trimer[11] (PDB entry 1KKE; $Z = 4.3$; r.m.s. deviation = 3.2 Å for 91 residues) and the head domain of the adenovirus fiber fragment[13] (PDB entry 1QHV; $Z = 0.9$; r.m.s. deviation = 3.1 Å for 84 residues). The head domains of phage p2, reovirus and adenovirus[11,13] share a double Greek-key motif in their topology but lack any sequence similarity. The fourth β-strand of the first Greek-key motif is replaced in the phage p2 RBP head by an extended, irregular conformation (**Fig. 2d**). The phage p2 head shares no structural similarity with the RBP heads of two other *E. coli* viruses: phage T4 (ref. 17) and the tailless phage PRD1 (ref. 18). In the former case, the RBP head has a fold assembling three intertwined β-stranded subunits[17], whereas in the latter case, it consists of a five-bladed irregular propeller domain[18]. Other RBP heads from mammalian viruses fold and assemble in different ways from those of reo- and adenoviruses[19–21]. As an example, the rhesus rotavirus VP8 domain has a galectin fold that has lost its original carbohydrate-binding site, so that it binds sialic acid at a different position from that of the classical galectins[22].

The head domains in the trimer are parallel to the three-fold axis and form a compact structure in which each head domain contacts the other two in the trimer (**Fig. 1b**,**c**). The interaction area among the subunits involves mainly β-strands 3, 7, 4 and 5. Each subunit head has a water-accessible surface area of about 6,000 Å^2, 1,860 Å^2 of which is buried, representing 31% of its whole surface. This value is far above average values generally observed, which indicates strong interactions among subunits (see below). The phage p2 RBP head trimer is tightly assembled, with very little space between the subunits (**Fig. 1c**). The same dense packing is observed in the head-domain trimer of gp12 of phage T4, in which 1,550 Å^2 of each monomer's surface area is buried from a total surface of 4,490 Å^2 (34%). In contrast, although p2, adenovirus and reovirus have similar head-domain assemblies, the latter two are more loosely packed: in reovirus, the interdomain-interacting surface is half that observed in lactococcal bacteriophage p2.

Complex of RBP with the neutralizing llama VHH5

The llama anti-RBP VHH5 binds the p2 RBP at the distal part of the phage tail and can suppress phage infection[10], suggesting that VHH5 may function by binding at or near the receptor-binding site of p2. To identify the receptor-binding area of p2, we cocrystallized RBP with VHH5. During crystallization, proteolytic cleavage occurred between the neck and the head domains, yielding the phage p2 RBP trimeric head in complex with the neutralizing VHH5. VHH5 binding does not affect the structure of the trimeric head assembly (0.3 Å r.m.s.

