Sequence Analysis and Characterization of pOM1, a Small Cryptic Plasmid from *Butyrivibrio fibrisolvens*, and Its Use in Construction of a New Family of Cloning Vectors for *Butyrivibrios*†

MARY ALICE HEFFORD, 1* YASUO KOBAYASHI, 2 SHARON E. ALLARD, 2 ROBERT J. FORSTER, 1 AND RONALD M. TEATHER 1

Centre for Food and Animal Research, Agriculture and Agri-food Canada, Ottawa, Ontario, Canada K1A 0C6, 1 and Faculty of Bioresources, Mie University, Tsu 514, Japan 2

Received 24 December 1996/Accepted 23 January 1997

As a preliminary step in the development of vector systems, we have isolated and begun to characterize small, cryptic plasmids from several strains of the rumen bacterium *Butyrivibrio fibrisolvens*. We present here the complete nucleotide sequence of *Butyrivibrio* plasmid pOM1, which was isolated from *B. fibrisolvens* Bu49. While it is very similar in size to the previously characterized *Butyrivibrio* plasmids pRJF1 and pRF2, pOM1 exhibits a restriction pattern which is quite distinct. Analysis of sequence data reveals that pOM1 contains only two open reading frames of significant length (ORF1 and ORF2), both of which are required for self-replication and maintenance. The protein encoded in ORF1 shows homologies with Pre (plasmid recombination enzyme) proteins encoded in plasmids from gram-positive organisms such as *Staphylococcus aureus*, *Streptococcus agalactiae*, *Lactobacillus plantarum*, and *Bacillus thuringiensis*. The putative translation product of ORF2, on the other hand, resembles Rep (replication) proteins of a different group of gram-positive plasmids, for which the *Staphylococcus* plasmid pSN2 is a prototype. Unlike the other characterized *Butyrivibrio* plasmids, pOM1 appears to replicate via a rolling-circle mechanism. Experimental evidence showing the presence of a single-stranded replication intermediate consistent with this mechanism is presented. pOM1 has been used in the construction of a new *Escherichia coli*-*B. fibrisolvens* shuttle vector, pSMerm1, which has been successfully used to introduce a cloned gene into *B. fibrisolvens* harboring the pRJF1 plasmid.

Much of the molecular biological work on rumen bacteria to date has focused on the isolation of genes of potential industrial importance (e.g., those involved in fiber degradation). Genetic engineering of rumen bacteria themselves has lagged behind engineering of common bacteria from other sources, largely because of the lack of effective vector and DNA transfer systems. With the identification and characterization of potential plasmid vectors from at least three species of rumen bacteria (8, 17, 26, 42, 49, 50) and the development of electroporation methods to introduce DNA into essentially any bacterial cell, this situation is beginning to change. Our efforts have focused on developing vector and transformation systems for *Butyrivibrio fibrisolvens*, a rumen bacterial species which, by classical methods at least, has been found at fairly high population densities in many ruminants (21).

A survey of *B. fibrisolvens* isolates from domestic dairy cattle and wild white-tailed deer showed that many strains do indeed harbor small plasmids that are potential cloning vectors (12a). We have recently sequenced two such plasmids, pRJF1 (17) and pRF2 (26). These appear to encode only proteins required for their own replication and maintenance. Both plasmids contain the same replication origin and rep (replication protein) gene and appear to replicate via a theta mechanism. Both have been successfully used to introduce an antibiotic resistance marker into *B. fibrisolvens* OB156 (3, 26).

pOM1, a plasmid from *B. fibrisolvens* Bu49, was isolated and cloned in *Escherichia coli* by Mann et al. (32). While it is very similar in size to pRJF1 and pRF2 (approximately 2.8 kb), pOM1 exhibits a restriction pattern which is quite distinct. In addition, probes for functional regions of pRJF1 do not hybridize with pOM1 plasmid DNA. It was therefore possible that pOM1 could form the basis of a second, compatible plasmid vector family in *B. fibrisolvens*. To help define the nature of pOM1 replication and facilitate its further development as a cloning vector, the sequence of pOM1 was determined. Because the original *E. coli* clones containing the complete sequence of pOM1 in pBR322 appeared to be unstable and subject to rearrangement and/or deletions, we report here an independent cloning and sequencing of pOM1 plasmid DNA, a study on its replication, and our preliminary work to construct a shuttle vector based on this plasmid.

**MATERIALS AND METHODS**

Bacterial strains and plasmids. pOM1 was originally isolated from *B. fibrisolvens* Bu49 and characterized by Mann et al. (32). *B. fibrisolvens* Bu49 that was cured of plasmid (Bu49c) was kindly provided by Kari Gobius (Commonwealth Scientific and Industrial Research Organisation, St. Lucia, Australia). *E. coli* TBI (F- araD139 araE239 proA+ proC+ metB supE44 thi-1 leu-2 mtl-1 lacIq Z15898 [F' proAB lacIq Z15898]) and DH5αF- lacIq Z15898 (F- lacIq Z15898 [F' proAB lacIq Z15898]) were used. Plasmid vectors pUC18 and pUC19 (42) were used.

Plasmid isolation, cloning, and DNA sequencing. pOM1 was isolated from *B. fibrisolvens* Bu49 by a modified alkaline lysis procedure (17) in which the crude plasmid preparation was treated with cetyltrimethylammonium bromide (CTAB) to remove polysaccharides (12). Plasmid preparations of pOM1 were further purified by electrophoresis on a 1.0% agarose gel with a Tris-borate-EDTA running buffer (31). The band containing the plasmid DNA was excised from the gel and purified with a 0.22-μm-mesh-diameter spin column (Costar, Cambridge, Mass.) (46).
FIG. 1. Complete nucleotide sequence of pOM1. ORFs enclosing the putative Pre and Rep proteins are indicated. Putative SD, −10, and −35 sequences, and the regions showing homology to the sequences within the replication region of pSN2 and the RSA site, are indicated by thick lines under the sequence. The AT-rich region (nt 2524 to 2800) is indicated with a thin underline. Inverted repeat sequences are indicated by arrows under the sequence and designated by letters.

