Lysogenic bacteriophage M1 from *Selenomonas ruminantium*: isolation, characterization and DNA sequence analysis of the integration site

Judy P. E. Cheong and John D. Brooker

**INTRODUCTION**

Rumen micro-organisms play an essential role in ruminant nutrition (Hungate, 1966) and the bacterial population density may be as high as $2 \times 10^{10}$ c.f.u. ml$^{-1}$ in a normal sheep (Warner, 1962). In view of the high bacterial density, it is not surprising to find bacteriophages in the rumen and indeed a large number of lytic and temperate phages have been reported (Orpin & Munn, 1974; Inversor & Millis, 1976; Lockington et al., 1988; Klieve & Bauchop, 1988; Klieve et al., 1989; Styriak et al., 1991; Klieve & Swain, 1993). However, despite potential effects of bacteriophage infections on the ruminal bacterial population, there have been few detailed molecular analyses of any ruminal phage. Gregg et al. (1994) described the *attP* region and integrase gene of the temperate bacteriophage *4AR29* from *Prevotella ruminicola* AR29 but little further work on this phage has been reported.

*Selenomonas ruminantium* is a ubiquitous Gram-negative ruminal species that plays a major role in the fermentation of a variety of carbohydrates, the breakdown of protein (Wallace & Brammall, 1985) and the hydrolysis of urea (Wonzny et al., 1977). Because of its relatively large and consistent population in the rumen, *S. ruminantium* is one of the target species for genetic engineering to increase digestive efficiency in ruminants. However, genetic studies in this species have progressed slowly. Some studies have described endogenous plasmids of *S. ruminantium* (Zhang et al., 1991), and a temperate bacteriophage has been described previously (Lockington et al., 1988). However, this bacteriophage did not survive long-term storage and has been lost from the culture collection. A potential plasmid shuttle vector has been developed for this species (Zhang et al., 1991), but despite many attempts, transformation with this or any other vector has not been demonstrated for any strain of *S. ruminantium*. Because plasmid-borne genes may be less stable than chromosomal genes, new genetic traits may be stabilized by inserting new genes into the host bacterium’s chromosome. This would also eliminate the problem of incompatibility with plasmids.
Integration-proficient vectors for insertion of gene sequences into a host chromosome have been constructed successfully in other bacterial systems using site-specific integrases of temperate phages (Dupont et al., 1995; Lee et al., 1991; Kuhstoss et al., 1991). It may be possible to apply that approach to Selenomonas. Before this can be achieved, however, detailed information on the integration-mediating sequence (attP) and the associated integrase gene of a Selenomonas bacteriophage is required.

In this report we describe the isolation and characterization of a temperate bacteriophage from S. ruminantium and the identification and sequencing of the attP site and integrate gene of the phage.

**METHODS**

**Bacterial strains and media.** A selective medium was used for the isolation of S. ruminantium from crude rumen samples. It comprised a modified NB medium (Nili & Brooker, 1995) plus 2% (w/v) manitol. ML12 was identified as S. ruminantium subspecies lactilytica based on its morphology, metabolic products and the restriction map of its 16S rDNA compared with other defined strains of S. ruminantium, including S23, HD4, FB32 and 2385 (unpublished). Strain ML12 was subsequently shown to be a lysogen, harbouring bacteriophage M1, and was routinely grown in brain heart infusion (BHI) broth (25 mM Tris/HCl, 10 mM EDTA, pH 6-8) and suspension in 1 ml TE buffer. Lysozyme (10 mg; Boehringer Mannheim) was added and the suspension was incubated for 30 min at 50°C before isolating phage DNA from the lysate as described in Ausubel et al. (1994).

