Nucleotide sequence and structural organization of the small, broad-host-range plasmid pC1411 from Leuconostoc lactis 533

Aidan Coffey, Aidan Harrington, Karen Kearney, Charles Daly and Gerald Fitzgerald

INTRODUCTION

The genus Leuconostoc represents a diverse group of heterofermentative lactic acid bacteria of considerable industrial importance. Traditionally, they have been used with other lactic acid bacteria in various food and wine fermentations where, in many cases, a number of beneficial qualities result from the dominance of Leuconostoc species (Mathieu et al., 1993). For example, in dairy fermentations, the ability of some Leuconostoc species to produce diacetyl and carbon dioxide is important because of their contribution to flavour and eye-hole formation in Dutch-type cheeses. Like other lactic acid bacteria, many Leuconostoc species harbour one or more natural plasmids of various sizes. To date, phenotypes such as lactose utilization (David et al., 1992), citrate permease activity (David, 1992) and bacteriocin production and immunity (Hastings et al., 1991) have been plasmid-linked in this genus but the majority of plasmids have no known function.

The study of small plasmids from a variety of bacteria has recently become a focus of research interest and in several cases this work has led to the development of families of vectors designed for specific purposes. The genetic organization and mode of replication of a number of small cryptic plasmids from some lactic acid bacteria have been described recently (Bates & Gilbert, 1989; Leer et al., 1992; Xu et al., 1991), and it is now known that these replicons are broadly similar to those from other Gram-positive (Gruss & Ehrich, 1989) and Gram-negative (Yang & McFadden, 1993) bacteria which replicate via single-stranded intermediates. They are collectively referred to as rolling circle (RC) plasmids. All RC plasmids harbour a gene encoding a plasmid replication initiation-termination protein (Rep), its target site, termed the double strand origin (DSO, formerly plus origin), and a

Abbreviations: SSO, single strand origin; DSO, double strand origin; RS, recombination-specific site; RC, rolling circle.

The GenBank accession number for the nucleotide sequence reported in this paper is L25529.

The nucleotide sequence of the Leuconostoc lactis 533 cryptic plasmid pC1411 (2926 bp) was determined. Analysis revealed the presence of three open reading frames (ORFs). ORF 1 was capable of encoding a 249 kDa peptide which shared homology with the replication initiation protein (RepB) from a number of Gram-positive rolling circle plasmids. ORF 2 could encode a peptide of 66 kDa which was homologous to the RepC protein of the lactococcal plasmid pWV01. A function could not be assigned to ORF 3, which was capable of encoding a 12.1 kDa peptide. Transcription-translation analysis indicated the presence of three peptides of the predicted molecular masses. A putative double strand origin of replication (DSO) was identified which showed strong similarity with the DSO of a number of Gram-positive plasmids including pEl94 from Staphylococcus. Structural analysis identified a number of direct and indirect repeats in addition to putative recombination-specific sites (RS and RS) in the non-coding region of pC1411. The observed characteristics suggest that this plasmid replicates using the rolling circle mechanism. pC1411, which could be introduced into Leuconostoc, Lactococcus, Streptococcus, Lactobacillus and Bacillus is the first plasmid from the genus Leuconostoc to be characterized in such detail.

Keywords: Leuconostoc lactis, cryptic plasmid

Food Microbiology Department and National Food Biotechnology Centre, University College, Cork, Ireland

0001-8880 © 1994 SGM 2263
specific sequence known as the single strand origin (SSO, formerly minus origin) which makes possible the conversion of ssDNA intermediates into dsDNA plasmid molecules. A different genetic organization is observed in plasmids which replicate using the theta mechanism (Hayes et al., 1991). In the genus Leuconostoc, there have been no reports concerning the mode of replication and stability of plasmid molecules. In this paper, we describe the cloning and sequence analysis of the cryptic plasmid pCI411 (2926 bp) from Leuconostoc lactis 533 (type species). Its genetic organization and mode of replication is deduced, and its host range, stability and homology with other plasmids are examined.