The sequence data were assembled and analyzed with the Wisconsin Sequence Analysis Package version 7.0 software (Genetics Computer Group) running on a VAX computer.

Oligonucleotide synthesis. Oligonucleotides for sequencing, primer extension, PCR amplification, and probe labeling were synthesized on a Pharmacia Gene Assembler (Pharmacia Biotech) but not further purified. The sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence.

A nested deletion library was made for pETB2 (18). The deleted mutants and intermediates was performed essentially as described by te Riele et al. (41), with minor modifications. Plasmid replication was tested both in E. coli host strains TB1 and E. coli host strains TB2 to yield plasmids pETB1 and pETB2. DNA was digested, dephosphorylated pUC19. Of the resulting clones, one, designated pETB2, was selected, and the deletions were sequenced.

Crude pOM1 DNA (without CTAB treatment) was also restricted with E. coli EcoRI and XmnI and PCR generation of probes were synthesized on a Pharmacia Gene Assembler (Pharmacia Biotech). Crude oligonucleotide preparations were desalted with a NAPS-10 column (Pharmacia Biotech) but not further purified. The sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence.

Sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence. The AT-rich region (nt 2524 to 2800) is indicated with a thin underline. Inverted repeat sequences are indicated by arrows under the sequence and designated by letters.

The The sequence data were assembled and analyzed with the Wisconsin Sequence Analysis Package version 7.0 software (Genetics Computer Group) running on a VAX computer.

Oligonucleotide synthesis. Oligonucleotides for sequencing, primer extension, and PCR generation of probes were synthesized on a Pharmacia Gene Assembler Special at a 0.2 μM scale with chemicals and the procedure supplied by Pharmacia Biotech. Crude oligonucleotide preparations were desalted with a NAP-10 column (Pharmacia Biotech) but not further purified. The sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence.

A nested deletion library was made for pETB2 (18). The deleted mutants and intermediates was performed essentially as described by te Riele et al. (41), with minor modifications. Plasmid replication was tested both in E. coli host strains TB1 and E. coli host strains TB2 to yield plasmids pETB1 and pETB2. DNA was digested, dephosphorylated pUC19. Of the resulting clones, one, designated pETB2, was selected, and the deletions were sequenced.

Crude pOM1 DNA (without CTAB treatment) was also restricted with E. coli EcoRI and XmnI and PCR generation of probes were synthesized on a Pharmacia Gene Assembler (Pharmacia Biotech). Crude oligonucleotide preparations were desalted with a NAPS-10 column (Pharmacia Biotech) but not further purified. The sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence.

Sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence. The AT-rich region (nt 2524 to 2800) is indicated with a thin underline. Inverted repeat sequences are indicated by arrows under the sequence and designated by letters.

The sequence data were assembled and analyzed with the Wisconsin Sequence Analysis Package version 7.0 software (Genetics Computer Group) running on a VAX computer.

Oligonucleotide synthesis. Oligonucleotides for sequencing, primer extension, and PCR generation of probes were synthesized on a Pharmacia Gene Assembler Special at a 0.2 μM scale with chemicals and the procedure supplied by Pharmacia Biotech. Crude oligonucleotide preparations were desalted with a NAP-10 column (Pharmacia Biotech) but not further purified. The sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence.

A nested deletion library was made for pETB2 (18). The deleted mutants and intermediates was performed essentially as described by te Riele et al. (41), with minor modifications. Plasmid replication was tested both in E. coli host strains TB1 and E. coli host strains TB2 to yield plasmids pETB1 and pETB2. DNA was digested, dephosphorylated pUC19. Of the resulting clones, one, designated pETB2, was selected, and the deletions were sequenced.

Crude pOM1 DNA (without CTAB treatment) was also restricted with E. coli EcoRI and XmnI and PCR generation of probes were synthesized on a Pharmacia Gene Assembler (Pharmacia Biotech). Crude oligonucleotide preparations were desalted with a NAPS-10 column (Pharmacia Biotech) but not further purified. The sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence.

Sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence. The AT-rich region (nt 2524 to 2800) is indicated with a thin underline. Inverted repeat sequences are indicated by arrows under the sequence and designated by letters.

The sequence data were assembled and analyzed with the Wisconsin Sequence Analysis Package version 7.0 software (Genetics Computer Group) running on a VAX computer.

Oligonucleotide synthesis. Oligonucleotides for sequencing, primer extension, and PCR generation of probes were synthesized on a Pharmacia Gene Assembler Special at a 0.2 μM scale with chemicals and the procedure supplied by Pharmacia Biotech. Crude oligonucleotide preparations were desalted with a NAP-10 column (Pharmacia Biotech) but not further purified. The sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence. The AT-rich region (nt 2524 to 2800) is indicated with a thin underline. Inverted repeat sequences are indicated by arrows under the sequence and designated by letters.

The sequence data were assembled and analyzed with the Wisconsin Sequence Analysis Package version 7.0 software (Genetics Computer Group) running on a VAX computer.

Oligonucleotide synthesis. Oligonucleotides for sequencing, primer extension, and PCR generation of probes were synthesized on a Pharmacia Gene Assembler Special at a 0.2 μM scale with chemicals and the procedure supplied by Pharmacia Biotech. Crude oligonucleotide preparations were desalted with a NAP-10 column (Pharmacia Biotech) but not further purified. The sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence. The AT-rich region (nt 2524 to 2800) is indicated with a thin underline. Inverted repeat sequences are indicated by arrows under the sequence and designated by letters.

The sequence data were assembled and analyzed with the Wisconsin Sequence Analysis Package version 7.0 software (Genetics Computer Group) running on a VAX computer.