**DNA sequencing.** The 23 kb EcoRV fragment of phage M1 present in pEH2.3 was digested with HindIII and subcloned into pUC19 for sequence analysis. Plasmid DNA was prepared for sequencing by the alkaline lysis PEG precipitation procedure described in the Applied Biosystems Taq DyeDeoxy Terminator Cycle Sequencing kit user notes. The DNA was sequenced both manually [using Sequi Therm EXCEL DNA sequencing kit and an automated sequencer (Applied Biosystems model 373A)] and automatically [using the Applied Biosystems DNA Sequencing kit and an automated sequencer (Applied Biosystems model 373A)]. Sequences that could not be obtained using the M1 universal primers were determined using custom synthesized oligonucleotides (Table 1) and the DNA Sequencing kit by Terminator Cycle Sequencing Ready Reaction (Perkin Elmer) and automatically [using the Applied Biosystems DNA Sequencing kit and an automated sequencer (Applied Biosystems model 373A)]. Sequences that could not be obtained using the M1 universal primers were determined using custom synthesized oligonucleotides (Table 1) and the DNA Sequencing kit by Terminator Cycle Sequencing Ready Reaction (Perkin Elmer). Computer analysis of nucleotide sequence was carried out by the Australian National Genomic Information Service (ANGIS) and Basic Local Alignment Search Tool (BLAST). DNA structural analysis was carried out using Gap (Brown et al., 1990).

**RESULTS**

**Characterization of bacteriophage M1**

Bacteriophage M1 was identified in S. ruminantium strain ML12 after treatment of a late-exponential-phase culture with 1 μg mitomycin C ml⁻¹. On a lawn of ML12 cells, small turbid plaques that became more turbid with time were observed.

To obtain a phage-free host, a prophage curing experiment was performed on ML12 lysogens. ML12

**Table 1. Nucleotide sequence of primers used for DNA sequencing of the phage M1 attP region**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Nucleotide sequence (5'–3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>TGGGAAATGCTGGGGCGGTA</td>
</tr>
<tr>
<td>B</td>
<td>CCAGAAGTATGGCATCCACAT</td>
</tr>
<tr>
<td>C</td>
<td>GAGCCGGCAAGGCAAAGTA</td>
</tr>
<tr>
<td>D</td>
<td>TACTACTGGCCAACGGCCTC</td>
</tr>
<tr>
<td>E</td>
<td>GAGCCGGTTCGACAGGTAG</td>
</tr>
<tr>
<td>F</td>
<td>AGGTTGAAACATGGGGAAG</td>
</tr>
<tr>
<td>G</td>
<td>CCTAATGCGTATGTGTTGTTTTC</td>
</tr>
<tr>
<td>H</td>
<td>ATGCCGCTTGATGTCGCCCT</td>
</tr>
<tr>
<td>I</td>
<td>CTCCTGCGTTTCCCT</td>
</tr>
<tr>
<td>J</td>
<td>ACTTGCTTTTTCCGTTCCT</td>
</tr>
</tbody>
</table>

above method except the cells were washed in 5 ml TE buffer after lysozyme treatment to remove free phage particles. Cloning of phage fragments in pUC19 was carried out as described in Ausubel et al. (1994).

**PCR.** PCR reaction mixtures contained 1 mM dNTP mixture, 2.5 μl PCR buffer (10 x), 1 μl AmpliTaq (1 unit per 100 μl), 25 pmol each primer and 25 ng template DNA in a final volume of 25 μl. Primers from the 1.3 kb cloned phage DNA fragment were 5'CGCAATCACGGGGAGAAC3' and 5'CAGTCCAATCAGGGGGAAC3'. DNA samples were amplified using a Perkin-Elmer Cetus Thermal Cycler at 94°C for 1 min, 55°C for 2 min and 72°C for 3 min for 35 cycles. A final step of 10 min at 72°C was included to ensure the complete extension of primers.

**Electron microscopy.** Bacteriophage-infected cells were fixed in 4% (v/v) paraformaldehyde, 1.25% (v/v) gluteraldehyde in PBS buffer [145 mM NaCl, 9 mM Na2HPO4, 1.3 mM NaH2PO4, 1.5% (w/v) sucrose]. The cells were centrifuged and washed gently twice (30 min each) in PBS buffer and the pellets were suspended in 2% (w/v) osmium tetroxide solution in PBS buffer for 2 h followed by dehydration through a series of ethanol washes. Stepwise infiltration was carried out with 50% (w/v) Spurr's resin in ethanol for 8 h and three changes with 100% Spurr's resin (8 h each). The cells were then cured in fresh 100% Spurr's resin at 60°C for 24 h. Embedded preparations were sectioned with a diamond knife in ultrathin sections (50 nm) using a Reichert Ultracut E microtome and mounted on collodion-coated 400-mesh copper grids. The preparations were examined using a Phillips CM100 electron microscope at an accelerating voltage of 80 kV, and images were recorded on Kodak no. 4489 electron microscope film.
Bacteriophage M1 from *Selenomonas ruminantium*