### METHODS

**Bacterial strains and plasmids.** All bacterial strains and plasmids used in this study are listed in Table 1. Leuconostoc strains were routinely subcultured in modified MRS (mMRS) broth (Difco) at 30 °C. Lactococcus lactis subsp. lactis MG1614 was grown at 30 °C in M17 medium (Terzaghi & Sandine, 1975) in which glucose (0.5%, w/v) replaced lactose (GM17). Escherichia coli strains were grown in L broth (Sambrook et al., 1989), with shaking at 37 °C. Bacillus subtilis was grown in LB medium at 30 °C. Selective media contained antibiotics at the following concentrations: vancomycin, 100 µg ml⁻¹; chloramphenicol, 10 µg ml⁻¹ for E. coli and Lactobacillus casei; 5 µg ml⁻¹ for Leuconostoc strains, Lc. lactis subsp. lactis and B. subtilis, and 2264
3.5 μg ml⁻¹ for Streptococcus salivarius subsp. thermophilus. Ampicillin (50 μg ml⁻¹) was used to maintain pUC19 in E. coli, and erythromycin (5 μg ml⁻¹) to maintain pC1214 in Leuconostoc.

**DNA preparation and plasmid analysis.** For rapid preparations of plasmid DNA from Leuconostoc, Streptococcus and Lactococcus strains the method of Anderson & McKay (1983) was used, with the following modifications for Leuconostoc species: cells were suspended in TES buffer [50 mM Tris/Cl (pH 8.0); 1 mM EDTA, 6.7% (w/v) sucrose] were treated with 15 μg lysozyme ml⁻¹ in 25 mM Tris/Cl (pH 8.0) at 37 °C for 15 min. After treatment with sodium dodecyl sulfate (20%, w/v) in 50 mM Tris/Cl (20 mM EDTA, pH 8.0), the suspension was left to stand for 20 min at 37 °C. Plasmid DNA was isolated from Lb. casei as described by Chassy & Flickinger (1987). For lysis of E. coli and B. subtilis, the method of Birnboim & Doly (1979) was used. Preparative amounts of plasmid DNA were obtained by scaling up these procedures followed by cesium chloride/ethidium bromide density gradient ultracentrifugation.

**Restriction endonuclease analysis, molecular cloning techniques and nucleotide sequence analysis.** Restriction endonucleases were purchased from Boehringer. DNA digestion and general cloning procedures were as outlined by Sambrook *et al.* (1989). Restriction analysis was performed as described previously (Coffey *et al.*, 1991). pC1411 DNA fragments generated using *DraI*, *HindIII*, *SstI*, *HpaI* and *HaeIII* were cloned into *E. coli* TG1 using pUC19. The nucleotide sequence was determined by sequencing double-stranded plasmid DNA in two orientations by the dideoxy chain termination method (Sanger *et al.*, 1977), using Sequenase version 2.0 (United States Biochemicals) and TaqTrack (Promega). Synthetic 17-mer DNA primers were prepared on a DNA synthesizer (Applied Biosystems). LaserGene (DNASTAR) and MicroGenie (Beckman) were used for computer-assisted sequence analysis.

**Electroporation and transformation.** *E. coli* electro-transformation was performed using a Gene Pulser apparatus and the conditions outlined in the manufacturer's instruction manual (BioRad). *Leuconostoc* cultures were prepared for electro-transformation as follows. Cultures were grown overnight at 30 °C in mMRS broth supplemented with 40 mM DL-threonine and then diluted (1:50) in fresh broth and grown at 30 °C for 3–4 h. Cells were harvested by centrifugation at 4 °C, washed twice in ice-cold 10% (w/v) sucrose (Sigma), resuspended in 0.01 culture volume of ice cold 10% sucrose and held on ice until required. Up to 0.1 volume of plasmid DNA was thoroughly mixed with 100 μl of cell suspension in a chilled Gene Pulser cuvette (electrode separation of 2 mm). Cells were exposed to a single electric pulse (peak voltage 2.5 kV; capacitance 25 μF; resistance 200 Ω) which generated a peak field strength of 125 kV cm⁻¹. Immediately after delivery of the pulse, the cells were added to 5 ml mMRS broth and incubated at 30 °C for 2 h prior to spreading on mMRS agar plates containing the relevant antibiotics. Colonies were visible after 48 h. *Strep. salivarius* subsp. *thermophilus* 054 was electro-transformed as described by Slos *et al.* (1991), *Lc. lactis* subsp. *lactis* MG1614 as described previously (Coffey *et al.*, 1991), *Lb. casei* 20012 as described by Chassy & Flickinger (1987) and *B. subtilis* DB104 as described by Veemanenperä (1989).

