Genomic Characterization of *Ralstonia solanacearum* Phage φRSB1, a T7-like Wide-host-range Phage

Takeru Kawasaki, Mio Shimizu, Hideki Satsuma, Akiko Fujiwara, Makoto Fujie, Shoji Usami and Takashi Yamada*

Department of Molecular Biotechnology, Graduate School of Advanced Sciences of Matter,
*Hiroshima University, Higashi-Hiroshima 739-8530, Japan*

Running title: NOTES

*Corresponding author:
Dr. Takashi Yamada
Department of Molecular Biotechnology,
Graduate School of Advanced Sciences of Matter,
Hiroshima University, 1-3-1 Kagamiyama,
Higashi-Hiroshima 739-8530, Japan
Tel/Fax: +81-82-424-7752.
E-mail: tayamad@hiroshima-u.ac.jp
Abstract

φRSB1 is a wide-host-range, T7-like bacteriophage that infects and efficiently lyses the phytopathogenic bacterium Ralstonia solanacearum. The φRSB1 genome comprises 43,079-bp dsDNA (61.7% G+C) with 325-bp terminal repeats and contains 47 open reading frames. Strong activity of tandem early promoters and wide specificity of phage promoters of φRSB1 were demonstrated.

Text

The phytopathogenic gram-negative bacterium Ralstonia solanacearum causes bacterial wilt disease in many important crops (10). Recently, Yamada et al. (9, 12, 21) isolated and characterized various kinds of bacteriophage that specifically infect R. solanacearum strains belonging to different races and/or biovars. These phages may be useful as a tool not only for molecular biological studies of R. solanacearum pathogenicity but also for diagnosis and biocontrol of bacterial wilt. In this study, we report the genome and characteristic features of a new phage φRSB1. φRSB1 was isolated from a soil sample from a tomato crop field, and was selected for its ability to form large clear plaques on plate cultures of R. solanacearum strain M4S (for details of bacterial strains see 21). Plaques formed on assay plates (CPG) were 1.0-1.5 cm in diameter. This phage has a wide host range and infected 13 of 15 strains tested, including strains of races 1, 3 and 4 and of biovars 3, 4, and N2. Under laboratory conditions (in standard one-step growth assays) host cells of R. solanacearum strains lyses after 2.5-3 h p.i. (with an eclipse period of 1.5-2h), releasing approximately 30-60 pfu new phage particles per cell (burst size) (data not shown). Electron microscopic observation of negatively stained phage particles revealed short-tailed icosahedral structures resembling those of the family Podoviridae. The particles consisted of a head approximately 60 nm in diameter and a stubby
The φRSB1 genome is linear double-stranded (ds) DNA of approximately 43.0 kbp in size as determined by pulsed-field gel electrophoresis (data not shown). Since no oligomeric forms were formed after heat-treatment, cohesive ends are absent. The sequence of the φRSB1 genome was determined using DNA purified from phage particles by shotgun sequencing and primer walking strategy (9). Sequences were assembled into a circular contig of 42,754 bp, suggesting the presence of long terminal repeats. The precise sequence of the repeat was determined by direct sequencing of genomic DNA with outward-directed primers, located outside the possible terminal repeat region. The final sequence of the φRSB1 genome is 43,079 bp, and includes direct terminal repeats of 325 bp. The φRSB1 genome size is comparable with that of Pseudomonas aeruginosa phage φKMV (42,159 bp; accession no. AJ505558) and a little larger than coliphages T7 (39,937 bp; accession no. NC_00164) and T3 (38,208 bp; accession no. NC_003298). The size of the φRSB1 terminal repeats is smaller than that of φKMV (414 bp) but larger than that of T7 (160 bp) or T3 (231 bp). The G + C content of the genome is 61.7%. This value is lower than the G + C values of the large and small replicons of the R. solanacearum GMI1000 genome (67.04% and 66.86%, respectively) (18). Potential ORFs consisting of more than approximately 50 codons and starting with ATG, GTG, or TTG were identified using the Orfinder and DNASIS programs. The presence of a Shine-Dalgarno ribosome-binding sequence preceding the initiation codon was taken into account for ORF prediction. Possible functions were assigned to ORFs by searching through databases using BLAST, BLASTX, and BLASTP programs (1). Accordingly, a total of 47 potential ORFs oriented in the same direction were assigned on the genome (Fig. 1A and Supplementary data). To find homologous sequences, nucleotide sequences from φRSB1 DNA were used to searched databases with BLAST and BLASTX programs. Patchy or local
homologies were detected in the genomic sequences of various phages including