**Fig. 1.** Transmission electron micrograph of phage M1. (a) M1 particle on the cell surface of ML12; bar, 50 nm. (b) M1 particles within ML12; bar, 200 nm. Arrows show groups of phage heads.

lysogens were exposed to various treatments that reduced the c.f.u. by at least 75%, including UV irradiation, incubating the cells at 25 °C or incubating the cells in BHI medium containing 1 μg mitomycin C ml⁻¹ at 39 °C. No cured derivatives were obtained. Since low temperatures often induce lysis of cells harbouring lysogenic bacteriophages (Lubitz et al., 1984), we examined the effects of temperature on a mid-exponential-phase culture of the ML12 lysogen. At 39 and 42 °C, after 15 h incubation, a very low amount of lysis occurred. When the incubation temperature was lowered to 37 °C, lysis was visible but optimal lysis occurred at an incubation temperature of 25 °C. At 15 °C and below, lysis was reduced, but this was probably a consequence of reduced cell metabolism at lower temperatures. By transmission electron microscopy, phage M1 was described as a small lambda-like particle with a 30 nm diameter icosahedral head and a tail of approximately 25 nm (Fig. 1).

Twenty *S. ruminantium* isolates from existing laboratory stocks and 50 isolates from crude rumen fluid were screened for phage M1 infection. No plaques were apparent on bacterial lawns of strains tested. Southern blot analysis of DNA isolated from these strains after incubation with M1 showed no homology to phage M1 DNA.

To detect bacteriophage M1 in the rumen, ten sheep (five on normal shed ration, five on a lucerne diet) were examined. A 1.3 kb M1 DNA probe (pUCM1.3) was obtained from a *ClaI* library of M1 DNA inserted in pUC19 and cloned in *Escherichia coli*. The probe was sequenced and forward and reverse 20-mer PCR primers were synthesized. Total DNA from crude rumen samples from the ten sheep was tested for the presence of M1 DNA by PCR amplification of the 1.3 kb fragment, agarose gel electrophoresis, Southern blotting and hybridization to pUCM1.3. No M1 DNA signal was detected. Bacteria from the crude rumen samples were therefore selectively cultured *in vitro* in mannitol-containing NB medium. When total chromosomal DNA from the selected bacterial population was used (in duplicate) as a PCR template, a 1.3 kb fragment that hybridized to pUCM1.3 was detected in one out of the five rumen samples collected from sheep fed a normal shed ration, and from two out of the five rumen samples collected from sheep fed a lucerne diet (Fig. 2). To eliminate the possibility that unknown factors extracted from rumen fluid may interfere with the PCR reaction, an experiment was carried out where separate PCR reactions containing 50 or 100 ng rumen fluid DNA were spiked with control DNA. Addition of rumen fluid DNA had no effect on the PCR product from the control DNA (result not shown). In addition, there was no difference in amplification of M1 DNA from lysogens grown in BHI ±20% rumen fluid.

**Bacteriophage DNA analysis**

Total bacteriophage DNA from ML12 was isolated and mapped. Restriction enzymes EcoRI and *ClaI* cleaved the M1 genome into 16 and 10 restriction fragments, respectively; *NotI* and *SacI* cleaved the phage DNA twice; *BstEII* and *HpaI* each cleaved the DNA once. *HindIII* and *BamHI* cleaved the phage DNA too frequently to be used for restriction mapping. No restriction sites were found for *PstI*, *BglII*, *DraI* and *XhoI*. The positioning of some sites was confirmed by probing Southern blots of restriction digests with radiolabelled phage M1 fragments isolated from agarose
Fig. 3. Restriction map of bacteriophage M1 DNA, with an enlargement of the int and attP regions. E, EcoRl; C, ClaI; S, Scal; N, NotI; B, BstII; H, HindIII; EV, EcoRV. Numbers 77, 1052 and 1285 refer to the start and end of ORF1 (int) and attP.