**Southern hybridizations.** After fractionation on 0.7% agarose gels, DNA was transferred to nitrocellulose filters by the method of Southern (1975). DNA was labelled using the enhanced-chemiluminance (ECL) gene detection system (Amersham). Probe labelling, hybridization and washing steps were done according to the manufacturer's instructions.

**Detection of ssDNA in whole-cell lysates of *Lc. lactis* and *Leuconostoc*.** This was achieved using the method described by Leenhouts *et al.* (1991).

**In vitro transcription-translation analysis.** Protein products of pC1411 were identified from in vitro transcription-translation studies using a prokaryotic DNA-directed translation kit (Amersham) based on an S30 cell extract from *E. coli*. 35S-labelled proteins were electrophoresed on 0.1% SDS/15% polyacrylamide gels. Gels were fixed in 7% (v/v) acetic acid, impregnated with Amplify (Amersham), dried and exposed to Hyperfilm MP autoradiography film (Amersham) for 48–96 h with an intensifying screen. 14C-methylated proteins (Amersham) were used as molecular mass markers.

**RESULTS AND DISCUSSION**

**Isolation and characterization of pC1411**

*Leuconostoc lactis* 533 is the type strain of this species (Garvie, 1986) and contains two plasmids of 12.5 kb and 2.9 kb, the latter designated pC1411. Total plasmid DNA was isolated from this strain and co-transformed with the 12.4 kb erythromycin resistance plasmid pC1214 (Caplice *et al.*, 1987) into the plasmid-free *Leu. lactis* AHO22. *Em<sup>R</sup>* transformants were obtained at a frequency of 3.52 × 10<sup>4</sup> (μg DNA)⁻¹, of which 11% were shown to harbour pC1411. Repeated subculture at 37 °C was used to cure the marker plasmid from *Em<sup>R</sup>* transformants. A representative derivative containing pC1411 alone was designated *Leu. lactis* AH411. A map of pC1411 showing relevant restriction sites is presented in Fig. 1. In order to locate the genetic elements necessary for replication of pC1411, various constructs were generated by subcloning parts or all of the plasmid into the replication probe vector pCI341 (Hayes *et al.*, 1990), whose replication in *E. coli* HB101 is

![Fig. 1. Map of pC1411 showing relevant restriction sites used for cloning and location of features derived from nucleotide sequence analysis: open reading frames (ORF), inverted repeat structures (IR) with ΔG values, direct repeats (DR), location of putative double strand origin (DSO) and recombination sequences (RS).](image-url)
supported by the pBR322 replicon, which is not functional in the Leuconostoc background. Thus the isolation of CmR transformants in Leu. lactis AH022 implied that a functional pCI411 replicon was present (Table 1). Cloning of the entire plasmid using the unique SstI site and subsequent transformation into E. coli resulted in the isolation of pCI431, which was the only recombinant plasmid which could also replicate in Leuconostoc (Rep*). When parts of pCI411 DNA were removed, generating the constructs pCI430 and pCI432 (Table 1 and Fig. 1), the ability to replicate was abolished, indicating that the region from HincII (934) to HincII (1442) and HpaI (1442) to SstI (2926) contained DNA which was essential for plasmid replication.

