Morphology and genome sequence of phage φ1402
A dwarf myovirus of the predatory bacterium
Bdellovibrio bacteriovorus

Hans-W. Ackermann,1 Henry M. Krisch2,3 and André M. Comeau4,*

1Department of Microbiology; Faculty of Medicine; Laval University; Québec-Océan; Département de Biologie; and Institut de Biologie Intégrative et des Systèmes (IBIS); Université Laval; Québec, QC Canada; 2Centre National de Recherche Scientifique; 3Université de Toulouse; UPS; Laboratoire de Microbiologie et Génétique Moléculaires; Toulouse, France

Key words: Bdellovibrio phage, dwarf myovirus, complete genome, terminase, capsomers

Abbreviations: ORF, open-reading frame; UTR, untranslated region

Introduction
Bdellovibrio bacteriovorus is a diminutive, motile, curved, rod-shaped, Gram-negative bacterium of the δ-Proteobacteria class. Bdellovibrios live in freshwater, seawater and soil, but are best known for their predatory activity in which they can penetrate into the periplasmic space of Gram-negative bacteria and utilize their constituents as carbon and energy sources.1 The first Bdellovibrio phage described in B. bacteriovorus was a tailless icosahedral virion with a ssDNA genome and a double capsid,2 an otherwise unheard of phage capsid structure that was probably an artifact of the uranyl acetate staining. Additional Bdellovibrio phages were isolated and described in reference 3–5 in 1972, including another ssDNA phage and two small phages of the Myoviridae family. Unfortunately, none of these viruses are currently available.

Thirty years later a new isolate, ϕMH2K, became the first Bdellovibrio phage to have its genome sequenced.6 This bacteriophage is morphologically similar to the classical coliphage ϕX174 of the Microviridae family, but phylogenetically it is quite distant from it, being more closely related to the Microviridae phage Ch-1 of Chlamydia. In 2000, B. Fane of the University of Arizona at Tucson isolated a small Bdellovibrio myovirus, ϕ1402, from sewage (Fane B, personal communication). Soon afterwards this phage’s morphology was examined by electron microscopy by H.W. Ackermann and its DNA was preserved in his laboratory collection at Laval University. Given the paucity of information available about the phages infecting this unusual and interesting group of predatory bacteria, we have now sequenced the genome of ϕ1402. Here we report on both phage ϕ1402’s morphology and genome sequence. This phage represents the smallest and simplest autonomous myovirus currently known. Of the 42 putative ORFs encoded by its genome, only two are clear homologs of any known phage genes.

Results and Discussion
Phage ϕ1402 has a contractile tail and thus is classified as a myovirus (Fig. 1). As measured on 20 particles, it has an elongated 68 x 40 nm head and a tail that is, in its extended state, 62 x 17 nm. The tail has a neck that lacks a collar and, hence, collar-associated appendages. In the contracted state, the tail measures 20 x 20 nm and shows a base plate of 17 x 2 nm that separates from the
The DNA pac site which is diagnostic of phages using a Restriction digest profiles (data not shown) identified a fragment containing a pac site which is diagnostic of phages using a terminally redundant, head-full packaging strategy. The DNA sequence encodes 42 ORFs, which are nearly all located on the positive strand and have no significant overlaps. The genome has no obvious phylogenetic relationship to any other phage in public databases. Remarkably, only two of its ORFs can clearly be considered as phage-related: ORF11, the large subunit of the phage terminase and ORF33 which is related to various fibrous phage proteins (Table 1). Both these sequences are among the most highly conserved and ubiquitous sequences in phage genomes. Five other \( \phi 1402 \) proteins had significant \( E \)-values \( (\leq 10^{-4}) \) when compared to cellular proteins: ORF01 is related to ssDNA-binding proteins; ORF10 belongs to an RNA-binding Conserved Domain Database (CDD) protein family; and three ORFs have homologies to hypothetical cellular proteins. The remaining 35 proteins (83%) are ORFans—database entries without any known homologs. Iterative PSI-BLAST, applying less stringent selection criteria, found weak hits to three additional ORFs (ORF21/24/38). These are putatively identified as dehydrogenases/reductases, potassium/proton antiporters and peptidoglycan-binding domain proteins, respectively, but the significance of such weak homologies is dubious.