Oligonucleotide synthesis. Oligonucleotides for sequencing, primer extension, and PCR generation of probes were synthesized on a Pharmacia Gene Assembler Special at a 0.2 μM scale with chemicals and the procedure supplied by Pharmacia Biotech. Crude oligonucleotide preparations were desalted with a NAP-10 column (Pharmacia Biotech) but not further purified. The sequences in gaps and on the second strand were obtained by using custom oligonucleotide primers as necessary to determine the complete nucleotide sequence. The AT-rich region (nt 2524 to 2800) is indicated with a thin underline. Inverted repeat sequences are indicated by arrows under the sequence and designated by letters.
use. Hybridizations and washings were with standard conditions and protocols. Approximately equal total radioactivities and specific activities were used for each hybridization.

**Southwestern analysis.** Total cellular proteins from *B. fibrisolvens* Bu49 containing pOM1 and from *E. coli* clones containing PETB2, pXMN, or pUC19 plasmid DNA were electroblotted onto a nylon membrane after sodium dodecyl sulfate-polyacrylamide gel electrophoresis and probed with the same 32P-labelled double-stranded DNA (dsDNA) Pre or Rep probes used for Northern analysis. Because DNA-binding proteins usually recognize dsDNA, the probes were not denatured prior to Southwestern analysis (47).

**Vector construction and transformation.** *E. coli*- *B. fibrisolvens* shuttle vectors, pSMerm1 and pSMerm4, were constructed from pETB2 by restriction with *Sst*I fragment or the blunted 2-kb *Kpn*I fragment of the *Enterococcus faecalis* plasmid pAM121, respectively, and used to transform *E. coli* TB1 cells by standard methods (31). (The larger pAM121 fragment contains both the ermC and ermM genes, while the smaller fragment contains the ermC gene but not the ermM gene.) pSMerm1 and pSMerm4 plasmid DNA were then isolated from *E. coli* hosts, and transformation by electroporation was attempted with pSMerm1 containing part of the leading-strand replication origin (oriL) recognized by its cognate Rep protein. In pOM1, a portion of this sequence (nt 2651 to 2686) is complementary to sequence thought to contain the promoter for the rep gene (nt 2614 to 2649) and can form an extended stem-loop structure (shown in detail in Fig. 5A).

**Plasmid segregation.** A recombinant *B. fibrisolvens* Bu49 transformant containing pSMerm1 was diluted 1:100 into LB medium without erythromycin and grown aerobically for 18 h. The culture was then diluted 1:100 in the same nonselective medium and grown for 8 h longer (i.e., approximately 12 generations in total) before serial dilutions were plated onto nonselective solid medium. After 24 h of incubation under standard conditions, 200 individual colonies from these plates were plated onto replica plates. The first (selective) plate contained 10 μg of erythromycin per ml, and the second (nonselective) plate contained no antibiotic.

**Nucleotide sequence accession number.** The sequence reported here has been submitted to GenBank (accession no. L31579).

**RESULTS**

**DNA sequence analysis.** The complete DNA sequence of pOM1 is shown in Fig. 1. Preliminary sequence analysis revealed the following features: (i) There is a highly AT-rich region (nucleotides [nt] 2524 to 2800) which contains several short direct and inverted repeat regions. (ii) There is an open reading frame (ORF1, nt 449 to 1711), which is thought to encode a plasmid recombination protein (Pre), and the recombination site (RSX, nt 266 to 287) recognized by this Pre protein.

(Homologies for the Pre sequence are shown in Fig. 2; those for the RSX site are shown in Fig. 3.) ORF1 is initiated at a methionine and preceded by typical Shine-Dalgarno (SD) and promoter (−10 and −35) sequences. (iii) A second open reading frame (ORF2, nt 2535 to 1912) encodes a protein highly homologous to the replication proteins (Rep) of the pSN2 family, which are involved in rolling-circle replication (RCR) of these plasmids. (Rep sequence homologies are shown in Fig. 4.) ORF2 is preceded by typical SD, −10, and −35 sequences but appears to be initiated at a leucine (TGG) codon. Because the homology of this putative Rep with other proteins in the vicinity of the TGG codon is high, there is some possibility of a sequencing error. This region was therefore resequenced with the pOM2347 primer used for primer extension. These data confirm the TGG and the in-frame TAA stop codons immediately upstream of the proposed start. In addition, the primer extension analysis indicates a transcriptional start site at nt 2623. (iv) A region of sequence 5′ to the putative rep gene of pOM1 (nt 2662 to 2759) strongly resembles (67% similarity) the sequence immediately 5′ of the rep gene in pSN2. In pSN2 this sequence was identified as containing part of the leading-strand replication origin (oriL) recognized by its cognate Rep protein. In pOM1, a portion of this sequence (nt 2651 to 2686) is complementary to sequence thought to contain the promoter for the rep gene (nt 2614 to 2649) and can form an extended stem-loop structure (shown in detail in Fig. 5A).

**Southern analysis.** Plasmid preparations from *B. fibrisolvens* Bu49 or *E. coli* TB1 containing pETB2 were analyzed for the presence of single-stranded DNA (ssDNA) replication intermediates by treatment with S1 nuclease and Southern blot analysis of electrophoretically separated plasmid forms with and without denaturation prior to transfer (Fig. 6). Ethidium bromide (EtBr) visualization reveals the expected, abundant plasmid bands for pOM1 and pETB2 without S1 nuclease treatment (Fig. 6A, lanes 1 and 3, respectively). On treatment with S1 nuclease, a new band migrating at the position expected for linearized dsDNA is seen (Fig. 6A, lanes 2 and 4). This dsDNA results from the known S1 cleavage of circular plasmid nicked during the isolation procedure and varies in intensity from preparation to preparation (data not shown). (In our experience, covalently closed circular pOM1 plasmid migrates anomalously when electrophoresed on a 1% agarose gel and actually runs more slowly than linearized double-stranded pOM1 DNA [Fig. 6A, lanes 1 and 2].) A gel similar to that shown in Fig. 6A was transferred to a nylon membrane without denaturation and probed with pOM1 DNA (Fig. 6B). Small amounts of DNA migrating with both circular and linear dsDNA are transferred (Fig. 6B). In addition, in the case of pOM1 DNA, a fast-migrating, S1-sensitive band is also seen (Fig. 6B, lane 1). This band presumably results from an ssDNA replication intermediate. In the case of pETB2, several S1-sensitive bands are seen (Fig. 6B, lanes 3 and 4). None of these, however, migrate more quickly than the supercoiled, monomeric form of pETB2.