---

**Fig. 2.** Nucleotide sequence of plasmid pCI411 and predicted amino acid sequences of ORFs 1, 2 and 3. Boxed, recombination sequences RS, and RS,, putative double strand origin, -10 and -35 sequences, and ribosome-binding sites (RBS); horizontal solid arrows, direct repeats; horizontal dashed arrows, inverted repeats; *, stop codons.
free energy of binding, $\Delta G$, of $-25 \text{ kcal mol}^{-1}$ ($-104.6 \text{ kJ mol}^{-1}$). It has been proposed that inverted repeat structures overlapping Rep promoters could constitute a target site for auto-regulation of Rep protein expression in the theta replicating plasmid pIP501 (Brantl & Behnke, 1992). Comparison of the three pCI411-encoded proteins with the Swiss-Prot protein sequence databank (EMBL, Heidelberg, Germany) and NBRF-PIR (Washington DC, USA) revealed significant similarity with proteins from Gram-positive replicons. ORF1 showed 46.3% homology to the replication initiation protein RepB from the Lactobacillus plantarum plasmid pLB4 (Bates & Gilbert, 1989). There was also significant homology to the replication initiation proteins of pADB201 of Mycoplasma mycoides (41.5%; Bergemann et al., 1989), pA1 from Lb. plantarum (28.7%; Vujicic & Topisirovic, 1993) and pE194 from Staphylococcus aureus (22.2%; Horinouchi & Weisblum, 1982). ORF2 showed 52.1% similarity to ORF C of pWV01, believed to encode the repressor protein regulating synthesis of the replication initiation protein (Leenhouts et al., 1991). There was no apparent similarity between ORF3 and other DNA or protein sequences in the data banks.

The homology of the translated peptide product of ORF1 to the replication initiation proteins of pLB4, pADB201, pA1 and pE194 argues in favour of pCI411 replicating by the RC mechanism. Where this system operates, a specific target site, the DSO, located upstream of the rep gene is nicked by the replication initiation protein. The sequence 5'-TACTACAACACCCCCCTATGT-3', which was homologous (one mismatch) to the DSO of RC plasmids pE194, pSH71, pLS1 and pWV01, was identified 440 bp upstream of the RepB start codon and is believed to be the DSO of pCI411. The observation that the plasmid construct pCI430 (Table 1 and Fig. 1), which lacks this region, could not replicate in Lactobacillus supports this proposal. ORF2 of pCI411, in addition to exhibiting strong homology to RepC of pWV01, is translationally coupled to ORF1 encoding RepB. In pWV01, pLB4 and pLS1 repressor proteins are also translationally coupled to the replication initiation proteins.

The SSO functions as the initiation site for DNA synthesis on the lagging strand. This site is normally located outside the minimal replicon and its deletion may cause a reduction in copy number, plasmid instability and accumulation of ssDNA intermediates (del Solar et al., 1987). Different sequences representing the lagging strand initiation site have been characterized: ssoA (formerly palA, Gruss et al., 1987), ssoU (formerly palU, Bron et al., 1988), soOT (formerly palT, Devine et al., 1989) and a new type not belonging to these three groups which is highly conserved in plasmids like pLP4 and pJL1 (Leer et al., 1992). Typical SSOs are palindromic sequences of up to 200–300 bp in length with the potential to form hairpin secondary structures. In addition, all ssoA homologous so-called RS$_n$ sites at the 3' base of the stem, containing the 18 bp consensus sequence 5'-AAGTTTTTCTCGGCGAT- AAA-3' (Gruss et al., 1987). When pCI411 was examined for such characteristics, a similar sequence with one