Fig. 4. Identification of cohesive ends in bacteriophage M1. (a) ClaI-digested bacteriophage M1 DNA unheated (lane 1) or heated at 70 °C for 5 min (lane 2). (b) ClaI-digested, heat-treated M1. Lane 3, gel profile; lane 4, Southern blot probed with 32P-labelled 5 kb ClaI fragment [arrowed in (a)]. DNA size markers (kb) are on the right of each panel.

Fig. 5. Integration of bacteriophage M1 DNA into the S. ruminantium genome. Total DNA from S. ruminantium ML12 and from purified bacteriophage was restricted with EcoRl and EcoRl/HindIII and separated by agarose gel electrophoresis. (a) Lanes: 1, EcoRl-restricted phage M1 DNA; 2, infected S. ruminantium ML12 DNA; 3 and 4, M1 and ML12 DNA, respectively, after Southern transfer and probing with [32P]dCTP-labelled 27.4 kb NotI fragment of M1. (b) Lanes: 1, EcoRl/HindIII-restricted phage M1 DNA; 2, EcoRl/HindIII-restricted infected S. ruminantium ML12 DNA; 3 and 4, M1 and ML12 DNA, respectively, after Southern transfer and probing with [32P]dCTP-labelled 2.3 kb EcoRl–EcoRV fragment of M1. (c) Lanes: 1, EcoRl/HindIII-restricted phage M1 DNA; 2, EcoRl/HindIII-restricted infected S. ruminantium ML12 DNA; 3 and 4, M1 and ML12 DNA, respectively, after Southern transfer and probing with [32P]dCTP-labelled 1 kb EcoRl–HindII fragment of M1. Numbers at the right of each panel represent the size of DNA fragments (kb).

gels. The phage genome was estimated at 48 kb and the restriction map is shown in Fig. 3.

The bacteriophage genome was shown to contain cohesive ends by comparing restriction patterns of phage DNA before and after heat denaturation at 70 °C for 5 min. The staining intensity of the ClaI 5.0 kb band was greater in the unheated sample (Fig. 4a, lane 1) than in the heated sample (Fig. 4a, lane 2). To locate the phage ends more precisely, a Southern blot of ClaI-digested M1 DNA was carried out. With the ClaI 5.0 kb fragment as probe, hybridization to the ClaI 4.0 and 1.0 kb bands (Fig. 4b) was demonstrated. These results localized the cohesive ends as shown in Fig. 3.

Localization of the integration region of phage M1

The integration site of the phage is predicted to be separated into two locations upon integration. Therefore, we examined Southern blots of phage and lysogen DNA for evidence of a DNA band shift consistent with an integration event. When the 27.4 kb NotI fragment of phage M1 was used as probe, a 7.5 kb EcoRI fragment present in phage M1 DNA was absent from EcoRI-digested ML12 lysogen DNA (Fig. 5a). EcoRV restricts the 7.5 kb EcoRI fragment at two sites, producing three fragments of 1.3, 2.3 and 3.9 kb. These DNA fragments were used separately as probes in Southern blots of EcoRI/HindIII doubly digested lysogen and phage M1 DNA. A band shift was observed only when the 2.3 kb fragment was used as a probe (Fig. 5b), localizing the
Bacteriophage M1 from *Selenomonas ruminantium*

(a) Mapping of the plasmid insert. The encoded amino acid sequence is shown above the nucleotide sequence at ORFs 1 and 2, starting with the first methionine of each ORF. The encoded amino acid sequence is marked by a double underscoring and the putative SD sequences are boxed.

(b) Location of the three pairs of imperfect inverted repeats separated by a 12 bp palindrome.

Fig. 6. Nucleotide sequence of the attP region of phage M1. (a) Restriction sites are indicated by a line under the sequence. The encoded amino acid sequence is shown above the nucleotide sequence at ORFs 1 and 2, starting with the first methionine of each ORF. The attP core sequence is marked by a double underscoring and the putative SD sequences are boxed. (b) Location of the three pairs of imperfect inverted repeats separated by a 12 bp palindrome.