**Northern analysis.** Total RNA was isolated from *B. fibrisolvens* Bu49 and the two *E. coli* clones of pOM1 DNA, pPETB2 and pXMN, and blotted, after electrophoresis, onto nylon membranes (Fig. 7). The RNA was probed with restriction fragments representing the pre gene and its upstream sequence (Pre probe), the rep gene and its upstream sequence (Rep probe), the intergenic region (Inter probe), the AT-rich region...
immediately upstream of the rep gene (AT-rich probe), or the entire pOM1 plasmid (Whole probe). (Maps of plasmids pETB2 and pXMN indicating the Pre, Rep, Inter, AT-rich, and Whole probes are shown in Fig. 7A. Details on the sequences in these probes are in Fig. 1.) Signals from all probes except the AT-rich probe (Fig. 7C) were apparent in the lanes containing E. coli clones pETB2 and pXMN (Fig. 7B to D, lanes 1 and 2) after relatively short exposures (typically 4 to 5 h) of the X-ray film to the membrane, whereas considerably longer exposures (20 to 30 h) were needed for the detection of RNA isolated from B. fibrisolvens Bu49 (Fig. 7B to D, lanes 3).

Probing of total RNA with the Pre probe revealed a message (ca. 1.5 kb in length) in both B. fibrisolvens and the E. coli clone pXMN (Fig. 7B, lane 2) after relatively short exposures (typically 4 to 5 h) of the X-ray film to the membrane, whereas considerably longer exposures (20 to 30 h) were needed for the detection of RNA isolated from B. fibrisolvens Bu49 (Fig. 7B to D, lanes 3).

Probing of total RNA with the Pre probe revealed a message (ca. 1.5 kb in length) in both B. fibrisolvens Bu49 (Fig. 7B, lane 3) and the E. coli clone pXMN (Fig. 7B, lane 2). In addition to a band that may represent a transcript of the truncated pre gene in pETB2 (1.3 kb), several other transcripts are present in RNA from this clone (Fig. 7B, lane 1). These signals probably arise from transcription from the opposite strand as a result of the lac promoter of the pUC vector. These bands are absent in the RNA from the pXMN clone, in which the pre gene and the lac promoter are in the same orientation. The stronger signal from E. coli hosts is a reflection of the increased plasmid copy number of the pUC-based plasmids.

When the Rep probe is used for Northern analysis, a single band running just below the 1.28-kb marker is observed in RNA isolated from the native Butyrivibrio host (data not shown). This message is also observed in RNA isolated from each of the two E. coli hosts but only after extended exposure times (Fig. 7C, lanes 1 and 2). In addition, a second, slightly larger and slightly more intense band is observed in E. coli harboring the pETB2 plasmid (Fig. 7C, lane 1).

Results of Northern blot analysis with the Inter probe (Fig. 7D) or the Whole probe (data not shown) are, basically, the summation of the results from the Rep and Pre probes. The native Butyrivibrio host shows signals corresponding to transcripts expected from the rep gene seen on short exposure, signals that are very similar in size to those observed with the Pre probe are seen with longer exposure. Closer analysis of the sequence of the Rep probe reveals that, in addition to considerable sequence upstream of the rep gene, it does contain small regions with some homology to the Pre protein-coding sequence. When the small, and presumably more specific, AT-rich probe is used (Fig. 7C), the band running just below the 1.28-kb marker is once again observed in RNA isolated from the native Butyrivibrio host (lane 3). This band is also observed in RNA isolated from each of the two E. coli hosts but only after extended exposure times (Fig. 7C, lanes 1 and 2). In addition, a second, slightly larger and slightly more intense band is observed in E. coli harboring the pETB2 plasmid (Fig. 7C, lane 1).

Results of Northern blot analysis with the Inter probe (Fig. 7D) or the Whole probe (data not shown) are, basically, the summation of the results from the Rep and Pre probes. The native Butyrivibrio host shows signals corresponding to transcripts expected from the rep gene seen on short exposure, signals that are very similar in size to those observed with the Pre probe are seen with longer exposure. Closer analysis of the sequence of the Rep probe reveals that, in addition to considerable sequence upstream of the rep gene, it does contain small regions with some homology to the Pre protein-coding sequence. When the small, and presumably more specific, AT-rich probe is used (Fig. 7C), the band running just below the 1.28-kb marker is once again observed in RNA isolated from the native Butyrivibrio host (lane 3). This band is also observed in RNA isolated from each of the two E. coli hosts but only after extended exposure times (Fig. 7C, lanes 1 and 2). In addition, a second, slightly larger and slightly more intense band is observed in E. coli harboring the pETB2 plasmid (Fig. 7C, lane 1).