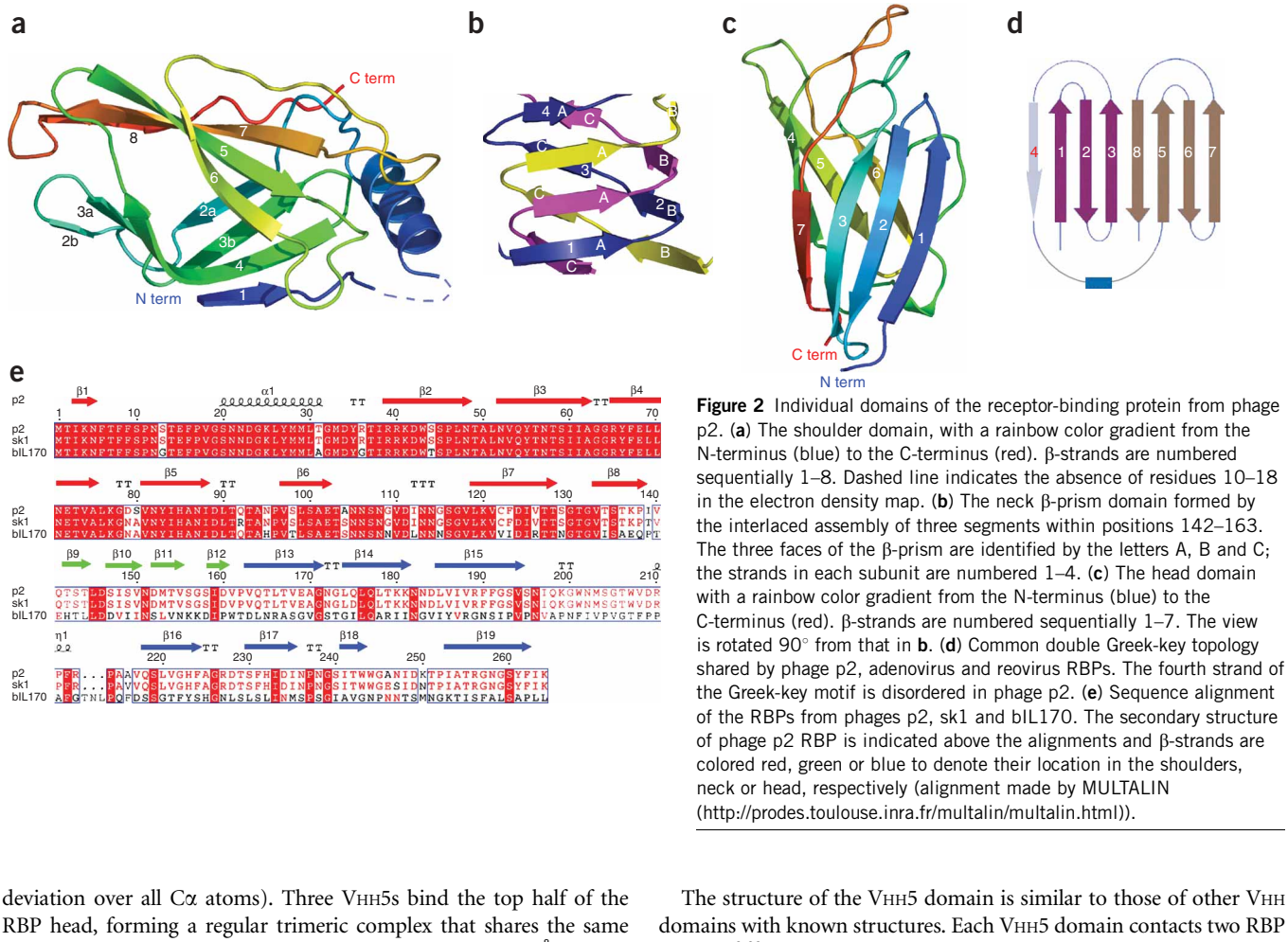

Figure 2 Individual domains of the receptor-binding protein from phage p2. (**a**) The shoulder domain, with a rainbow color gradient from the N-terminus (blue) to the C-terminus (red). β-strands are numbered sequentially 1–8. Dashed line indicates the absence of residues 10–18 in the electron density map. (**b**) The neck β-prism domain formed by the interlaced assembly of three segments within positions 142–163. The three faces of the β-prism are identified by the letters A, B and C; the strands in each subunit are numbered 1–4. (**c**) The head domain with a rainbow color gradient from the N-terminus (blue) to the C-terminus (red). β-strands are numbered sequentially 1–7. The view is rotated 90° from that in **b**. (**d**) Common double Greek-key topology shared by phage p2, adenovirus and reovirus RBPs. The fourth strand of the Greek-key motif is disordered in phage p2. (**e**) Sequence alignment of the RBPs from phages p2, sk1 and bIL170. The secondary structure of phage p2 RBP is indicated above the alignments and β-strands are colored red, green or blue to denote their location in the shoulders, neck or head, respectively (alignment made by MULTALIN (http://prodes.toulouse.inra.fr/multalin/multalin.html)).

deviation over all Cα atoms). Three VHH5s bind the top half of the RBP head, forming a regular trimeric complex that shares the same three-fold axis as unbound RBP, with an overall width of 90 Å (**Fig. 3a** and **Supplementary Fig. 1** online). Each VHH5 contacts two RBP subunits, but the footprint on one subunit is larger than that on the other (**Fig. 3a**,**b**). The surface area covered in the interaction is 720 Å^2 on the VHH domain and 835 Å^2 on the RBP head trimer, in the range of values observed with classical or camel antibodies. The interaction surface includes seven charged or polar noncharged residues and five nonpolar residues (**Supplementary Table 1** online).