The highly compact \( \phi 1402 \) genome (95% coding) contains no tRNAs and dot-plot analysis reveals no significant direct- or inverted-repeat sequences. A comparison of the \( \phi 1402 \) genome and that of its host revealed that both had similar GC contents and codon usage patterns (data not shown). The intergenic regions of the phage genome contained no obvious consensus promoter sequences; however, the host promoter sequences of Bdellovibrio spp. have not yet been defined either. The \( \phi 1402 \) genome contains a single large (-1 kb) untranslated region (UTR) located between the end of the forward-transcribed genes and the small cluster of genes transcribed leftward. Although this UTR has a substantially reduced GC ratio, there are no additional indications that it corresponds to an \( attP \) site of a temperate phage. \( \phi 1402 \) always behaved as a standard lytic phage, with no particularities regarding plaque morphology, stock production, nor latent period (Fane B, personal communication). However, BLASTn did find an identical 16 bp sequence (GTA ACT CCT CAA GAA T) in the phage UTR and in a \( Bdellovibrio bacteriovorus \) intergenic sequence (coordinates 3,528,980–3,528,995 on NC_005363) that is located between an asparaginase gene and a gene for a hypothetical protein, thus quite unlike the site of many temperate phage \( attP \) sites that are often in close proximity to tRNA sequences. The function of this large phage UTR region, therefore, remains unknown. Although an investigation of the DNA sequence with Ori-Finder was inconclusive, there are 6 putative shifts in GC skew indicating possible origins of replication, with the strongest candidate being near the 0 coordinate.

In conclusion, phage \( \phi 1402 \) differs from other myoviruses by its extremely small size, the presence of visible capsomers, and its very small dsDNA genome. Tailed-phage capsids are generally smooth; the visualization of capsomers after negative staining is exceptional in tailed phages and many indicate structural differences separating \( \phi 1402 \) from other phages. Among the myoviruses, only the satellite coliphages P4 and \( \phi R73 \) have smaller DNAs, of 11.6 and 12.7 kb, respectively. Both of these have small isometric heads (~45 nm diameter) and tails of approximately 140 nm in length. A phylogenetic analysis of the \( \phi 1402 \) terminase protein (Fig. 3) further reinforces the notion that this unusual Bdellovibrio phage has diverged considerably from all other currently known phage groups. In the absence of any evidence that \( \phi 1402 \) is a satellite phage, it appears to represent the smallest known autonomous myovirus and thus an interesting example of reductionism carried to the extreme. It will be interesting to identify all the functions that \( \phi 1402 \) has been able to jettison on its evolutionarily reductionist pathway and, equally important, what few genes it was constrained to retain. It seems likely that even this small “simplistic” myovirus would require at least 25 genes to encode all the essential virion components and the associated genes for correct assembly of the virion. Such a conservative estimate leaves only 17 functions unaccounted for in the \( \phi 1402 \) genome, but many of these must be involved in the phage’s takeover of host macromolecular synthesis and phage transcription/replication. It will be interesting to see what the few...
remaining unaccounted for phage functions do and how they do it with such modest genetic resources.

**Materials and Methods**

**Electron microscopy.** The phage was pelleted at 25,000 x g for 1 hour, using a Beckman high-speed centrifuge and a JA-18.1 fixed-angle rotor (Beckman, 347824). The phage pellet was washed twice in neutral 0.1 M ammonium acetate and re-suspended. The phages were then deposited on copper grids with carbon-coated Formvar films and stained with 2% uranyl acetate (pH 4) and 2% phosphotungstic acid (pH 7.2). They were then examined in an electron microscope (Philips, EM300) whose magnification was calibrated using T4 phage tails as size standards.

**DNA extraction and sequencing.** Phage DNA was extracted 3 times with 1:1 v/v phenol: chloroform, amended with 0.25 M (final concentration) sodium acetate, and then precipitated immediately with 100% room-temperature ethanol. After a spin at 15,000 g for 3 min, the pellet was washed with 70% ice-cold ethanol and re-spun. The pellet was dried at 37°C for 1 h and re-suspended in sterile Milli-Q water and left to rehydrate at 4°C. The resulting pure DNA was used for restriction digests, bar-coded library construction and 454 pyrosequencing that was performed according to the manufacturer’s instructions on a quarter picotiter plate of a GS-FLX sequencer (Roche) at the IBIS/Université Laval Plate-forme d’Analyses Génomiques.

**Bioinformatics.** Raw reads were assembled using the GS De Novo Assembler (Roche), resulting in one final contig with a 330-fold coverage. Analysis of the genome was done with the following programs: (1) GLIMMER (www.ncbi.nlm.nih.gov/genomes/MICROBES/glimmer_3.cgi; >100 nt; bacterial genetic code) and GeneMark (exon.gatech.edu/GeneMark; heuristic approach for prokaryotes and viruses; >90 nt) for ORF determinations; (2) tRNA search using tRNAscan-SE (lowelab.ucsc.edu/tRNAscan-SE); (3) Java Word Frequencies and Java Dot Plot Alignments (athena.bioc.uvic.ca) for the exploration of DNA “words”/patterns; (4) Graphical Codon Usage Analyser (gcua.schoedl.de) and the Codon Usage Database (www.kazusa.or.jp/codon) for the exploration of codon usage patterns; (5) the BLAST tools at NCBI (blast.ncbi.nlm.nih.gov) for the characterization of genes/proteins and untranslated regions (UTRs) of the DNA; (6) various phylogenetic tools of the Mobyle Project.