Results of Northern blot analysis with the Inte probe (Fig. 7D) or the Whole probe (data not shown) are, basically, the summation of the results from the Rep and Pre probes. The native Butyrivibrio host shows signals corresponding to transcripts expected from the rep gene seen on short exposure, signals that are very similar in size to those observed with the Pre probe are seen with longer exposure. Closer analysis of the sequence of the Rep probe reveals that, in addition to considerable sequence upstream of the rep gene, it does contain small regions with some homology to the Pre protein-coding sequence. When the small, and presumably more specific, AT-rich probe is used (Fig. 7C), the band running just below the 1.28-kb marker is once again observed in RNA isolated from the native Butyrivibrio host (lane 3). This band is also observed in RNA isolated from each of the two E. coli hosts but only after extended exposure times (Fig. 7C, lanes 1 and 2). In addition, a second, slightly larger and slightly more intense band is observed in E. coli harboring the pETB2 plasmid (Fig. 7C, lane 1).
cription of both the rep and the pre genes. The rep gene transcript is more abundant. Only the pre gene gives rise to a detectable transcript in the two E. coli clones on short exposure. In addition to bands attributable to rep and pre transcripts, a small RNA transcript (approximately 300 nt) is observed when B. fibrisolvens Bu49 RNA is probed with the Inter probe (Fig. 7D, lane 3). When the total plasmid was used as a probe, the 300-nt transcript was not observed in exposures in which the rep and pre transcripts were easily visible.

Southwestern analysis. Southwestern analysis with the Rep probe detects a single DNA-binding protein in the native host of pOM1, B. fibrisolvens Bu49, which, on the basis of its size (Mr ~ 52,300), is predicted to be the Rep protein (Fig. 8A, lane 5). In E. coli, a larger protein (Mr ~ 32,000), binds dsDNA sequences within this probe (Fig. 8A, lanes 2, 3, and 4). This larger protein, however, is detected in the E. coli clone containing only pUC19 DNA (Fig. 8A, lane 2) and appears to be unrelated to the expression of pOM1 genes. When the Pre probe is used either alone (Fig. 8B) or in combination with the Rep probe (data not shown), several DNA-binding proteins are detected in all E. coli clones, including the pUC19 control (lane 2). Once again, these proteins appear to be unrelated to the expression of pOM1 genes. Southwestern analysis with the Pre probe did not result in the detection of any protein bands in the native host of pOM1, B. fibrisolvens Bu49 (Fig. 8B, lane 5).

Vector construction and transformation. Shuttle vectors (pSMerm1 or pSMerm4) containing all or part (ermAM or ermA, respectively) of the erythromycin resistance marker from the E. faecalis plasmid pAMβ1 were used to transform E. coli TB1 cells. pSMerm1 was shown in Fig. 9A. pSMerm1 and pSMerm4 plasmid DNAs were then isolated from E. coli hosts and used to electroporate B. fibrisolvens Bu49, Bu49c, or OB156. Erythromycin-resistant colonies appeared on plates from the transformation into OB156 and Bu49c within 2 days of plating, while no colonies grew on the plates from the Bu49 transformation until after 6 days of incubation. The delayed appearance of the latter colonies and their small size suggested that they represent spontaneous erythromycin-resistant mutants of Bu49 rather than transformants.

The presence (or absence) of the recombinant shuttle vector pSMerm1 or pSMerm4 was confirmed by gel electrophoresis of total plasmid DNA from erythromycin-resistant clones and Southern blot analysis with a DIG-labelled probe representing the ampicillin resistance marker of pUC19. As expected, colonies from attempts at transformation into B. fibrisolvens Bu49 contained only the wild-type pOM1 plasmid DNA and no detectable amounts of the recombinant plasmid, pSMerm1 or pSMerm4 (Fig. 9B and C, lanes 2 and 3). Strain OB156, however, was successfully transformed with both pSMerm1 (Fig. 9B and C, lanes 4) and pSMerm4 (lanes 5). Preparations of total plasmid DNA from these isolates contained both pRJF1, the resident plasmid in B. fibrisolvens OB156, and the recombinant plasmid, pSMerm1 or pSMerm4. The cured B. fibrisolvens strain Bu49c was also successfully transformed as indicated by the detection of pSMerm plasmids with the pUC probe (Fig. 9C, lanes 8 and 9).

Analysis of segregational and structural stabilities. Recombinant Bu49c clones containing pSMerm1 were grown for approximately 12 generations on nonselective medium. Colonies isolated on nonselective medium were then tested for plasmid retention by plating on selective, erythromycin-containing medium. All (200 of 200) colonies tested showed antibiotic resistance, indicating retention of the plasmid in the absence of selective pressure. Plasmid DNA was prepared from 10 of these colonies. Restriction patterns obtained with multiple enzymes indicate that the structure of the plasmid isolated from each of these clones grown without selective pressure is iden-
tical to that of the pSMerm1 plasmid used for the initial transformation.

DISCUSSION

Naturally occurring plasmids have been used to form the basis of plasmid vectors in many bacterial systems. The potential of the pOM1 plasmid to form the basis of a vector system for the rumen bacterium \textit{B. fibrisolvens} Bu49 was noted when the plasmid was first isolated (32). In our own laboratories, the small, cryptic plasmids pRJF1 and pRJF2 were used to develop the vectors pBHerm and pYK4, which have been used to introduce and maintain new genetic information in butyrivibrios (3, 26). Our earlier work using restriction maps and probing with fragments representing functional regions of pRJF1 DNA indicated that pOM1 was unrelated to plasmids pRJF1 and pRJF2 and probably contained a different replication origin. We therefore decided to characterize this plasmid in the hope that it could be developed into a second, compatible vector family for bacteria of the genus \textit{Butyrivibrio}. The original clones of pOM1 DNA in \textit{E. coli} vectors pBR322 and pAT153 were graciously supplied to us by G. Hazlewood but, in our hands, appeared to undergo both deletions and rearrangements, particularly between pOM1 sequence and the pBR322 sequence encoding tetracycline resistance. We therefore proceeded to isolate and restrict pOM1 and to clone the resulting fragments into the pUC vectors, which lack the \textit{tet} gene of pBR322. Modification of the plasmid isolation procedure to include removal of polysaccharides by CTAB treatment greatly improved the quality of DNA obtained from Bu49 (similar improvements were not seen, however, with other \textit{B. fibrisolvens} strains tested) and allowed more facile restriction and cloning of the plasmid DNA. Despite some discrepancies among the original restriction maps of the

![FIG. 5. Potential stem-loop structures. (A) Sequence upstream of the putative \textit{rep} coding region (nt 2603 to 2710). The –10 and –35 sequences for \textit{rep} are indicated. Part of the sequence is 67% homologous to the sequence upstream of the \textit{rep} gene in \textit{p}SN2 (see text). All parts of the stem-loop structure are within the AT-rich region of pOM1 (nt 2524 to 2800). (B) nt 136 to 287 of pOM1, including invert repeat units D, E, F, and H. The core (RS\textsubscript{A}) recognized by the putative Pre protein is indicated.]