The structure of the VHH5 domain is similar to those of other VHH domains with known structures. Each VHH5 domain contacts two RBP subunits[8,23], predominantly with its CDR2 and CDR3 (**Fig. 3c** and **Supplementary Table 1**). The surface of the RBP head interacting with VHH5 contains seven charged or polar residues: Asn202, His232, Asp234, Asn236, Ser240 and Thr242 on one monomer and Arg256 on the other monomer. Nonpolar residues are also observed: Trp201, Trp207, Trp242 and Pro237 on the first monomer and Ala225 on the second monomer. The interface is quite polar, with a total of 15 hydrogen bonds and an ionic bond. However, the interface also

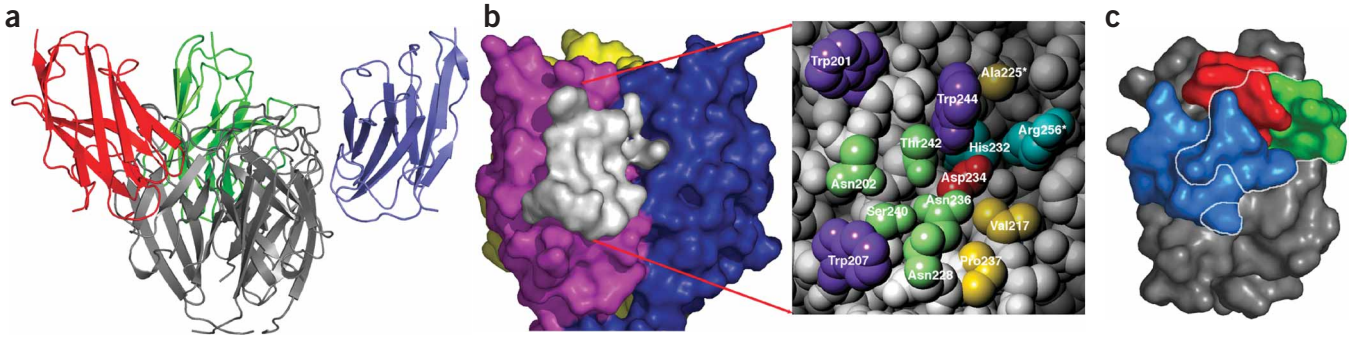

Figure 3 The complex of the phage p2 RBP head with VHH5. (**a**) The RBP–VHH5 complex. The three head domains, the only parts of RBP observed in the structure of the complex, are in gray. The three VHH5s are colored red, green and blue. (**b**) Views of the surface of the RBP-head trimer that interacts with VHH5 (white). Inset, spheres representation of the RBP-head residues interacting with VHH5. Residues labeled with and without asterisks are from different monomers. (**c**) VHH5 CDR1 (red), CDR2 (green) and CDR3 (blue). The surface interacting with the RBP is circled by a white line encompassing part of the CDRs.

contains three tryptophan residues from RBP and two tyrosines and a phenylalanine from VHH5 making nonpolar interactions. The amino acid composition of the head's VHH5-binding site is reminiscent of those of sugar-binding sites, such as those observed in lectins[24] or rotavirus VP8[22]. These often associate patches of aromatic residues that accommodate the nonpolar faces of saccharides and polar residues that function as specific hydrogen bond donors or acceptors[22,25]. However, the RBP fold is dissimilar to those of all lectin or galectin folds currently known.

DISCUSSION

To date, the complete genome sequence is available for only two lactococcal phages of the 936 group, namely sk1 (ref. 26) and bIL170 (ref. 27). Both phages share more than 80% DNA genome identity, yet they infect different *L. lactis* strains. It has been shown recently that swapping the RBPs of phages bIL170 and sk1 is sufficient to modify the host range of chimeric phages[5]. As indicated previously, the RBP of phage sk1 shares 97% identity with the RBP of p2 (**Fig. 2e**) and they infect the same *L. lactis* strains. Notably, whereas the RBP shoulder domains of phages p2 and bIL170 share ~90% amino acid identity, this figure drops below 15% for the necks and heads (**Fig. 2e**). As each virus evolves along with its host cells to ensure efficient entry and replication, the drastic sequence difference observed between the RBP heads of phages p2 and bIL170 points to the possibility that domain shuffling accounts for the differences observed in receptor specificity and hence in host recognition[5]. The high sequence similarity between the receptor-binding regions of phage T4 (gp37) and the λ gpStf tail-fiber proteins is another recognized example of phage gene swapping[28]. In fact, the tailed double-stranded DNA phages seem to be exchanging genetic information throughout the biosphere[29,30], which may explain the assembly of phage p2 RBP: preexisting viral building blocks or parts, such as the shoulder, the neck or the head, might originate from different precursors, such as reoviruses or adenoviruses for the head and phage T4 for the neck.

