Journal of Structural Biology 172 (2010) 75–84

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Solution and electron microscopy characterization of lactococcal phage baseplates expressed in *Escherichia coli*

Valérie Campanacci [a,*], David Veesler [a], Julie Lichière [a], Stéphanie Blangy [a], Giuliano Sciara [a], Sylvain Moineau [b,c], Douwe van Sinderen [d,e], Patrick Bron [f], Christian Cambillau [a,*]

[a] *Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR 6098 CNRS and Universités Aix-Marseille I & II, Campus de Luminy, Case 932, 13288 Marseille Cedex 09, France*
[b] *Groupe de Recherche en Écologie Buccale (GREB) and Félix d'Hérelle Reference Center for Bacterial Viruses, Faculté de Médecine Dentaire, Université Laval, Québec City, Québec, Canada G1V 0A6*
[c] *Département de Biochimie et de Microbiologie, Faculté des Sciences et de Génie, Université Laval, Québec City, Québec, Canada G1V 0A6*
[d] *Department of Microbiology, National University of Ireland, Cork, Ireland*
[e] *Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland*
[f] *Centre de Biochimie Structurale, INSERM U554/CNRS UMR 5048, 29 rue de Navacelles, 34090 Montpellier, France*

ARTICLE INFO

Article history:
Received 29 October 2009
Received in revised form 4 February 2010
Accepted 7 February 2010
Available online 11 February 2010

Keywords:
Operon expression
Lactococcus lactis phage
Multi-protein complex
Multi-angle light scattering
Receptor binding protein
Baseplate
Electron microscopy

ABSTRACT

We report here the characterization of several large structural protein complexes forming the baseplates (or part of them) of *Siphoviridae* phages infecting *Lactococcus lactis*: TP901-1, Tuc2009 and p2. We revisited a "block cloning" expression strategy and extended this approach to genomic fragments encoding proteins whose interacting partners have not yet been clearly identified. Biophysical characterization of some of these complexes using circular dichroism and size exclusion chromatography, coupled with on-line light scattering and refractometry, demonstrated that the over-produced recombinant proteins interact with each other to form large (up to 1.9 MDa) and stable baseplate assemblies. Some of these complexes were characterized by electron microscopy confirming their structural homogeneity as well as providing a picture of their overall molecular shapes and symmetry. Finally, using these results, we were able to highlight similarities and differences with the well characterized much larger baseplate of the myophage T4.

1. Introduction

During the course of several structural genomics programs, we designed several methods to easily express hundreds of single proteins in *Escherichia coli* and screen for the best conditions in order to obtain them in a soluble form and prone to crystallization (Alzari et al., 2006; Campanacci et al., 2003; Graslund et al., 2008; Vincentelli et al., 2003, 2004, 2005). However, the over-expression of individual subunits of a multi-protein complex in *E. coli* can result in inclusion body formation, improper folding, degraded products or cell growth inhibition due to product toxicity. In contrast, co-expression of two or more subunits of such complexes can enhance expression yield, solubility, correct folding and result in higher activities (Buddha et al., 2004; Dokland et al., 2002; Li et al., 1997; Stebbins et al., 1999; Strong et al., 2006). Different strategies can be applied for co-expression in *E. coli* such as the use of several vectors, a single vector with a single promoter, a single vector with different promoters, or a combination of all of these approaches (for review, see Perrakis and Romier (2008)). Recently, co-expression of four genes coding for the Type IV secretion system made it possible to crystallize it and solve its 3D structure (Chandran et al., 2009).

In this contribution, we describe the expression and characterization of the overall assembly of different baseplates or baseplate components belonging to three lactococcal phages (TP901-1, Tuc2009 and p2) infecting different strains of the low GC content gram-positive lactic acid bacterium *Lactococcus lactis* (Fig. 1). Gene expression of lactococcal phages is tightly and timely controlled during the replication cycle. In several cases, it was shown that such phage genes are organized in at least three sequential clusters

Abbreviations: RBS, ribosome binding site; TMP, tape measure protein; Dit, distal tail protein; Tal, tail-associated lysin; BppU, baseplate upper protein; BppL, baseplate lower protein; RBP, receptor binding protein; SEC/MALS/RI, size exclusion chromatography coupled with on-line static light scattering and refractometry; CD, circular dichroism; EM, electron microscopy; TEV, tobacco etch virus.

