Sequence Analysis of Replication Origin of Plasmid pLS11 of *Bacillus subtilis* IFO 3022


Microbial Genetics Division, Institute of Genetic Resources, and * Department of Food Science and Technology, Faculty of Agriculture, Kyushu University, Hakoizaki, Fukuoka 812, Japan
** Department of Applied Microbial Technology, Kumamoto Institute of Technology, Ikeda, Kumamoto 860, Japan

Received July 10, 1991

The structure of a 1.6-kb *Sphl* — *HindIII* DNA sequence necessary and sufficient for the replication of a 8.6-kb plasmid pLS11 of *Bacillus subtilis* IFO 3022, which is responsible for γ-polylglutamate production, has been characterized by using a trimethoprim (Tmp)-resistance gene derived form *B. subtilis* TTK24 chromosomal DNA as a selective marker. The 1.6-kb DNA sequence contains a *rep* gene encoding the protein (333 amino acids) essential for initiation of replication and a possible origin of replication. The predicted REP protein of pLS11 has an overall homology with the REP proteins of pUH1 (74.8% identity), pBAA1 (92.8%), and pFTB14 (78.7%) in *Bacillus* spp., pLP1 (42.1%) and pLAB1000 (36.3%) in *Lactobacillus* spp., and pUB110 (35.3%) and pC194 (37.4%) in *Staphylococcus aureus*, but has not any similarity with the REP protein of the staphylococcal plasmid pT181.

We have reported that a 5.7-kilobase pair (kb) plasmid designated pUH1, which encodes the γ-glutamyltranspeptidase gene responsible for γ-polylglutamate production, is distributed widely in a number of *Bacillus subtilis* (natto) strains isolated from a fermented soybean food, natto.1,2) *B. subtilis* and *B. subtilis* (natto) should be considered as one species, but these two bacilli are found to be classified separately on the basis of whether biotin is essential for growth or not.3,4) *B. subtilis* IFO 3022, which did not require biotin for the growth but produced a small amount of viscous substances, harbored a single plasmid pLS11 (8.6 kb). 5) Recently, electron microscopy showed that a heteroduplex molecule between pUH1 and pLS11 contains 1.11- and 1.24-kb double-stranded termini.6) The 2.0-kb *BsrEII* DNA fragment of pUH1, including the 1.11-kb double-stranded region, contains a 999-bp open reading frame, a promoter for the open reading frame, and a possible replication origin upstream of the promoter.7)

Analysis of the organization of various plasmids isolated from Gram-positive bacteria, such as pT181,8) pC194,9) pE194,10) pFTB14,11) pLP1,12) pUB110,13) pBAA1,14) and pLAB1000,15) has shown that all of the information necessary for replication is located on fragments of about 1.5 kb. These fragments harbor a *rep* gene, encoding a protein essential for the initiation of replication (REP) and its corresponding target site.

This communication reports that the 1.6-kb fragment of pLS11 contains a 999-bp open reading frame, a promoter for the open reading frame, and a possible replication origin upstream of the promoter. Significant homology was observed between the amino acid sequence predicted from the 999-bp open reading frame and those of

Fig. 1. Derivation of Plasmids Used in This Study.
The chain of solid circles in the diagram indicates the DNA segment containing the *Tmp* gene of *B. subtilis* TTK24. Heavy and thin lines represent the regions of pBR322 and pTL12, respectively. Double lines represent the DNA fragment of pLS11. A, AarI; Bg, BglII; E, EcoRI; H, HindIII; P, PstI; S, SphI.

Fig. 2. Structure and Replication Activity of the Derived Fragments of the 1.6-kb ori Fragment.
The open arrow in the restriction map indicates an open reading frame found in the sequence data (see Fig. 3). + and — indicate, respectively, ability and inability to replicate in the *B. subtilis* host. Derivatives were made by deletion with the indicated restriction enzyme or by Bal31 digestion (B) from the *BglI* site of pNH2.
similar putative REP proteins encoded by the other well-known gram-positive replicons.

Materials and Methods

**Bacterial strains and plasmids.** *Escherichia coli* JM101 and *B. subtilis* JM112 (arg-15, leuB8, thr-5, recE4) were used for a cloning host. The plasmids pLS11 and pATE1 were described previously, and plasmid pTL12 carrying the dihydrofolate reductase gene was kindly provided by Tanaka and Kawano.

**Media.** LB broth and Panassay broth (Difco) for both *B. subtilis* and

---

**Fig. 3.** Nucleotide Sequence of the 1.6-kb SpII-HindIII DNA Fragment.

Nucleotide residues are numbered in the 5'-to-3' direction, beginning with the 5'-end residue originated from the SpII site. The deduced amino acid sequence is given below the nucleotide sequence. Putative promoter elements (−33, −10 and the ribosome binding site) are underlined.
E. coli. Spizizen minimal medium for B. subtilis, and M9 minimal medium for E. coli were the same as those described previously.\textsuperscript{27} The cells carrying trimethoprim-resistant (Tmp\textsuperscript{r}) plasmids were grown in AA medium\textsuperscript{17} containing 1 mg/ml of Tmp.