It is generally believed that virus origins are multiple (or polyphyletic), in contrast to the unique origins of the three kingdoms of cellular life (monophyletic). The huge number of viruses in the biosphere (generally assumed to outnumber cells by a factor of ~10) may be organized into a small number of viral lineages that each include members infecting hosts belonging to different domains of life[31]. It has also been suggested that a common architecture might be the hallmark of a viral lineage spanning the three domains of life[11,31–34]. It has been shown previously that the major coat protein of coliphage PRD1 resembles that of human adenovirus, raising the possibility of a common origin[35]. The similar topology reported here among the viral RBP heads of phage p2, adenoviruses and reoviruses provides support for a common origin of these genes. The nature of this viral evolutionary process may result from lateral gene transfer in the same ecological niche (for instance, in the milk of mammals) or from distinct virus lineages appearing at different times during evolution that have independently chosen these motifs from a common original pool.

METHODS

Structure of phage p2 RBP. Phage p2 RBP was cloned, expressed and purified as described previously[9,10]. Crystallization of the selenomethionine (SeMet) RBP was performed at 20 °C using hanging drop vapor diffusion. A volume of 1 μl of protein (6–9 mg ml^{-1}) was mixed with 1 μl of reservoir solution containing 0.85–1.1 M ammonium sulfate in 0.1 M MES (pH 6.5). Crystals were improved by microseeding. Crystals contained one RBP trimer per asymmetric unit (V_m = 3.05 $Å^3$ Da^{-1}). After cryo-cooling with 30% glycerol (v/v), a SAD experiment was performed on beamline ID14-EH4 (European Synchrotron Radiation Facility) at the Se K-edge (λ = 0.9794 Å). SAD data were collected from a single crystal using an ADSC Quantum 4 detector. Data were indexed and integrated using MOSFLM[36,37] and scaled using SCALA[36], and structure-factor amplitudes were calculated using TRUNCATE[36] (**Table 1**). SOLVE[38] readily identified the 15 Se picks arising from the 15 methionines in the asymmetric unit (excluding N-terminal ones). Initial phases were calculated with SOLVE[38] and were improved by solvent flattening using RESOLVE[38]. The structure of an almost complete subunit was built and used to find the positions of the two other subunits in the asymmetric unit with AMoRe[39]. Cycles of ARP/wARP[40] and unrestrained TLS refinement with REFMAC5 (ref. 41) alternating with rebuilding using Turbo-Frodo[42] led to the final trimer structure (**Table 1** and **Supplementary Fig. 2** online). The final RBP model counts 88.2% and 11.8% residues in allowed and generously allowed regions of the Ramachandran chart[43].

Structure of the complex of phage p2 RBP with VHH5. Crystallization assays were performed with the RBP trimer and the llama antibody VHH5. A few crystals appeared after 1 year. Molecular replacement with AMoRe[39] failed when the complete RBP trimer was used as starting model, but gave a clear solution when the trimeric head domain was used (c = 0.53, R = 0.46). The position of a VHH5 molecule was then identified and those of the other two

Table 1 Data collection, phasing and refinement statistics

	Native RBP	SeMet RBP	RBP–VHH5 complex
Data collection			
Space group	$P2_12_12_1$		$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	76.11, 96.25, 149.49		74.76, 83.94, 137.8
		Peak (SAD)	
Wavelength (Å)	0.9794	0.9794	0.9794
Resolution (Å)	40.0–2.3 (2.3–2.36)	40.0–2.5 (2.5–2.6)	20.0–2.7 (2.7–2.77)
R_{sym} or R_{merge}	0.09 (0.33)	0.075 (0.28)	0.097 (0.30)
I / σI	6.7 (2.0)	8 (2.6)	5.6 (2.0)
Completeness (%)	100 (99)	100 (99)	99.0 (98)
Redundancy	3.5 (2.6)	4.2 (2.9)	2.3 (2.1)
Refinement			
Resolution (Å)	20.0–2.3 (2.3–2.36)		15.0–2.7 (2.7–2.77)
No. reflections	48,171		23,055
R_{work} / R_{free}	0.22 / 0.25 (0.27 / 0.29)		0.25 / 0.29 (0.32 / 0.39)
No. atoms			
Protein	5,820		5,317
Water	236		—
B-factors			
Protein	23.0		23.72
Water	40		—
R.m.s. deviations			
Bond lengths (Å)	0.009		0.006
Bond angles (°)	1.14		0.883

One crystal was used for native and SeMet data sets and RBP–VHH5 complex. Highest-resolution shell is shown in parentheses.

deduced by applying three-fold symmetry. The final model contains a head-domain trimer (residues 158–264) and three VHH5 molecules ($V_m = 2.95$ Å^3 Da^{-1}). Analysis of crystal packing reveals that no space is left for the neck or shoulders, indicating that cleavage might have occurred during the long crystallization time, a fact confirmed by mass spectrometry. The same refinement procedure was followed as for native RBP (**Table 1**). Figures were drawn with PyMol (http://pymol.sourceforge.net/).

