Mol Gen Genet (1985) 200:193–198

MGG

ISL*1*: a new transposable element in *Lactobacillus casei*

Mariko Shimizu-Kadota[1], **Mayumi Kiwaki**[1], **Hideo Hirokawa**[2], **and Nobuo Tsuchida**[1]*

[1] Yakult Central Institute for Microbiological Research, Kunitachi, Tokyo 186, Japan
[2] Life Science Institute, Sophia University, Kioicho, Chiyoda-ku, Tokyo 102, Japan

Summary. The genome structures of a temperate *Lactobacillus* phage, ϕFSW, and its virulent mutants, ϕFSVs, were examined by restriction, heteroduplex and nucleotide-sequence analyses. The results showed that two out of three ϕFSVs had the same 1.3 kbp insertion (designated as ISL*1*) at different positions in the ϕFSW sequence. ISL*1* was 1,256 bp long and contained at least two long open reading frames of 279 and 822 bases on one strand. Inverted repeats were found at the termini of the ISL*1* which was bracketed by 3 bp direct repeats of the ϕFSW sequence. From this evidence, we concluded that ISL*1* was a transposable element in *Lactobacillus casei.*

Introduction

Transposable elements have now been identified from a variety of organisms (Shapiro 1983). These elements share a common structure; they carry at least one open reading frame coding for a transposase and a terminal inverted repeat serving as a *cis*-acting sequence and both of these seem to be essential for transposition of the elements. Moreover, a direct repeat of the target sequence is generated at the termini as a consequence of transposition. Bacterial transposable elements are classified into two groups: transposons and insertion sequences (IS). Transposons carry at least one additional gene besides that involved in transposition and therefore are generally larger than ISs whose sizes are usually between 0.8 and 1.8 kilobase pairs (kbp) (Kleckner 1981; Iida 1983). Most work on transposable elements has been done in gram-negative bacteria, especially in *Escherichia coli,* while studies on elements in gram-positive bacteria have been limited.

During the course of studying the temperate phage ϕFSW of *Lactobacillus casei* (a gram-positive, lactic-acid-producing bacterium) we found that: (1) in the three virulent mutants, ϕFSV-A, ϕFSV-B and ϕFSV-C, the genomes of the two latter were 1.3 kbp larger than that of wild-type ϕFSW in the region between the 22.4 kbp *Bgl*II site and 23.3 kbp *Eco*RI site on the ϕFSW genome, in contrast to the ϕFSV-A genome which is indistinguishable from that of ϕFSW (Fig. 1). (2) In the enlarged regions the genomes of both ϕFSV-B and ϕFSV-C had incorporated DNA sequences derived from a portion of the host chromosome. (3) The enlargements of the ϕFSV genomes were not caused by illegitimate excision of the ϕFSW prophage, since the enlarged region was more than 7 kbp apart from the prophage attachment site on the host chromosome (Shimizu-Kadota and Tsuchida 1984; Shimizu-Kadota et al. 1984).

In this study, we have analysed the DNA structures of the ϕFSW and ϕFSV genomes in detail and found that the enlargement of both ϕFSV genomes was due to the insertion of a 1.3 kbp transposable element designated as ISL*1* (previously denoted as the V-element).

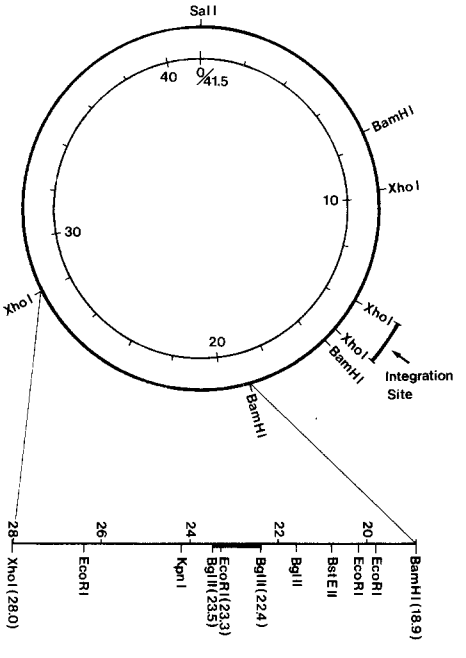

Fig. 1. A restriction map of ϕFSW DNA. The *thick bar* indicates the fragment cloned in pMSK41. The numerical scale indicates the distance (kbp) clockwise from the unique *Sal*I site on the circularly permuted genome. The prophage integration site to the host chromosome is shown (Shimizu-Kadota and Tsuchida 1984; Shimizu-Kadota et al. 1984)