**DNA manipulations.** Plasmid DNAs from B. subtilis were prepared and purified as described previously.\textsuperscript{27} Restriction endonucleases, T4DNA ligase, and bacterial alkaline phosphatase were purchased from Takara Shuzo Co., Ltd., and used as recommended by the manufacturer. Degradation of DNA with exonuclease Bal31 (Takara Shuzo Co., Ltd.) was done by the procedure of Legerst\textsuperscript{e} et al.\textsuperscript{\textsuperscript{49}}

**Methods for transformation and assessment of ori function.** E. coli was transformed by the method of Morrison,\textsuperscript{199} and B. subtilis was transformed by using protoplasts cells.\textsuperscript{200}

**DNA sequencing.** DNA fragments were subcloned into plasmids pUC18 and pUC19, and DNA sequencing was done by the dideoxy chain termination method\textsuperscript{211} with sequence (United States Biochemical Corporation, Ohio, U.S.A.). Nucleotide and amino acid sequences were analyzed by the Hitachi DNASIS system.

**Results and Discussion**

**Delimitation of the replication origin**

To facilitate the identification of the replication region of pLS11, we used the trimethoprim-resistant (Tmp\textsuperscript{r}) dihydrofolate reductase gene of B. subtilis 168. A schematic presentation of the constructed plasmids is given in Fig. 1. The source of the dihydrofolate reductase-coding gene was a Tmp\textsuperscript{r} strain, TTK24, of B. subtilis 168\textsuperscript{197} and has been cloned in the pBR322 plasmid of Escherichia coli. DNAs from pTL12, carrying the Tmp\textsuperscript{r} dihydrofolate reductase gene, which was constructed by Tanaka and Kawano,\textsuperscript{160} and pBR322 were both treated with EcoRI and HindIII, mixed and ligated by T4 ligase, then pATE1 was constructed. Plasmid pLS11 was digested with SphI and HindIII, and then the ends were filled in with the Klenow fragment to generate blunt ends. The DNA fragments were mixed and ligated to the AatI site of pATE1 by T4 ligase, and then added to B. subtilis MI112 protoplasts. Several Tmp\textsuperscript{r} colonies were obtained on AA agar plates containing Tmp (1 mg/ml) and one of them was used for further study. A plasmid, pNH2, carried in such a Tmp\textsuperscript{r} colony had a molecular size of 8.2 kb (Fig. 1). The physical map of pNH2 using various restriction enzymes is shown in Fig. 1.

To define the boundaries of a functional unit of pLS11 replication, the 1.6-kb SphI–HindIII fragment of pLS11 was digested with selected restriction endonucleases to obtain a set of overlapping DNA fragments. The digests filled in with Klenow fragment were ligated with pATE1, introduced into E. coli by transformation, and selected for ampicillin resistance. The plasmid DNA preparations containing each generated fragment were tested for replication in B. subtilis.

The results are summarized in Fig. 2. The constructed plasmid with the 1.4-kb SphI–HaeIII fragment (fragment 2 in Fig. 2) could replicate in the B. subtilis host, but the recombinant plasmid preparations containing the small fragments (fragment 3 and 4 in Fig. 2) could not replicate.

**Nucleotide sequence of the replication region**

The nucleotides of the 1.6-kb SphI–HindIII fragment were sequenced by the method of Sanger.\textsuperscript{211} Though the strategy is not shown, the nucleotides of both stands were sequenced using numerous restriction fragments to give enough overlapping regions (Fig. 3). Looking for possible open reading frames (ORFs), we found only one large frame (Fig. 2) designated rep, which consisted of 999 bp and encoded a protein molecule with 333 amino acids with a M, of 37,138. Several conserved regulatory sequences similar to the E. coli and Bacillus consensus promoter sequences\textsuperscript{22} were observed 5' upstream of the rep ORF.

Plasmid pNH2 was digested with BglI and then treated for 20 min with exonuclease Bal31 under the conditions in which about 50 bp per min were removed from each end of the DNA molecule. After ligation by T4 ligase, the DNA was transformed into E. coli and subsequently introduced into B. subtilis by protoplast transformation. Tmp\textsuperscript{r} transformants, which contain 1.4-kb fragment 5 in Fig. 2, were obtained at high efficiencies with a 168-bp deletion plasmid generated with Bal31 (fragment 5 in Fig. 2), while no transformants were obtained with a similar 255-bp deletion plasmid (fragment 6 in Fig. 2). It suggests that the putative replication origin of pLS11 is between position 168 and 255. To discover the sequences essential for replication, a homology search was done to find whether there are sequences within this region conserved in the registered Gram-positive replicons. A 34-bp sequence was conserved in four plasmids, pLS11, pUH1, pBA1, and pFTB14 in Bacillus spp. (Fig. 4), and the 14-bp sequence within this region was conserved in all eight plasmids in Gram-positive bacteria including Lactobacillus spp. and Staphylococcus spp. This conserved 14-bp sequence is found in the 55-bp region of pC194 shown by Gros et al.\textsuperscript{23} to have origin activity. Within this 14-bp sequence, in addition, the sequence CTTGATA is the sequence at which nicking of the plus-strand occurs in the initiation of replication of the coliphage φX174,\textsuperscript{24} and this conserved sequence was found in a hairpin region of φX174.