* Corresponding authors. Fax: +33 491 266 720.
E-mail addresses: valerie.campanacci@afmb.univ-mrs.fr (V. Campanacci), cambillau@afmb.univ-mrs.fr (C. Cambillau).

1047-8477/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2010.02.007

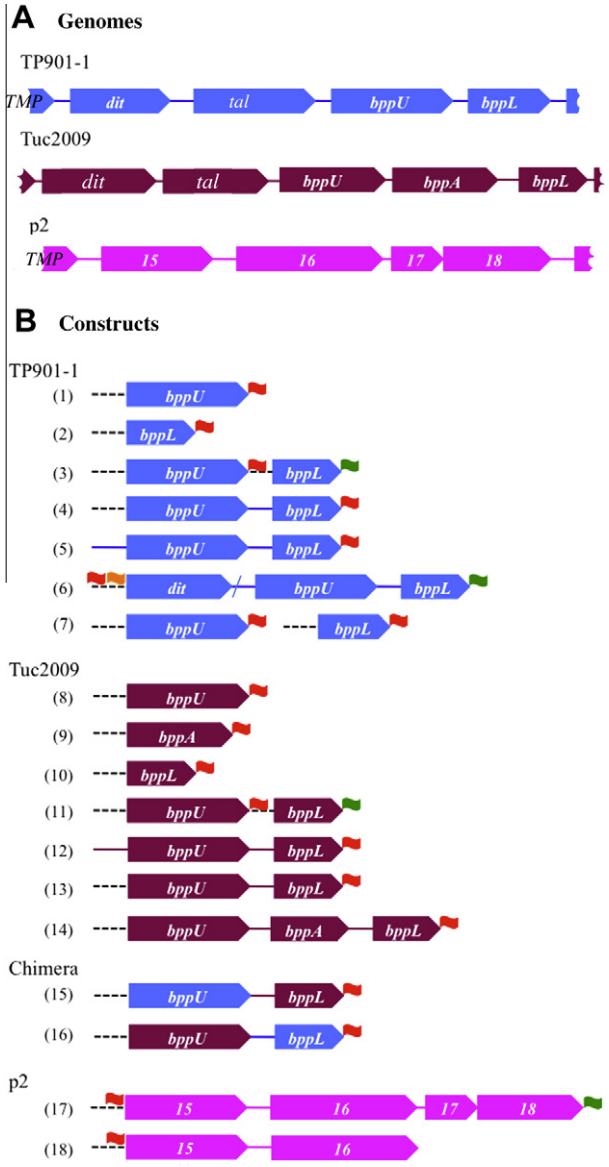

Fig. 1. Schematic representation of the *orfs* constituting the baseplate of the lactococcal phages TP901-1, Tuc2009 and p2 and the expression constructs of their complexes. (A) The structural genes encompassing the *orfs* belonging to the segment following the TMP and finishing at the RBP. (B) Schematic view of the various constructs generated during this study. The numbers in parenthesis correspond to the lines in Table 2. The red, green and orange flags represent His_6 tag, Strep-tag and TRX fusion, respectively. Full lines and dashed lines represent endogenous and engineered RBS, respectively.

that are temporally transcribed. The early genes are expressed just after infection and encode proteins involved in phage replication and host control; the middle genes specify proteins involved in DNA packaging and morphogenesis; and the late genes mostly encode proteins allowing final virion assembly and bacterial host lysis (Duplessis et al., 2005; Madsen and Hammer, 1998). Besides, phage genomes are densely packed as more than 90% of the complete sequence correspond to coding sequences (Hatfull, 2008). Each operon often contains clustered genes having limited distance between each other. For example, in the *L. lactis*-infecting phage TP901-1, genes that encompass *orf37* through to *orf41* all overlap, whereas for *orf41* through to *orf49* (*bppL*) the distance between the stop codon of one gene and the ATG codon of the following gene varies between 1 and 55 bp, with four non-coding intergenic regions of 12 ± 3 bp (Brondsted et al., 2001; Johnsen et al., 1996). Similar densely packed genomes are also observed in other lactococcal phages (Villion et al., 2009) including Tuc2009 (Seegers et al., 2004).