* *Present address:* Department of Oral Microbiology, Dental School, Tokyo Medical and Dental University, Bunkyo-ku, Tokyo 113, Japan

Offprint requests to: M. Shimizu-Kadota

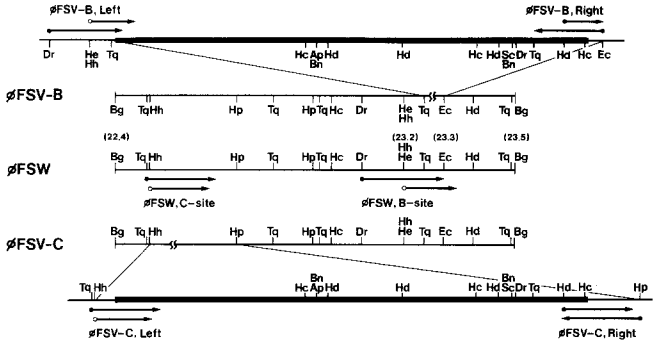

Fig. 2. Fine restriction map of ϕFSW DNA in the region between the 22.4 and 23.5 kbp *Bgl*II sites (map positions of some restriction sites are given) and of the corresponding ϕFSV-B and ϕFSV-C DNAs. Sites for *Hpa*II and *Hha*I are indicated except for those located within the inserts in ϕFSV-B and ϕFSV-C. The sequencing strategy is also shown. An *arrow* indicates the extent and direction of one sequence reaction; *closed circles* stand for sites of 5′ labelling and *open circles* for those of 3′ labelling. Abbreviations: Ap, *Apa*I; Bg, *Bgl*II; Bn, *Ban*II; Dr, *Dra*I; Ec, *Eco*RI; Hc, *Hin*cII; Hd, *Hin*dIII; He, *Hae*II; Hh, *Hha*I; Hp, *Hpa*II; Sc, *Sac*I and Tq, *Taq*I

Materials and methods

Bacteria and phage strains. Group B *Lactobacillus casei* strain C239, which was used as the host of the temperate phage ϕFSW, and ϕFSW-TI, which is a thermoinducible mutant of ϕFSW, have been described previously (Shimizu-Kadota and Sakurai 1982). Phages ϕFSV-B and ϕFSV-C are two out of three independently isolated, virulent mutants of ϕFSW (Shimizu-Kadota et al. 1983). Strain HB101 of *Escherichia coli* K12 and its plasmid vector pTS101 (Sako 1985) were used to clone the restriction fragments of phage DNAs. The 5.6 kbp pTS101 DNA can be cleaved into 3.9 kbp and 1.7 kbp fragments at the unique sites for *Bgl*II and *Pst*I.

Preparation of DNAs and restriction mapping. Phage ϕFSW was prepared from a lysate of C239 (ϕFSW-TI) by heat induction. Phage DNAs were extracted with phenol from purified phage particles as described previously (Shimizu-Kadota et al. 1983). Plasmid DNAs in HB101 cells were extracted as cleared lysates and further purified by two succesive CsCl-ethidium bromide density-gradient centrifugations.

Restriction endonucleases were purchased from Takara Shuzo (Kyoto, Japan), Nippon gene (Tokyo, Japan), Toyobo (Osaka, Japan) and New England BioLabs. (Beverly, MA, USA). DNA was digested under the conditions suggested by the suppliers. DNAs cleaved with restriction endonuclease(s) were electrophoresed in agarose or acrylamide gel slabs. The size markers used were *Hin*dIII digests of lambda DNA and *Hae*III digests of ϕX174 RF DNA (BRL, Rockville, Md, USA). The map position on the circularly permuted ϕFSW genome was indicated as a distance (kbp) clockwise from the unique *Sal*I site.