![FIG. 6. Evidence for ssDNA intermediates in pOM1 plasmid replication. (A) EtBr-stained agarose gel electrophoresis of preparations of pOM1 or pETB2 plasmid DNA without (lanes 1 and 3, respectively) or with (lanes 2 and 4, respectively) treatment with S1 nuclease. Lane m, \textit{HindIII} digest of lambda DNA. (B) A similar gel was transferred to a nylon membrane without prior denaturation and probed as described in the text. Bands representing circular (ccc), linearized double-stranded (ds), or single-stranded (ss) forms of pOM1 or pETB2 plasmid DNA are indicated, as is the HMW moiety in pOM1 preparations.]

1076 HEFFORD ET AL.
pBR322-based plasmids, which, in that work, were attributed to cloning anomalies (32), we are satisfied from our restriction and sequence analysis that we have cloned, sequenced, and characterized the same plasmid, namely, pOM1.

Analysis of the sequence of pOM1 (Fig. 1) showed only two ORFs of any reasonable length. Our searches of the databases with the putative translation product of ORF1 showed considerable sequence homology with plasmid recombination (Pre) proteins of plasmids from a number of gram-positive organisms. The second ORF, which is transcribed in the opposite direction, showed highest homologies with replication (Rep) proteins from the pSN2 family of plasmids isolated from *Staphylococcus* strains. These plasmids are part of a larger class of plasmids isolated from gram-positive organisms which have been shown to replicate via a rolling-circle mechanism involving an ssDNA intermediate (14). The two ORFs were thus tentatively designated rep and pre, and pOM1 was tentatively assigned to the general class of plasmids isolated from gram-positive organisms which use a rolling-circle replication (RCR) mechanism. Northern analysis indicates that both the rep and the pre genes give rise to stable transcripts in *B. fibrisolvens* Bu49 carrying the native pOM1 plasmid and that, as might be expected, the rep transcript is more abundant than the pre transcript.

Pre proteins resembling the putative Pre protein of pOM1 have been found in plasmids isolated from a diverse group of gram-positive organisms. While the amino acid sequences of the Pre proteins themselves are highly homologous in their N-terminal regions (43), the plasmids encoding these Pre proteins are not necessarily closely related. Genes encoding Pre proteins have been found in some members of most of the families of RCR plasmids but not in all members of any one family of plasmids. In each case, the presence of a pre gene is associated with a recombination site, RS_A, exhibiting a conserved core sequence (13). pOM1 contains an inverted repeat sequence (nt 266 to 287) which, except for one nucleotide, is
identical to the core sequence of RSₐ (Fig. 3). The presence and location of this RSₐ sequence lend further support to our assignment of this ORF as encoding a Pre protein which presumably acts at the RSₐ site.

RCR plasmids from gram-positive microorganisms have been classified in families on the basis of sequence homologies in the replication proteins (35). The putative Rep protein encoded by ORF2 of pOM1 exhibits some sequence homology with Rep proteins of all these families, suggesting that, unlike Butyrivibrio plasmids pRJF1 and pRJF2, pOM1 replicates by an RCR mechanism. The most significant sequence homologies were between Rep of pOM1 and proteins from the pSN2 family of plasmids (23). All members of the pSN2 family were isolated from closely related strains of Staphylococcus (35). Sequence conservation between these proteins and the Rep protein of the distantly related B. fibrisolvens Bu49, therefore, may be indicative of the amino acids required for proper folding and functioning of this family of Rep proteins.

Given the homologies to plasmids known to replicate by RCR mechanisms, we sought to obtain additional experimental evidence that pOM1 replicates in a similar fashion. Several critical steps in RCR of circular dsDNA have been identified (14). First, replication initiation in PCR plasmids is characterized by recognition and nicking of DNA at ori₁. This triggers leading-strand replication by displacement of the plus strand and polymerization of a new plus strand by 3'-OH extension from the nick. Second, termination site recognition and nicking generate a fully replicated strand and an ssDNA monomer of the displaced strand. Third, the ends of the ssDNA are ligated to form a circle, and fourth, a minus-strand origin (MO) is recognized for the conversion of ssDNA to dsDNA. Steps 1 through 3 are thought to be mediated by the plasmid Rep protein; step 4 is thought to be mediated through host cell factors. Thus, in addition to homologies to rep genes from RCR plasmids, the presence of an ori⁺ sequence, an MO sequence, and a single-stranded replication intermediate can all be taken as evidence for an RCR mechanism.