Accession codes. Protein Data Bank: Coordinates have been deposited with accession codes 2BSD (phage p2 RBP) and 2BSE (RBP–VHH5 complex).

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

This work was supported by the Marseille-Nice Genopole, by the European Union's Structural Proteomics in Europe program (fifth PCRDT, QLG2-CT-2002-00988) and by a grant from the Natural Sciences and Engineering Research Council of Canada. C. Huyghe is greatly acknowledged for protein production and D. Tremblay (Laval University, Quebec, Canada) for phage p2 DNA.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Moineau, S., Tremblay, D. & Labrie, S. Phages of lactic acid bacteria: from genomics to industrial applications. *ASM News* **68**, 388–393 (2002).
2. Labrie, S. & Moineau, S. Complete genomic sequence of bacteriophage ul36: demonstration of phage heterogeneity within the P335 quasi-species of lactococcal phages. *Virology* **296**, 308–320 (2002).
3. Jarvis, A.W. *et al.* Species and type phages of lactococcal bacteriophages. *Intervirology* **32**, 2–9 (1991).
4. Valyasevi, R., Sandine, W.E. & Geller, B.L. The bacteriophage kh receptor of *Lactococcus lactis* subsp. *cremoris* KH is the rhamnose of the extracellular wall polysaccharide. *Appl. Environ. Microbiol.* **56**, 1882–1889 (1990).
5. Dupont, K., Vogensen, F.K., Neve, H., Bresciani, J. & Josephsen, J. Identification of the receptor-binding protein in 936-species lactococcal bacteriophages. *Appl. Environ. Microbiol.* **70**, 5818–5824 (2004).
6. Duplessis, M. & Moineau, S. Identification of a genetic determinant responsible for host specificity in *Streptococcus thermophilus* bacteriophages. *Mol. Microbiol.* **41**, 325–336 (2001).
7. Schafer, A., Geis, A., Neve, H. & Teuber, M. Bacteriophage receptors of *Lactococcus lactis* subsp. *diacetylactis* F7/2 and *Lactococcus lactis* subsp. *cremoris* Wg2–1. *FEMS Microbiol. Lett.* **62**, 69–73 (1991).
8. Muyldermans, S., Cambillau, C. & Wyns, L. Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains. *Trends Biochem. Sci.* **26**, 230–235 (2001).
9. Ledeboer, A.M. *et al.* Preventing phage lysis of *Lactococcus lactis* in cheese production using a neutralizing heavy-chain antibody fragment from llama. *J. Dairy Sci.* **85**, 1376–1382 (2002).
10. De Haard, H.J. *et al.* Llama antibodies against a lactococcal protein located at the tip of the phage tail prevent phage infection. *J. Bacteriol.* **187**, 4531–4541 (2005).
11. Chappell, J.D., Prota, A.E., Dermody, T.S. & Stehle, T. Crystal structure of reovirus attachment protein sigma1 reveals evolutionary relationship to adenovirus fiber. *EMBO J.* **21**, 1–11 (2002).
12. Burmeister, W.P., Guilligay, D., Cusack, S., Wadell, G. & Arnberg, N. Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. *J. Virol.* **78**, 7727–7736 (2004).
13. van Raaij, M.J., Mitraki, A., Lavigne, G. & Cusack, S. A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. *Nature* **401**, 935–938 (1999).
14. van Raaij, M.J., Schoehn, G., Burda, M.R. & Miller, S. Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. *J. Mol. Biol.* **314**, 1137–1146 (2001).
15. Kanamaru, S. *et al.* Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**, 553–557 (2002).
16. Holm, L. & Sander, C. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478–480 (1995).
17. Thomassen, E. *et al.* The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. *J. Mol. Biol.* **331**, 361–373 (2003).
18. Xu, L., Benson, S.D., Butcher, S.J., Bamford, D.H. & Burnett, R.M. The receptor binding protein P2 of PRD1, a virus targeting antibiotic-resistant bacteria, has a novel fold suggesting multiple functions. *Structure (Camb).* **11**, 309–322 (2003).
19. Fass, D. *et al.* Structure of a murine leukemia virus receptor-binding glycoprotein at 2.0 angstrom resolution. *Science* **277**, 1662–1666 (1997).