**attP** site to that fragment. The 2.3 kb EcoRI-EcoRV fragment was therefore ligated with linearized pBR322 and cloned in *E. coli*. Mapping of the plasmid insert suggested that the integration site is located within 300 bp of an internal HindIII site (Fig. 3). HindIII restricted the 2.3 kb cloned fragment into 1.0 and 1.3 kb fragments. These DNA fragments were used separately as probes in Southern blots of EcoRI/HindIII doubly

**Table 1**

<table>
<thead>
<tr>
<th>Lanes</th>
<th>Nucleotide Sequence</th>
<th>ORF</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ATTATGCTGATGTTCTGGTC</td>
<td>ORF1</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>ACCCATGATCGATGTTCTGGTC</td>
<td>ORF2</td>
<td>?</td>
</tr>
</tbody>
</table>

**Figure 6**

- **attP** region of phage M1.
- **Restriction sites** indicated by a line under the sequence.
- The encoded amino acid sequence is shown above the nucleotide sequence at ORFs 1 and 2, starting with the first methionine of each ORF.
- The attP core sequence is marked by a double underscoring and the putative SD sequences are boxed.

**Figure 6 (b)**

- Location of the three pairs of imperfect inverted repeats separated by a 12 bp palindrome.

**Figure 6 (c)**

- Mapping of the plasmid insert suggested that the integration site is located within 300 bp of an internal HindIII site (Fig. 3). HindIII restricted the 2.3 kb cloned fragment into 1.0 and 1.3 kb fragments. These DNA fragments were used separately as probes in Southern blots of EcoRI/HindIII doubly

---

**Table 1**

<table>
<thead>
<tr>
<th>Lanes</th>
<th>Nucleotide Sequence</th>
<th>ORF</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ATTATGCTGATGTTCTGGTC</td>
<td>ORF1</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>ACCCATGATCGATGTTCTGGTC</td>
<td>ORF2</td>
<td>?</td>
</tr>
</tbody>
</table>
digested lysogen and phage M1 DNA. A band shift was observed only when the 1 kb fragment was used as a probe (Fig. 5c), localizing the attP site to that fragment.

**Nucleotide sequence of the M1 attP region**

In most bacteriophages, site-specific recombination functions (attP, int and xis) are tightly clustered on a relatively small stretch of DNA (Tojo et al., 1996). Therefore, we sequenced the cloned 2.3 kb EcoRI-EcoRV fragment which includes the predicted attP site (Fig. 6a). This sequence contains two large ORFs. ORF1 extends from bp 77 to 1052 and potentially encodes 323 amino acids. This was preceded by a possible Shine-Dalgarno (SD) sequence (GTGGGAG) located at position 68. ORF2 extends from bp 1517 to 2360, potentially encoding a protein of 281 amino acids, and is preceded by an SD sequence (GTAGGA) at position 1482. Eight smaller ORFs surrounded the attP site, and four of these were preceded by an SD sequence.

Nucleotide and protein sequences of the two main ORFs were examined for homology to sequences recorded in GenBank, EMBL, DDB and PDB databases and GenBank Cds translations, PDB, SWISS-PROT, Spudate and PIR databases, respectively, using BLAST search. The only detected homology between the M1 2.3 kb sequence and other published DNA or protein sequences was in ORF1, which was found to contain 23 bp which were 83% homologous with the integrase gene of Myxococcus xanthus phage Mx8. Argos et al. (1986) reported that there were three perfectly conserved amino acid residues, His-396, Arg-399 and Tyr-433, in the integrase gene family. Fig. 7 shows the corresponding segment of the putative phage M1 int.

Residues His-232, Arg-235 and Tyr-268 correspond with the highly conserved residues, and in three other positions, amino acids matched those reported in up to eight other integrase genes (Brown et al., 1990). These matches were found in the only region of the protein common across the integrase gene family. Downstream of, but overlapping with, the 3' end of the proposed integrase gene, between bases 1000 and 1280, lies the region mapped as containing the attP site. Structural analysis of this region using Dotplot revealed three different imperfect inverted repeats of 14, 6 and 18 bp, positioned equidistant from a 12 bp palindrome (Fig. 6b). No other region of the 2.3 kb sequence showed a similar concentration of inverted repeats surrounding a putative core sequence.