Single-stranded replication intermediates in RCR plasmids can be detected by Southern blot analysis of total plasmid DNA isolated from growing bacterial cultures (41). A replication intermediate (ssDNA) is identified on the basis of three criteria: (i) it migrates faster than the supercoiled monomeric plasmid form, (ii) it is sensitive to S1 nuclease, and (iii) it binds to the membrane without prior denaturation of the DNA. Using these criteria, we have identified only one plasmid form in pOM1 in B. fibrisolvens preparations, which appears to represent an ssDNA replication intermediate (Fig. 6). This putative ssDNA intermediate represents only a small proportion of the plasmid.
the total DNA in the preparation (data not shown) but is the major form detected when the DNA is transferred without prior denaturation (Fig. 6B, lane 1). Small amounts of other plasmid forms were detected on membranes when the separation gel was not treated with denaturant prior to the transfer. The signal arising from these other plasmid forms varied with the culture age and from preparation to preparation and was never proportional to the amount of the DNA (estimated from EtBr staining) present in these bands (data not shown). These secondary bands thus appear to result from nicked forms of dsDNA in which short, single-stranded regions are exposed either during plasmid preparation or in the displacement-polymerization step of the replication. As expected, S1 cleaves these nicked but double-stranded DNA forms to linearize the dsDNA, and a band running with linearized double-stranded plasmid becomes apparent in S1-treated samples (Fig. 6A, lanes 2 and 4). In addition, Fig. 6A, lane 1, shows a high-molecular-weight (HMW) (>220-kb) DNA moiety which is sensitive to S1 digestion. This HMW DNA moiety was not transferred or detected by our probe, and its identity is not clear.

Hybrid plasmids containing sequences from the RCR plasmid pC194 of *Staphylococcus* and the theta-replicating *E. coli* plasmid pBR322 can replicate via an ssDNA intermediate (RCR) in *E. coli* (19, 40, 41). The ori\(^+\) replication origin of pOM1 was therefore tested in *E. coli* by using the plasmid pETB2, a hybrid of *Eco*RI-restricted pOM1 and pUC19. Plasmid preparations contained nicked dsDNA with exposed regions of ssDNA, as evidenced by the transfer of some of this material to membranes in the absence of DNA denaturation and the appearance of a new band migrating with linearized plasmid dsDNA upon S1 treatment (Fig. 6A, lanes 3 and 4). No ssDNA band migrating faster than the supercoiled monomeric pETB2 plasmid is present. Data from Northern analysis indicate the presence of transcripts consistent with some expression of both the *rep* and *pre* genes in *E. coli* clones containing pOM1 DNA. In *B. fibrisolvens* *rep* transcripts are more abundant than *pre* transcripts; in *E. coli* the situation is reversed. Thus, while the pOM1 ori\(^+\) may be active in *E. coli*, it appears that much of pETB2 plasmid replication in *E. coli* occurs via the pUC19 origin.

In addition to an ssDNA replication intermediate, RCR implies the presence of both a plus-strand replication origin (ori\(^+\)) and an MO. In RCR plasmids studied to date, both the ori\(^+\) and MO are characterized by extended regions of hyphenated dyad symmetry. ori\(^+\) sequence characteristics are, as expected, closed related to characteristics of the Rep proteins that recognize them. ori\(^+\) sequences within a given family of plasmids are similar (14, 35). In pSN2, the region immediately 5' to the *rep* coding sequence has been identified as the leading-strand replication origin (ori\(^+\)) (35). pOM1 contains a highly related sequence, situated slightly further upstream from its putative *rep* gene (Fig. 1), which was identified as a possible ori\(^+\) in this plasmid. A portion of this pSN2-like sequence is complementary to the putative promoter region of the pOM1 *rep* gene and has the potential to form a single stem-loop structure (nt 2603 to 2710 [Fig. 5A]). There is, however, a second region of extended dyad symmetry immediately preceding the RS\(_A\) site (nt 138 to 259 [Fig. 5B]) in the pOM1 DNA sequence with the potential to form three consecutive stem-loop structures, which is reminiscent of triplet stem-loop structural motifs in the nick sites (ori\(^+\)) of RCR plasmids of the pTi181 family (22, 34, 48). Because the first step in RCR involves the recognition and nicking of dsDNA at the ori\(^+\) by the Rep protein, we tested the ability of dsDNA containing one or the other of these two structured regions to bind to protein from bacteria containing pOM1 DNA. Without extended exposure times, only the Rep probe detects a protein attributable to the presence of pOM1 DNA in *B. fibrisolvens* Bu49, and that protein is of a size consistent with translation of *rep* mRNA. This result suggests that the Rep probe contains the plus origin of pOM1. The fact that this protein is not detected in lysates of *E. coli* clones containing pOM1 DNA sequence is consistent with the suggestion from Northern analysis that the *rep* gene is not efficiently transcribed in *E. coli* hosts. Both the entire pSN2-like sequence and the potential single stem-loop structure are contained within the Rep probe. The second structural motif (the triple stem-loop motif) is contained within the sequence of the Pre probe. This probe did not bind the putative Rep protein and so is less likely to be involved in initiation of single-strand replication of pOM1.

Interestingly, although the Pre probe contains the potential RS\(_A\) recognition site for the Pre protein, it does not bind a protein that can be attributed to translation of the *pre* mRNA in either *E. coli* or *B. fibrisolvens* hosts. This result indicates that either the RS\(_A\) site is recognized as an ssDNA or that the Pre protein cannot renature sufficiently in the washing of the membrane after electrophoresis and transfer in sodium dodecyl sulfate.

The third requirement for RCR is the presence of an MO. Hosts containing RCR plasmids lacking an efficient MO are unable to convert ssDNA to double-stranded plasmid and tend to accumulate ssDNA replication intermediates (15). Several types of MO have been identified in RCR plasmids (4, 16, 25, 29, 35, 43). No sequences exhibiting strong homologies with any of the well-characterized MOs of RCR plasmids were identified in pOM1. MOs, however, vary considerably among RCR plasmids of the same family: pSN2 has a typical pal/A RS\(_E\)-type MO, whereas the RS\(_A\) sequence is entirely missing in the related plasmid pTCS1 (35). The recently characterized *Bacillus coagulans* plasmid pBC1 also has no extensive sequence similarity with any known MO but appears to effectively convert single-stranded to double-stranded plasmid (10). Because plasmid preparations from *B. fibrisolvens* Bu49 produce only trace amounts of ssDNA in proportion to plasmid copy number, we presume that pOM1 does contain an MO which allows effective conversion of ssDNA replication intermediates to double-stranded plasmids.