**Amino acid sequence homology of pLS11 REP protein with different Gram-positive replication proteins**

The amino acid sequence of the REP protein coding region of pLS11 was compared with a number of amino acid sequence of proteins registered in GenBank by the homology search system GENAS.\textsuperscript{35} Homologies between the predicted amino acid sequence of the REP protein of pLS11 and those of the REP protein of pUH1, pBA1, pFTB14, pLPl, pLAB1000, pUB100, and pC194 are illustrated in Fig. 5. The REP protein encoded on pLS11 showed a substantial degree of homology to three REP proteins in Bacillus plasmids: 92.8% identity with pLPl 5'-CAATCGCGCTCTCTTTTTATATATA CTatataga-3' pUH1 qAAATGCGCTCTCTTTTTATATATA CTatataga pBA1 CAATCGCGCTCTCTTTTTATATATA CTatataga pFTB14 aAAATGCGCTCTCTTTTTATATATA CTatataga pLPl 5TTCCTTTTTATATATA CTatataga pLAB1000 GTCCTTTTTATATATA CTataga pH10 GTCCTTTTTATATATA CTataga pC194 5TTCCTTTTTATATATA CTataga

Fig. 4. Homology at the Origin for Plus-strand Synthesis among Plasmids pLS11, pUH1,\textsuperscript{148} pBA1,\textsuperscript{144} pFTB14,\textsuperscript{131} pLPl,\textsuperscript{22} pLAB1000,\textsuperscript{151} pUB100,\textsuperscript{10} pC194,\textsuperscript{23} and phage φX174.\textsuperscript{24} Sequences from the regions corresponding to the pLS11 origin for replication are shown, and the same base of given replicons with that of pLS11 is capitalized. The 14-bp sequence conserved in all eight plasmids in gram-positive replicons are boxed. The sequence CTTGATA, which is the site at which nicking of the plus strand occurs in phage φX174, are boldface. The DNA bond between nucleotides G and A will be nicked by the REP protein and is indicated by an arrow.
Fig. 5. Amino Acid Sequence Comparison of the REPpLS1 Protein with Different Gram-positive Replication Proteins.

The original names of these proteins have been replaced by REP, followed by a letter code indicating the plasmid from which they originate. The amino acid identities with REPpUN1,13 REPpBAA1,13 REPpFTB14,13 REPpPL1,13 REPpLAB1000,13 REPpB110,13 and REPpC194 with REPpLS1 are indicated (*). Amino acid numbers follow the sequence of the REP protein from the amino-terminal methionine to the carboxy terminus. Gaps have been inserted to gain maximum matching. The one-letter amino acid code has been used. The tyrosine residue (marked with arrow) probably involved in the binding of the REP protein to the DNA is indicated by an arrow.
REPpBAA1, 78.7% with REPpFTB14, and 74.8% with REPpUH1. The REPpLS12 is similar to REPpLP1 (42.1% identity) and REPpAB1000 (36.3%) from lactic acid bacteria but also to REPpUB110 (35.3%) and REPpC194 (37.4%) from *S. aureus*. The REP protein of pFTB14 stretches for 1,017 bp, a promoter region for *rep* expression, and a possible replication origin for *rep* expression, which is upstream of the promoter. The *rep* product is trans-acting and essential for plasmid replication. The tyrosine, which acts as an active site of REP in the rolling circle mechanism, was conserved in all of the Gram-positive initiation proteins (Fig. 5). Khan et al. identified the start site of pT181 DNA synthesis within a 127-bp segment and showed that a 168 bp segment containing the replication start site is enough to initiate unidirectional replication. Furthermore, like REP protein of pT181, the protein of the *E. coli* plasmid R6K does not have any significant homology in its amino acid sequences with those of the REP proteins of pLS11, pUH1, and also pFTB14 deduced from the ORFs (data not shown). The degree of homology of REPpLS11 with the REP proteins of pLP1 and pLAB1000 in *Lactobacillus* spp. strains was almost the same as those with the REP proteins of pUB110 and pC194 in *S. aureus*. Taxonomic studies (based on the 16S rRNA similarities) showed that *Bacillus* spp. and *S. aureus* strains are more closely related to each other than *Streptococcus* spp. and *Lactobacillus* spp. The plasmid homologies suggest an exchange of plasmid replicons by recent horizontal transfer through the different genera, including *B. subtilis* (natto).

Acknowledgments. We are indebted to Dr. T. Tanaka for his generous supply of plasmid pTL12 and also to Dr. S. Kuhara for the computer search of protein sequences.

References