*Xanthomonas oryzae* phages Xop411 (accession no. DQ777876) and Xp10 (AY299121; 22),

*Pseudomonas aeruginosa* phages φKMV (accession no. AJ505558; 14), LKD16 and LKA1

(6), *Erwinia amylovora* phage Era103 (accession no. EF160123), and *Burkholderia cepacia*

phage BcepF1 (accession no. AY616033). All of these are members of *Podoviridae*. The

genome of coliphage T7, the representative of T7-like viruses of the *Podoviridae*, generally

consists of three functional gene clusters; one for early functions (class I), one for DNA

metabolism (class II), and the other for structural proteins and virion assembly (class III) (8). These gene clusters are essentially conserved in the φRSB1 genome. Figure 1A shows

putative ORFs identified on the φRSB1 genome compared with ORFs from other phages;

*Xanthomonas* phage Xop411 (giving the highest local similarities), *Pseudomonas* phage

φKMV (showing marginal similarity but longest regions of similarity), and coliphage T7. The

mosaic genetic relationship of φRSB1 indicates frequent recombinations on the φRSB1

ancestral genome during its evolution, in the way suggested for tailed phages and their

prophages (2-5, 11).

One of the characteristic features found in the φRSB1 gene organization is that the

predicted gene for RNA polymerase (RNAP) of φRSB1 (orf26) is not located in the early

gene region (class I) but at the end of the class II region (Fig. 1A). Another exception is the

gene for DNA ligase (DNAL); orf25 encoding the φRSB1 DNA ligase is in front of the RNAP

ORF (orf26), whereas the gene encoding T7 DNAL is downstream of the gene for RNAP at

the end of the class I cluster (8). In *Pseudomonas* phages φKMV, LKD16, and LKA1, the

dna ligase gene is upstream of the gene for DNAP in the class II gene cluster (Fig. 1A).

Similarly to the T7 genome, structural proteins are predicted to be encoded in the class III

gene cluster of the φRSB1 genome. Purified φRSB1 particles gave at least 10 protein bands
on SDS-PAGE (Fig. 2). Each band was extracted from the gel and was subjected to N-terminal amino acid sequencing (19). The N-terminal sequence of each protein always started from the second amino acid residue of its corresponding ORF, except for ORF35, which included the first methionine (Fig. 2). In addition to known structural proteins, ORF35, ORF36 and ORF46 were identified as structural proteins. In this way, all predicted protein in the class III-structural region were identified, except for the scaffolding protein (ORF31) and a possible tail fiber protein ORF38, which may be lost during purification of the phage particles. In the case of the largest structural protein (170-180 kDa), determination of the N-terminal sequence was unsuccessful using standard methods, possibly because of modification at the N-terminus. However, it most likely corresponds to ORF37, judged from its exceptionally large size (174 kDa); there is no other candidate for this size. ORF37 may encode a tail protein with a transglycosylase domain.