**DISCUSSION**

Bacteriophage M1 from S. ruminantiurn strain ML12 was classified as belonging to Bradley's group A and is consistent with other phages of S. ruminantiurn (as described by Hazelwood et al., 1983). It possesses cohesive ends and a linear map of 48 kb. The presence of the phage was detected after PCR amplification of enriched cultures of mannitol-fermenting bacteria, but not in crude rumen samples. This suggests that the natural population of the phage may be low, perhaps having a limited host range in only a few strains of S. ruminantiurn. A separate control experiment showed that this result was not an underestimation caused by interference in the PCR reaction by inhibitory factors potentially present in the rumen fluid DNA preparations. The attP region of phage M1 was located by band shift analysis and was mapped to between nucleotide positions 1024 and 1243, with the core sequence predicted to be located between positions 1068 and 1204. However, the site showed no homology to any published DNA or amino acid sequence. Inverted repeat sequences are commonly found near the attP site of several temperate phages (Gregg et al., 1994). In phage λ DNA, sites for attachment of integrase are present as short sequence repeats around attP, inverted about the point where strand exchange occurs in the integration process with additional sites in close proximity on both sides of the inverted repeat. Phage M1 appears to differ slightly from this structure in that three sets of inverted repeats surround a 12 bp palindrome that may function as a core sequence for integration. Since no other regions of the 2.3 kb sequence reveal similar structural motifs, we speculate that the above organization functions as the attP region for bacteriophage M1.

Sequence analysis of the ORFs near the integration site revealed very little homology with other published sequences. This was not unexpected since recombinase genes of even closely related phages show very little direct sequence homology (reviewed by Argos et al., 1986). However, the integrase protein of many temperate phages contains three perfectly conserved amino acid residues, His-396, Arg-399 and Tyr-433 (Gregg et al., 1994; Brown et al., 1990; Boyce et al., 1995; Dupont...
et al., 1995; Hauser & Scocca, 1992). Since the predicted amino acid sequence translated from ORF1 of phage M1 also contained these three residues, and at the same relative spacing to that described above, and three other amino acids were conserved in M1 and at least eight other temperate phages, it is therefore highly likely that ORF1 encodes the integrase protein. Other proteins involved in integration, e.g. xis, could be encoded by one of the several nearby ORFs. ORF2 (843 bp), for example, is preceded by a possible SD sequence and is located in a region downstream of the attP site. However, no homology with other xis genes could be found when the DNA or amino acid sequence of the smaller ORFs was compared with published sequences of other bacteriophages. Since the attP site of most temperate phages is usually located directly adjacent to both the int and xis genes (Leong et al., 1986), it is tempting to speculate that ORF1 encodes int, and one of the smaller ORFs encodes xis, with the attP site located nearby. This could be verified by cloning and sequencing the attR, attL and attB regions of the lysogen chromosomal DNA. Comparison of the sequences at attP, attR, attL and attB regions should reveal a common core sequence which could vary in size from 9 bp (Lactococcus lactis phage BK5; Boyce et al., 1995) to 46 bp (E. coli phage P22; Leong et al., 1986). Unfortunately, attempts to carry out these experiments have been inhibited by difficulties in establishing representative genomic libraries of S. ruminantium DNA. Despite many different approaches, the generation of a full genomic library from this species has proved elusive.

Future work with bacteriophage M1 will now focus on insertion of the 2:3 kb sequence containing attP and int into a plasmid vector designed for transformation of S. ruminantium. Recombination into the bacterial chromosome would confirm the data presented here and provide a means of achieving direct chromosomal insertion of foreign genes.

ACKNOWLEDGEMENTS

This project was supported by a grant under the Australian Research Council (small grant) scheme.

REFERENCES


Received 19 December 1997; revised 16 April 1998; accepted 12 May 1998.