Regulation of plasmid replication often involves both the interaction of a small RNA countertranscript with the mRNA from the replication protein and the action of the Cop protein (9). The intercistronic (*cop*-*rep*) region of the RCR plasmid pE194, for example, has been shown to be transcribed bidirectionally: in one direction as part of the synthesis of the *cop*-*rep* message and in the other to give rise to a 65-nt RNA countertranscript (6, 27). The 350-nt *cop* transcript was first detected by Northern blot analysis of RNA from wild-type pE194 plasmid and copy number mutants (45) and has since been shown to encode a small (55-amino-acid) repressor protein which binds a stem-loop structure in the *cop*-*repF* promoter region (5, 27). We were therefore intrigued by the detection of a small transcript (approximately 300 nt in length) in our Northern analysis with the complete intergenic (*rep*-*pre*) region of pOM1 as a probe. Reanalysis of the pOM1 DNA sequence in light of this result did reveal a small potential ORF (putative protein product of 120 amino acids) encoded entirely within the sequence of the *Inter* probe that is not contained in the AT-rich probe. Unlike the *cop* gene of pE194, this ORF would be transcribed in the direction opposite to that for the *rep* gene and is not preceded by any recognizable SD, −10, or −35 sequence motifs. While the putative translation product of this ORF is predicted to be quite basic (24 Lys and Arg residues, 10
Glu and Asp residues), no DNA-binding protein of the appropriate molecular size (approximately 13 kDa) was detected with either the Rep or Pre probe. (Between them, the Rep and Pre probes contain all of the DNA sequence in the rep-pre intergenic region, i.e., all of the sequence in the Inter probe.) In addition, we were unsuccessful in detecting this small transcript with the Pre probe (which contains almost the entire putative coding region of this ORF) or, indeed, with the entire pOM1 DNA sequence as a probe for Northern analysis.

Based on sequence similarity and analysis of replication mode and origin, the pOM1 plasmid resident in *B. fibrisolvens* Bu49 appears to be related to the RCR plasmids commonly found in gram-positive organisms and to be unrelated to the other *B. fibrisolvens* plasmids, pRJF1 and pRJF2, which were used in the construction of the cloning vectors pHerm and pYK4 (3, 26). pOM1, therefore, appeared to be a good candidate to form the basis of a second family of vectors for *B. fibrisolvens*. The pHerm vectors created used the pOM1-pUC19 hybrid plasmid pETB2 and sequence from pAM1 carry the *B. fibrisolvens* resistance marker. Thus, the authors obtained the sequence information required for replication in *E. coli* (from pUC19) and in *B. fibrisolvens* (from pOM1), as well as genes encoding ampicillin resistance (for selection in *E. coli*) and erythromycin resistance (for selection in *B. fibrisolvens*).

Like pETB2, pSerm constructs replicate and are stably maintained in *E. coli* hosts. When we attempted to introduce pSerm plasmids into *B. fibrisolvens* Bu49, in which the parental plasmid, pOM1, is already resident, we did not obtain any erythromycin-resistant colonies carrying the recombinant plasmid. While this result is perhaps not unexpected, it is in contrast with our previous experience using pHerm or pYK4 vectors to transform a *B. fibrisolvens* strain containing the parental plasmid, pRFJ1. In the latter cases, transformants containing the recombinant plasmid and no parental plasmid were isolated at relatively high frequencies (3, 26). We therefore conclude that either *B. fibrisolvens* strain OB156 is easily cured of plasmid pRFJ1 during electroporation or the recombinant plasmids have a replicative advantage over the resident plasmid which allows them to become established. The substantially higher yields of pBHerml relative to pRFJ1 observed in plasmid preparations from OB156 suggest that the latter is the case. A similar situation does not exist in the case of pSerm vectors and their parental plasmid, pOM1.

pSerm1 plasmid DNA could be successfully used to transform *B. fibrisolvens* Bu49c, a derivative of Bu49 cured of pOM1. Our inability to transform the parental strain, Bu49, with pSerm constructs appears, therefore, to be the result of plasmid incompatibility.

pSerm vectors could also be introduced into *B. fibrisolvens* OB156, without curing of the resident plasmid pRFJ1. Erythromycin-resistant isolates selected after transformation contained both pSerm and pRFJ1 plasmid DNAs, indicating that the resident plasmid is not displaced and that the pOM1 and pRFJ1 replication origins are indeed different and compatible. (The complementary experiment, i.e., the transformation of Bu49 containing the pOM1 resident plasmid with vectors based on pRFJ1, was also successful.) Because the pre gene in pETB2 (and therefore in both pSerm constructs) is interrupted by pUC19 sequence, the ability of pSerm to be maintained in *B. fibrisolvens* also indicates that truncation of the Pre protein by removal of 42 amino acids at its C terminus does not seriously affect its ability to replicate in this host. Since the strong sequence homologies among known Pre proteins (and, therefore, inferred functional homologies) reside entirely within the N-terminal portions of these proteins, the relative unimportance of the C-terminal portion of the pOM1 Pre protein is not entirely unexpected.

There is evidence that some RCR plasmids tend to be segregated nonstably or prone to sequence deletions or rearrangements, properties that limit their utility as cloning vehicles (see references 9, 14, and 35 and references therein). Our data, however, indicate that this is not the case with pSerm vectors. Taken together, the data presented here indicate that the native *Butyrivibrio* plasmid, pOM1, can form the basis of a second, independent, and compatible family of cloning vectors, the pSerm family. Vectors from this family, used in conjunction with those of the pBHerml family, will allow the study of interactions of two cloned genes or groups of genes in butyrivibrios.

REFERENCES


Downloaded from http://aem.asm.org/ on October 29, 2017 by guest