T7-like phages are generally known as absolute lytic phages, with a few exceptions, such as integrase-coding phages, e.g. prophage 3 of Pseudomonas putida (17), and the cyanophage P-SSP7 (16). Sometimes nucleotide sequences related to T7-like phages are found in conjunction with other temperate phages such as λ−like phages that are integrated in various bacterial genomes (2-5, 11). BLAST and BLASTX database searches using the φRSB1 sequence revealed a significantly homologous region (at the nucleotide sequence level) in the genome of Burkholderia pseudomallei 1710b (accession no. CP000124). A matrix comparison plot showed that this homology is extended to a 20 kbp region (1710b positions 1,740,980-1,761,000) (data not shown). In the 1710b genome, this region is embedded in a large (85 kbp) prophage sequence (1710b positions 1,719,000-1,804,000), which is related to λ−like phages such as B. pseudomallei phage φ1026b (7) and B. thailandensis phage φE125 (20). The homologous region of 1710b-prophage contains 8 ORFs encoding a DNA primase,
DNA helicase, DNA ligase, DNA polymerase, exonuclease, and RNA polymerase etc. These correspond to the class II genes of φRSB1 as shown in Fig. 1B. Interestingly, the putative φRSB1 promoters (see below) were also found in this 1710B region (positions 1743597-1743650 and 1744983-1745040; Fig. 1B). Both *Ralstonia* and *Burkholderia* belong to *Betaproteobacteria* and may share common bacteriophages (9). Database search also showed that a 10 kbp genomic region of the *R. solanacearum* GMI1000 (positions 1,661,000-1,672,000) contains a 1589-aa ORF (RSO5240) showing significant similarity to φRSB1 ORF37, which encodes a putative transglycosylase-tail protein (Supplementary data, E-value e-131). Amino acid sequence similarity extends to the entire region consisting of the N-terminal transglycosylase and C-terminal core or tail domains (1606 aa in φRSB1 ORF37).

This GMI1000 ORF is associated with two IS transposase sequences (ISRSO8 transposase A and B; RSO5237 and RSO5236, respectively) on the left side. Immediately to the right of this ORF, there is an ORF (RSO5241) for a putative integrase, which is closely associated with arginine tRNA (AGA), a possible att sequence (Fig. 1C). This structure indicates horizontal acquisition of this φRSB1 ORF by host cells, as well as some involvement of the phage integrase/att sequence and transposons in such an event. In the context of lysogenic conversion or introduction of a new fitness factor by phage in the pathogenic bacteria, the functions of φRSB1 ORF37 are interesting.

As shown in Fig. 3, several putative transcription promoter and terminators were identified in apparently noncoding regions (more than 100 bp long) in the φRSB1 genome. A typical prokaryotic promoter sequence (resembling *E. coli* σ^70^) was repeated five times (p1-p5) in a left 1,000-bp region without ORFs (Fig. 1A and Fig. 3A). In addition, a few other putative sequences of host σ^70^ promoter (p6-p8) were detected in front of ORFs 1, 17, and 39 (Fig. 1A and Fig. 3A). There are possible ρ-independent terminator-like sequences (Fig. 3C).
A terminator-like sequence (t2) present after ORF13 is located in the region that separates class I and class II genes (Fig. 1A). Another possible terminator (t3) is located immediately downstream of ORF32, encoding the major capsid protein. t4 is located in front of a putative promoter p8 for ORF40 and ORF41 (similar to the large subunit of a terminase). A final terminator (t5) was defined behind the last ORF47. The terminator positions of t2, t3, and t5 are consistent with those reported in Pseudomonas phages φKMV (14), LKD16 and LKA1 (6) (Fig. 1A). Searching for core promoter-like sequences conserved in phages T3, T7, or SP6 in the φRSB1 intergenic regions could not find any significant ones. Instead, three sets of common sequence elements were found in front of ORFs 16, 18, and 32 (designated φp1, φp2, and φp3, Fig. 1A). We found a set of sequence elements consisting of a GC-rich stretch, and TTGT, TCTGG, and CGGGCAC motifs preceding an AG-rich Shine-Dalgarno sequence (Fig. 3B). Activity of transcriptional promoters of both host- and phage-types thus predicted on the φRSB1 genome was examined using a GFP-expressing single-copy plasmid pRSS12 (13), where the lac promoter for gfp expression was replaced with a φRSB1 promoter sequence. When we tested bacterial σ70-type promoters p1-p4, p1-p5, and p1-p6, which are located tandemly at the beginning of the class I gene cluster, transformed cells of R. solanacearum strains always showed strong GFP fluorescence. Fluorescence was 3-15 times greater than that of pRSS12 with a lac promoter. Results with strain MAFF301558 as the host are shown in Table 1. Increased GFP intensity from p1-p4 via p1-p5 to p1-p6 clearly demonstrates actual promoter activities of these φRSB1 early promoters. φp1 is located at the beginning of the class II gene cluster, after the possible terminator t2 and in front of ORF16 encoding possible DNA primase. φp3 is located upstream of ORF32, which encodes the major capsid protein. Both φp1 and φp3 also function as promoters for bacterial RNA polymerase in φRSB1-uninfected R. solanacearum cells, but show lower activity compared with p1-p6
(Table 1). The promoter activity was also examined in *R. solanacearum* cells after infection with φRSB1. As strain MAFF301558 was found to be a low-efficient host, giving fewer titers of phage progeny, the host was changed to strain M4S. After infection with φRSB1, GFP fluorescence intensity retained at almost the same levels in cells containing the promoters, p1-p5, or p1-p6, whereas cells with φp1 or φp3 showed increased GFP fluorescence after 30 min p.i. to 90 min p.i. (Table 1). At 120 min p.i. cell lysis began. These results indicate that φp1 and φp3 are functional in transcription by both bacterial and phage RNAPs. Bacterial σ70-type promoters are not shut down but continue to function after infection.