Molecular cloning. The restriction fragments of the phage DNAs were cloned into *E. coli* cells as described by Maniatis et al. (1982). The recombinant plasmids obtained, pMSK41, pMSK126 and pMSK117, have the inserts of the 1.1 kbp *Bgl*II fragment of ϕFSW and the corresponding 2.4 kbp fragments of ϕFSV-B and ϕFSV-C at the *Bgl*II site of pTS101, respectively with the same orientation.

Electron microscopic heteroduplex analysis. Heteroduplex molecules of *Pst*I-cleaved recombinant plasmid DNAs were formed by incubation at 60° C for 2 h in SSC (0.15 M NaCl,

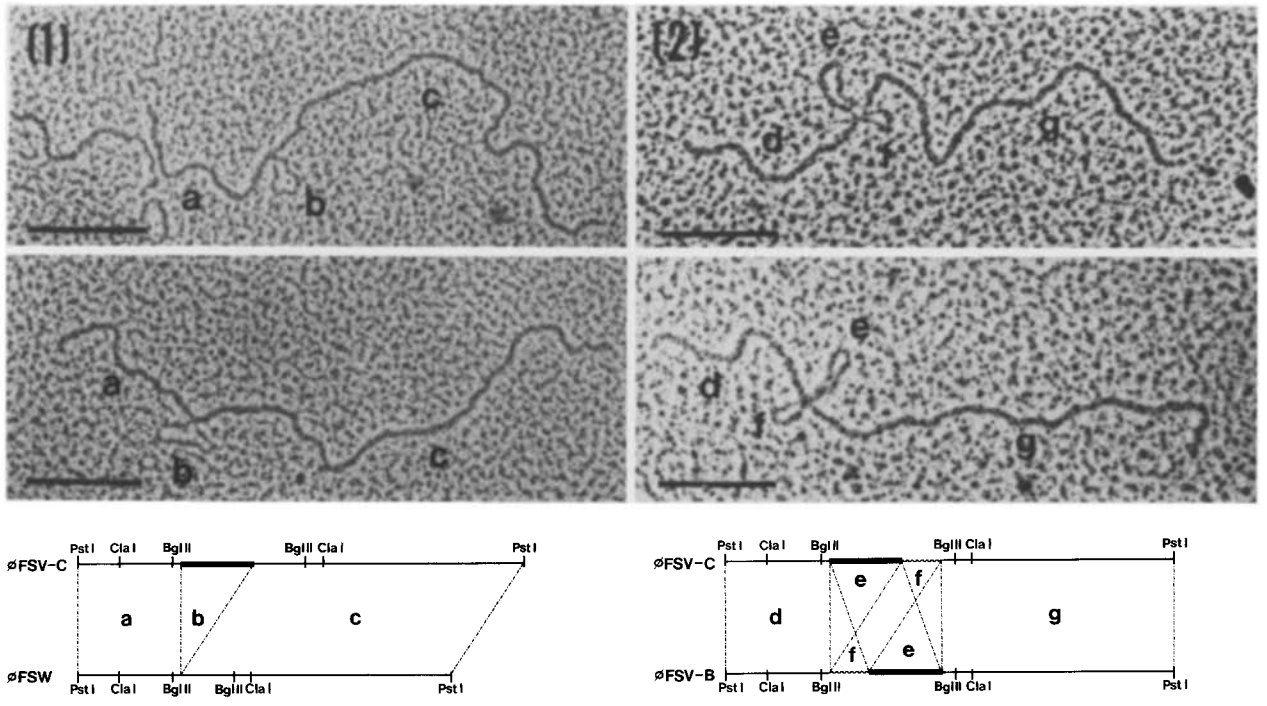

Fig. 3. Representative electron micrographs of heteroduplex DNA molecules of recombinant plasmids between (1) pMSK41 and pMSK117 and (2) pMSK126 and pMSK117; the *bar* represents 1 kbp. Schematic restriction maps including vector regions are presented at the bottom; *thick lines*, ISL*1*s and *wavy lines*, phage DNA segments between two insertion sites