---

**Sequence analysis of pCI411**

Plasmid pCI411 was sequenced to completion on both strands (Fig. 2). It comprises 2926 bp. Three significant open reading frames (ORFs) covering 38% of the plasmid were identified. ORF1, from position 1638 to 2264, could encode a peptide of 208 amino acids with a molecular mass of 24.9 kDa. ORF2 begins at nucleotide 2260, overlaps the ORF1 codon by four base pairs and extends to position 2431. It had the capacity to encode a 56 amino acid peptide with a molecular mass of 6.7 kDa. ORF3 extends from position 2444 to 2761 and could encode a 105 amino acid peptide of molecular mass 12 kDa. Transcription-translation analysis of ORF1 indicated the presence of proteins of molecular masses corresponding to those predicted (Fig. 3). The three prominent bands in the sample containing pCI411 are also evident in the sample containing the pCI411–pCI341 recombinant plasmid, which also has the product of the pCI414 cat gene. The length of background frequently encountered with the *in vitro* transcription-translation system is possibly due to the production of truncated derivatives of the larger gene products. All three ORFs have an ATG start codon, show the same orientation and are preceded by identifiable ribosome-binding sites at the consensus distance from the translational start reported for genes of Gram-positive bacteria (van de Guchte, 1991). A putative promoter region is located 261 bases upstream of ORF1. The putative –10 (TCTTAT) and the –35 regions (TGA-ACA) lie within a stem–loop structure with a calculated free energy of binding, $\Delta G$, of $-25 \text{ kcal mol}^{-1}$ ($-104.6 \text{ kJ mol}^{-1}$). It has been proposed that inverted repeat structures overlapping Rep promoters could constitute a target site for auto-regulation of Rep protein expression in the theta replicating plasmid pIP501 (Brantl & Behnke, 1992). Comparison of the three pCI411-encoded proteins with the Swiss-Prot protein sequence databank (EMBL, Heidelberg, Germany) and NBRF-PIR (Washington DC, USA) revealed significant similarity with proteins from Gram-positive replicons. ORF1 showed 46.3% homology to the replication initiation protein RepB from the Lactobacillus plantarum plasmid pLB4 (Bates & Gilbert, 1989). There was also significant homology to the replication initiation proteins of pADB201 of Mycoplasma mycoides (41.5%; Bergemann et al., 1989), pA1 from Lb. plantarum (28.7%; Vujicic & Topisirovic, 1993) and pE194 from Staphylococcus aureus (22.2%; Horinouchi & Weisblum, 1982). ORF2 showed 52.1% similarity to ORF C of pWV01, believed to encode the repressor protein regulating synthesis of the replication initiation protein (Leenhouts et al., 1991). There was no apparent similarity between ORF3 and other DNA or protein sequences in the data banks.

The homology of the translated peptide product of ORF1 to the replication initiation proteins of pLB4, pADB201, pA1 and pE194 argues in favour of pCI411 replicating by the RC mechanism. Where this system operates, a specific target site, the DSO, located upstream of the rep gene is nicked by the replication initiation protein. The sequence 5'-TACTACAACACCCCCCTATGT-3', which was homologous (one mismatch) to the DSO of RC plasmids pE194, pSH71, pLS1 and pWV01, was identified 440 bp upstream of the RepB start codon and is believed to be the DSO of pCI411. The observation that the plasmid construct pCI430 (Table 1 and Fig. 1), which lacks this region, could not replicate in Lactobacillus supports this proposal. ORF2 of pCI411, in addition to exhibiting strong homology to RepC of pWV01, is translationally coupled to ORF1 encoding RepB. In pWV01, pLB4 and pLS1 repressor proteins are also translationally coupled to the replication initiation proteins.

The SSO functions as the initiation site for DNA synthesis on the lagging strand. This site is normally located outside the minimal replicon and its deletion may cause a reduction in copy number, plasmid instability and accumulation of ssDNA intermediates (del Solar et al., 1987). Different sequences representing the lagging strand initiation site have been characterized: ssoA (formerly palA, Gruss et al., 1987), ssoU (formerly palU, Bron et al., 1988), soOT (formerly palT, Devine et al., 1989) and a new type not belonging to these three groups which is highly conserved in plasmids like pLP4 and pJL1 (Leer et al., 1992). Typical SSOs are palindromic sequences of up to 200–300 bp in length with the potential to form hairpin secondary structures. In addition, all ssoA homologous so-called RS$_n$ sites at the 3' base of the stem, containing the 18 bp consensus sequence 5'-AAGTTTTTCTCGGCGAT- AAA-3' (Gruss et al., 1987). When pCI411 was examined for such characteristics, a similar sequence with one
mismatch was located at position 423–440 within a direct repeat and 37 bp downstream of an inverted repeat region (346–360; 372–385, ΔG = −17 kcal mol⁻¹ (−71.1 kJ mol⁻¹). However, unlike ssoAs this region does not contain the consensus sequence 5'-TAGCGT-3' (del Solar et al., 1987) present in the loop of all ssoAs. As no other potential SSO motifs were identified on the plasmid, this region may be the SSO of pCI411.