![Figure 2. Genome sequence of φ1402. The circles, from outermost to innermost, represent: the scale in kilobases; the rightward-transcribed ORFs (blue); the leftward-transcribed ORFs (red); the %GC content (black for above-average and gray for below-average); the GC skew (orange for positive and purple for negative). Putative origins are indicated with red triangles and the 7 ORFs with assignable homologs are indicated.](image-url)
Table 1. \(\phi1402\) ORFs with identifiable homologs/gene functions

| ORF | Strand | Start | End | ORF length | aa length | E-value | Homologs/function* |
|-----|--------|-------|-----|------------|-----------|---------|-------------------|
| 01  | +      | 1     | 1098| 1098       | 365       | 6e-07   | Erf family protein [Enterococcus faecalis]; SSB-like pfam04404 |
| 10  | +      | 3886  | 4224| 339        | 112       | 7e-08   | ASCH domain protein [Listeria monocytogenes]; RNA-binding cd186302 |
| 11  | +      | 4253  | 5797| 1545       | 514       | 2e-11   | Phage terminase, large subunit [Enterococcus phage phiFL2A] |
| 14  | +      | 7582  | 8316| 735        | 244       | 2e-07   | Hypothetical protein [Populus trichocarpa]; many other conserved hyp. proteins |
| 32  | +      | 17666 | 18874| 1209       | 402       | 8e-04   | BNR domain protein [Azotobacter vinelandii] |
| 33  | +      | 18884 | 19576| 693        | 230       | 2e-16   | Phage tail collar domain protein [Ralstonia phage RSL1]; many phage fibers |
| 39  | +      | 21591 | 22475| 885        | 294       | 3e-29   | Hypothetical protein [Microcystis aeruginosa] |

**Weak hits**

| ORF | Strand | Start | End | ORF length | aa length | E-value | Homologs | |
|-----|--------|-------|-----|------------|-----------|---------|----------|
| 21  | +      | 11093 | 12373| 1281       | 426       | >1e-04  | Many dehydrogenases/reductases |
| 24  | +      | 13284 | 14273| 990        | 329       | >1e-04  | Potassium/proton antiporter [Bradyrhizobium] |
| 38  | +      | 20863 | 21594| 732        | 243       | >1e-04  | Peptidoglycan binding domain protein [Roseiflexus] |

*Generally listed are the closest phage (if any) homologs of identifiable function, followed by identifiable cellular homologs, and then hypothetical proteins (if no function could be deduced) from blastp searches against the nr (or phage restricted) database using an E-value cutoff of <10^{-4}. Also included are additional hits/homologs of interest and hits to the Conserved Domain Database (CDD) with their cd/pfam identifiers.

Figure 3. Large subunit terminase phylogeny. Neighbor-joining tree of the \(\phi1402\) terminase with its 7 closest phage homologs. Values at the nodes indicate the results of 100 bootstrap replicates and the scale bar indicates 0.1 substitutions per site.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgements

The authors thank Dr. Bentley M. Fane (BIO5 Institute, University of Arizona-Tucson) for his kind gift of phage \(\phi1402\) and Dr. Brian Boyle at the IBIS/Université Laval Plate-forme d’Analyses Génomiques (Québec, QC) for sequencing and assembly advice. H.M.K.’s research was primarily supported by
intramural funding from the INSB of the CNRS and by supplementary funding from the Kribu Foundation.

References
1. Socket RE. Predatory lifestyle of *Bdellovibrio bacteriovorus*. Annu Rev Microbiol 2009; 63:523-39; PMID:19575566; http://dx.doi.org/10.1146/annurev.micro.091208.073346.
2. Hashimoto T, Dödrich DL, Conti SF. Isolation of a bacteriophage for *Bdellovibrio bacteriovorus*. J Virol 1970; 5:97-8; PMID:4191817.
3. Althauer M, Samsonoff WA, Anderson C, Conti SE. Isolation and preliminary characterization of bacteriophages for *Bdellovibrio bacteriovorus*. J Virol 1972; 10:516-23; PMID:4116140.
4. Schindler J, Ludvik J. A new bacteriophage specific for a saprophytic mutant of *Bdellovibrio bacteriovorus*. Acta Virol 1972; 16:501-2; PMID:4118526.
5. Varon M, Levisohn R. Three-membered parasitic system: a bacteriophage, *Bdellovibrio bacteriovorus* and *Escherichia coli*. J Virol 1972; 16:519-25; PMID:452554.
6. Brendlinger KL, Hafenstein S, Novak CR, Fane BA, Borgon R, McKenna R, et al. *Microviridae*, a family divided: isolation, characterization and genome sequence of dMH2K, a bacteriophage of the obligate intracellular parasitic bacterium *Bdellovibrio bacteriovorus*. J Bacteriol 2002; 184:1089-94; PMID:11807069; http://dx.doi.org/10.1128/jb.184.4.1089-94.2002.