Except for the production of receptor binding proteins (RBPs) (Ricagno et al., 2006; Spinelli et al., 2006a,b; Tremblay et al., 2006), our initial attempts to express the individual components of lactococcal phages and reassemble them in huge complexes failed in a large extent, due to lack of expression, inclusion body formation, protein instability or the absence of interactions. We therefore implemented a partial operon-based cloning strategy using a whole phage genomic segment that codes for two or more phage structural proteins, followed by over-expression in *E. coli*. This strategy yielded excellent expression levels of structural protein complexes from lactococcal phages suitable for biophysical and structural studies. This approach made it possible to obtain the baseplates from phages TP901-1 (ORF46, ORF48 and ORF49) and p2 (ORF15, ORF16 and ORF18). Moreover, we were able to produce chimerical baseplate protein complexes using a protein-shuffling approach between the two related phages TP901-1 and Tuc2009. We then characterized both individual proteins (when available) and complexes by circular dichroism (CD) and SEC/MALS/RI (Veesler et al., 2009a,b). Some of these complexes were also studied by electron microscopy (EM) providing a picture of their overall molecular shapes and symmetry. Finally, we used these results to compare lactococcal phage baseplates organization to the well-known *Myoviridae* phage T4.

2. Materials and methods

2.1. Cloning procedure

All primers used were salt-free and purchased from Eurofins MWG Operon. Primers were designed to introduce protein purification tags as well as *att* recombination sequences. *In silico* optimized primers were designed using the ExSenSo software (Care et al., 2008) and are listed in Table 1. DNA regions containing two or more genes (Fig. 1B) were amplified from genomic DNA using PrimeStar HS DNA polymerase (Takara Bio, Inc.) or Platinum Pfx DNA polymerase (Invitrogen) following manufacturers' instructions. For cloning of chimera or multiple, non-contiguous genes (to form an artificial operon, Fig. 1B), we amplified the individual genes with inclusion of an overlapping sequence before performing a second PCR reaction to achieve joining of the initial PCR products. PCR amplicons were cloned by Gateway recombination (Invitrogen) in pDEST™14 (Invitrogen), pDEST™17 (Invitrogen) or pETG-20A (a kind gift from Dr. Arie Geerlof, EMBL, Hamburg) vectors, as previously described (Vincentelli et al., 2003). The various constructs are listed in Table 2. All clones were confirmed by sequencing (GATC Biotech).

2.2. Protein production

Recombinant plasmids were transformed in *E. coli* Rosetta (DE3)pLysS (Novagen) or T7 Express I^q pLysS (New England Biolabs) strains. Cells were grown at 37 °C in Terrific Broth or M9 minimal medium until the OD_{600nm} reached 0.6, after which protein expression was induced with 0.5 mM IPTG overnight at 25 or 17 °C. Protein purification was performed as previously described (Sciara et al., 2008; Siponen et al., 2009). Briefly, after cell harvesting, lysis was performed by adding 0.25 mg/mL lysozyme, followed by a freezing/thawing cycle and sonication. Soluble proteins were separated from inclusion bodies and cell debris by a 30-min centrifugation step at 20,000*g*. Purification was performed on an ÄKTA FPLC system in two steps: a Ni^{2+}–NTA column (HisTrap Ni 5 mL, GE Healthcare) with a step gradient of 250 mM imidazole, followed by a preparative Superdex 200 HR 26/60 gel filtration. For