The occurrence of host σ70-type promoter sequences in the late gene clusters, class II and class III, and the low specificity of phage promoters further imply that expression of φRSB1 genes is highly dependent on the host RNAP. To determine whether host RNAP is involved in late stages of φRSB1 infection, rifampicin was added to φRSB1-infected cultures at various times post infection and the number of progeny phage was determined. The results are shown in Fig. 4. In samples that were incubated with rifampicin, more than 90% of phage progeny was obtained when the drug was added at 90 min p.i. or later, and no or very few progeny phages were obtained when the drug was added 75 min p.i. or earlier. These results indicate that a switch from host RNAP to φRSB1 RNAP occurs between 75 min p.i. and 90 min p.i. and that late stages of φRSB1 replication are independent of rifampicin. The late genes can be transcribed by rifampicin-resistant φRSB1 RNAP, at least in the presence of rifampicin.

This study was supported by the Industrial Technology Research Grant Program (04A09505) from the New Energy and Industrial Technology Development Organization (NEDO) of Japan. The sequence data for φRSB1 genomic DNA has been deposited in the DDBJ database under accession no. AB451219.
REFERENCES


characterization of Ralstonia solanacearum phage φRSA1 and its related prophage
(φRSX) in strain GMI1000. J. Bacteriol. **190**:143-156.

10. **Hayward, A. C.** 1991. Biology and epidemiology of bacterial wilt caused by


2007. Genomic characterization of the filamentous integrative bacteriophage
φRSS1 and φRSM1, which infect Ralstonia solanacearum. J. Bacteriol. **189**:5792-5802.

phytopathogenic Ralstonia solanacearum cells using green fluorescent protein-expressing

The genome of bacteriophage φKMV, a T7-like virus infecting Pseudomonas aeruginosa.

15. **Liao, Y. D., J. Tu, and T. T. Kuo.** 1987. Regulation of transcription of the Xp10 genome

2004. Transfer of photosynthesis genes to and from Prochlorococcus viruses.


18. **Salanoubat, M., S. Genin, F. Artiguenave, J. Gouzy, S. Mangenot, M. Ariat, A.**