**Additional structural features of pCI411**

Plasmid pCI411 contains two putative site-specific recombination sequences, RSₐ and RSₐ (Fig. 2). These sites have been studied in *Staph. aureus* plasmids, where they have been shown to be involved in stable cointegrate formation between small plasmids (Novick et al., 1984; Gennaro et al., 1987). The suggested 20 base RSₐ site of pCI411 showed significant homology with that of pE194 (three mismatches); however, a genetic determinant for MOB (mobilization) which is involved in DNA nicking at the RSₐ site in the pE194 system (Gennaro et al., 1987) was not found on the Leuconostoc plasmid. The plasmid recombination/cointegration system mediated by the MOB-RSₐ process is common in *Staphylococcus* and *Bacillus*, and the existence of MOB-RSₐ complexes has been demonstrated in pLB4 from *Lb. plantarum* (Bates & Gilbert, 1989) and pLAB100 from *Lactobacillus hilgardii* (Josson et al., 1990). Like pCI411, an RSₐ region apparently lacking a MOB determinant has been described for plasmid pA1 of *Lb. plantarum* (Vujcic & Topisirovic, 1993) and lactococcal plasmid pFX2 (Xu et al., 1991). In these instances it was speculated that, despite the absence of a MOB-determinant on the plasmid, the function may be encoded by the chromosome or a different plasmid in the original host, and hence cointegrate formation leading to conjugative mobilization may still occur. A number of imperfect direct repeats have been identified on pCI411 (Figs 1 and 2) and it is not known what their function, if any, might be. The potential role of all six of the stem and loop structures identified on pCI411 is also uncertain.

**Detection of ssDNA intermediates during plasmid replication**

Plasmids pCI411, and the control plasmids pNZ12 and pCI305, were assayed for their ability to produce ssDNA during replication. ssDNA was observed during replication of pCI411 as was also the case for pNZ12 which has previously been shown to produce ssDNA (Hayes et al., 1991). pCI305, a small theta-replicating plasmid (Hayes et al., 1991) employed as a negative control, did not generate ssDNA (data not shown). These observations strongly suggest that pCI411 replicates using the RC mechanism.

**Host range and homology of pCI411 with Leuconostoc sp. plasmids**

To determine whether pCI411 was homologous to other plasmids in various Leuconostoc species, pCI411 was labelled and probed against the plasmid profiles of seven strains of *L. lactis* and *Leuconostoc mesenteroides* (Table 1). Under the conditions of the experiment (high stringency) no hybridization was detected between pCI411 and the other Leuconostoc plasmids. pCI431, the CmR derivative of pCI411, was used to transform a number of genera by electroporation. Expression of this CmR gene, which is derived from the staphylococcal plasmid pCI49 (Gruss & Ehrlich, 1989), was previously demonstrated in all the genera chosen for the pCI411 host-range studies. The pCI411 replicon was electroporated into, and shown to function in, Leuconostoc mesenteroides subsp. mesenteroides NCDO2028, Leuconostoc paramesenteroides NCDO1612, *L. lactis* subsp. lactis MG1614, *Streptococcus salivarius* subsp. thermophilus A054, *B. subtilis* DB104 and *Lb. casei* 20012. The replication functions did not operate in *E. coli* HB101, suggesting that pCI411 has a broad host range only among Gram-positive genera. It is, however, interesting that the DNA-directed transcription-translation system, which is based on an *E. coli* S30 extract, was apparently able to transcribe and translate all three ORFs of pCI411, while the plasmid could not be shown to replicate in this genus.

**ACKNOWLEDGEMENTS**

This work was supported by the European Community BRIDGE programme (contract BIOT-CT91-0263). We thank S. Foley for her valuable criticism of this manuscript and T. O’Sullivan for his assistance.

**REFERENCES**


Plasmid pCI411 from Leuconostoc lactis


Received 21 December 1993; revised 24 March 1994; accepted 31 March 1994.