20. Gibbons, D.L. *et al.* Conformational change and protein-protein interactions of the fusion protein of Semliki Forest virus. *Nature* **427**, 320–325 (2004).
21. Modis, Y., Ogata, S., Clements, D. & Harrison, S.C. Structure of the dengue virus envelope protein after membrane fusion. *Nature* **427**, 313–319 (2004).
22. Dormitzer, P.R., Sun, Z.Y., Wagner, G. & Harrison, S.C. The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site. *EMBO J.* **21**, 885–897 (2002).
23. Desmyter, A. *et al.* Three camelid VHH domains in complex with porcine pancreatic alpha-amylase. Inhibition and versatility of binding topology. *J. Biol. Chem.* **277**, 23645–23650 (2002).
24. Bourne, Y. *et al.* Structures of a legume lectin complexed with the human lactotransferrin N2 fragment, and with an isolated biantennary glycopeptide: role of the fucose moiety. *Structure* **2**, 209–219 (1994).
25. Bourne, Y. *et al.* Three-dimensional structures of complexes of *Lathyrus ochrus* isolectin I with glucose and mannose: fine specificity of the monosaccharide-binding site. *Proteins* **8**, 365–376 (1990).
26. Chandry, P.S., Moore, S.C., Boyce, J.D., Davidson, B.E. & Hillier, A.J. Analysis of the DNA sequence, gene expression, origin of replication and modular structure of the *Lactococcus lactis* lytic bacteriophage sk1. *Mol. Microbiol.* **26**, 49–64 (1997).
27. Crutz-Le Coq, A.M., Cesselin, B., Commissaire, J. & Anba, J. Sequence analysis of the lactococcal bacteriophage bIL170: insights into structural proteins and HNH endonucleases in dairy phages. *Microbiol.* **148**, 985–1001 (2002).
28. George, D.G., Yeh, L.S. & Barker, W.C. Unexpected relationships between bacteriophage lambda hypothetical proteins and bacteriophage T4 tail-fiber proteins. *Biochem. Biophys. Res. Commun.* **115**, 1061–1068 (1983).
29. Tetart, F. *et al.* Phylogeny of the major head and tail genes of the wide-ranging T4-type bacteriophages. *J. Bacteriol.* **183**, 358–366 (2001).
30. Hendrix, R.W. Bacteriophages: evolution of the majority. *Theor. Popul. Biol.* **61**, 471–480 (2002).
31. Bamford, D.H. Do viruses form lineages across different domains of life? *Res. Microbiol.* **154**, 231–236 (2003).
32. Benson, S.D., Bamford, J.K., Bamford, D.H. & Burnett, R.M. Does common architecture reveal a viral lineage spanning all three domains of life? *Mol. Cell* **16**, 673–685 (2004).
33. Rice, G. *et al.* The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *Proc. Natl. Acad. Sci. USA* **101**, 7716–7720 (2004).
34. Gibbons, D.L. *et al.* Visualization of the target-membrane-inserted fusion protein of Semliki Forest virus by combined electron microscopy and crystallography. *Cell* **114**, 573–583 (2003).
35. Benson, S.D., Bamford, J.K., Bamford, D.H. & Burnett, R.M. Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. *Cell* **98**, 825–833 (1999).
36. Collaborative Computational Project, Number 4. The CCP4 suite: programs for crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–766 (1994).
37. Leslie, A.G. Integration of macromolecular diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 1696–1702 (1999).
38. Terwilliger, T. SOLVE and RESOLVE: automated structure solution, density modification and model building. *J. Synchrotron Radiat.* **11**, 49–52 (2004).
39. Navaza, J. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
40. Morris, R.J., Perrakis, A. & Lamzin, V.S. ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol.* **374**, 229–244 (2003).
41. Murshudov, G., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
42. Roussel, A. & Cambillau, C. The TURBO-FRODO graphics package. in *Silicon Graphics Geometry Partners Directory* 86 (Silicon Graphics, Mountain View, USA, 1991).
43. Laskowski, R., MacArthur, M., Moss, D. & Thornton, J. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 91–97 (1993).