19. Figure legends

20. FIG. 1. Genetic organization and comparative analyses of φRSB1 genome. (A) Comparison of the genomes of φRSB1 and T7-like phages. In each alignment, corresponding ORFs (horizontal arrows) are connected by shading. Three functional gene clusters, class I (green), class II (Yellow), and class III (orange) are indicated above the φRSB1 and the T7 maps and
corresponding ORFs are colored. Putative bacterial promoters, phage promoters, and
terminators of transcription are indicated under the φRSB1 map by p, φp, and t, respectively.
Promoters and terminators are also shown in the φKMV and T7 maps. Xop411, *Xanthomonas oryzae* phage (44,520 bp, accession no. DQ777876); φKMV, *Pseudomonas aeruginosa* phage (42,159 bp, AJ50558); coliphage T7 (39,937 bp, NC_00164). DNAP, DNA polymerase;
DNAL, DNA ligase; RNAP, RNA polymerase; MCP, major capsid protein; LYS, lysozyme.
(B) Class II region of φRSB1 genome (ORF16-ORF26) often shows high homology with
phage or prophage sequences of different phage groups. The φRSB1 region is aligned with the
corresponding prophage sequence of *Burkholderia pseudomallei* 1710b (accession no.
CP000124). (C) Region on the GMI1000 genome (positions 1,661,000-1,672,000) containing
a large φRSB1-ORF37-like ORF (RSO5240), which encodes a putative transglycosylase
protein. This GMI1000 region is flanked by two ORFs encoding ISRSO8 transposase A and B
on the left side and by integrase (RSO5241) and argnine tRNA (AGA) on the right side.

FIG. 2. Identification of φRSB1 virion proteins. Proteins from purified φRSB1 particles
were separated by SDS-PAGE (10% gel) and stained with Coomassie blue. Molecular size of
each marker protein (Amersham full-range MW markers and a LMW gel filtration calibration
kit) is indicated on the left. N-terminal amino acid sequence (5 residues) determined for each
φRSB1 protein band is shown on the right with its corresponding ORF. Amino acids in
parenthesis are obscure residues. Although the N-terminal sequence could not determined, the
largest protein, approximately 175-180 kDa, is predicted to be ORF37, as there is no other
candidate for this large size. A few small proteins were lost from the gel during
electrophoresis.
FIG. 3. Predicted regulatory sequences found in the φRSB1 genome. (A) *E. coli* σ^{70}-promoter-like sequences, (B) putative promoter sequences for φRSB1-encoded RNAP, and (C) putative terminators.

FIG. 4. Late stages of φRSB1 development are resistant to rifampicin. Cells of *R. solanacearum* M4S were infected with φRSB1 at moi 5. At the indicated times p.i., aliquots of the infected culture were withdrawn, and incubated with 100 µg/ml rifampicin for 2.5 h before CHCl₃-treatment and determination of phage titers (pfu) (open circle). For control, phage titers were also determined at indicated times after CHCl₃ addition without rifampicin-treatment (closed circle). The method is according to Liao et al. (15).
TABLE 1. Expression of *gfp* by *φ*RSB1 promoters

<table>
<thead>
<tr>
<th>Promoter (positions)</th>
<th>Strain MAFF301558</th>
<th>Strain M4S</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Relative intensity of GFP fluorescence$^a$</td>
<td>Time (min) post infection of <em>φ</em>RSB1</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>lac</td>
<td>3.5 ± 0.2 (1.0)$^b$</td>
<td>1.8 ± 0.1</td>
</tr>
<tr>
<td>p1-p4 (425-845)</td>
<td>10.0 ± 0.3 (3.0)$^b$</td>
<td>ND</td>
</tr>
<tr>
<td>p1-p5 (425-921)</td>
<td>14.3 ± 0.5 (4.1)$^b$</td>
<td>4.2 ± 0.2</td>
</tr>
<tr>
<td>p1-p6 (425-995)</td>
<td>53.0 ± 0.9 (15.2)$^b$</td>
<td>7.3 ± 0.3</td>
</tr>
<tr>
<td><em>φ</em>p1 (6895-7009)</td>
<td>11.0 ± 0.5 (3.2)$^b$</td>
<td>6.7 ± 0.3</td>
</tr>
<tr>
<td><em>φ</em>p3 (23515-23675)</td>
<td>7.2 ± 0.3 (2.1)$^b$</td>
<td>1.4 ± 0.1</td>
</tr>
</tbody>
</table>

$^a$The values are means ± standard errors for data from three independent experiments.

$^b$The ratio to *lac*-value (1.0) is in parenthesis.

$^c$ND: Not determined.
### A

**E. coli σ^{30}-like promoters**

<table>
<thead>
<tr>
<th>Positions</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1 508-552</td>
<td>CAGAGCTTGACAAGACACACAGCAGCACAGC---TACAGTGGAGTCACAC</td>
</tr>
<tr>
<td>p2 589-633</td>
<td>CAGAGCTTGACAAGCGGACGGACTAGCA---TACAGTGGAGTCCAGC</td>
</tr>
<tr>
<td>p3 709-753</td>
<td>GCAGCGTTGACAAGCGTACGGTGAAGCTGAGCTGAGGACGGGACACTAGC</td>
</tr>
<tr>
<td>p4 801-845</td>
<td>GTGCCTTGACAAGCGGACGGAGGTACGGACTGAGC---TACACTGCAGTCACGT</td>
</tr>
<tr>
<td>p5 877-925</td>
<td>GGATGCTTTGACATTGACGTTGAGGACGGGACAGCTGAGGACGGGACACTAGC</td>
</tr>
<tr>
<td>p6 934-982</td>
<td>CGCGCTTTGACCTACAGGTATAACAGGATACGTACGAGGACGGGACACTAGC</td>
</tr>
<tr>
<td>p7 8113-8156</td>
<td>GACCAACTCGACAACTGAGGCTTGGCTT---TGTGAGTCCAGATTCAGG</td>
</tr>
<tr>
<td>p8 38455-38498</td>
<td>TCCGTCCTTGACATTGACGTTGAGGACGGGACAGCTGAGGACGGGACACTAGC</td>
</tr>
</tbody>
</table>

**E. coli σ^{30}**

------TTGACA\((−35)\)-----------------TATAAT\((−10)\)------

### B

**Putative phage promoter sequences**

<table>
<thead>
<tr>
<th>Positions</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>φp1 6926-6960</td>
<td>CGA GCCGGGCATCATCCTTGCCCTGCTCTCTGGAGATTCTAGGAGCGGGCACCTT---30b---ATG</td>
</tr>
<tr>
<td>φp2 9688-9721</td>
<td>AGCTGGGGGCGGGGTATTGATCGGAATTGGAGCGGGTATGCT---29b---ATG</td>
</tr>
<tr>
<td>φp3 23561-23593</td>
<td>AGGCTGGGGGCGGGCGTTTTGGCCCTGCTCTGCTCCACAAATG</td>
</tr>
<tr>
<td>Consensus</td>
<td>---GCstretch---TTGT----TCTGG----------------CGGGGCAC------------ATG</td>
</tr>
</tbody>
</table>

### C

**Putative phage terminators**

<table>
<thead>
<tr>
<th>Positions</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>t1 1808-1853</td>
<td>GACGTACATGCCCCTTCTGCGGAGTGCGAGCACATGTCATT</td>
</tr>
<tr>
<td>t2 6465-6501</td>
<td>ACTGCCTGGCTGGATAACGGAGCTCTGAGCCGGCCGAC</td>
</tr>
<tr>
<td>t3 24690-24727</td>
<td>CCCAACTGCCGTCCAACACTCAAGGTGGGCGGCGATTTTT</td>
</tr>
<tr>
<td>t4 38428-38460</td>
<td>TGGGATACCCGATTGCTGATTGGGTATCCGTC</td>
</tr>
<tr>
<td>t5 42791-42864</td>
<td>GTGTGCGCGCGCGTCGTCGATGACGTACGAGTCGAGGTGCGCATGTCATT</td>
</tr>
</tbody>
</table>

**Consensus**

---GCstretch---TTGT----TCTGG----------------CGGGGCAC------------ATG