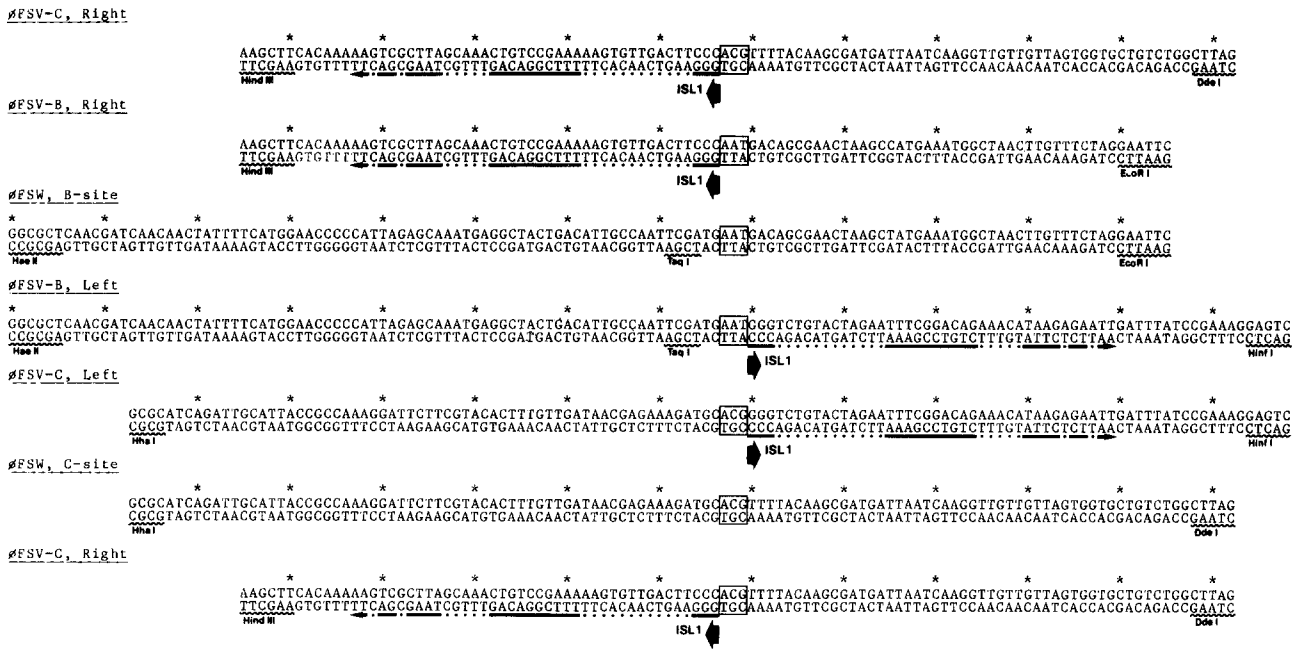

Fig. 4. Nucleotide sequences of the target and junction regions of the ISL*1* in ϕFSV-B and ϕFSV-C. The target sequences in ϕFSW and its duplications in ϕFSV-B and ϕFSV-C are *boxed*. The terminal inverted repeats are indicated by *solid arrows* under the sequences

0.015 M sodium citrate) after boiling at 100° C for 2 min, spread with cytochrome *c* to mount on carbon grids using the formamide technique of Davis et al. (1971) and shadowed with platinum-palladium (Pt:Pd = 80:20). Homoduplexes of the plasmid DNAs were used as internal length standards.

DNA sequencing. The cloned DNAs were sequenced by the method of Maxam and Gilbert (1980) in both strands using five reactions: G, A+G, A>C, C+T and C. The 5′ ends of the restriction fragments were labelled with [γ-^{32}P]ATP (Amersham, Bucks., England) and T4 polynucleotide kinase (Takara Shuzo) after removing the terminal phosphate with alkaline phosphatase from *E. coli* (Takara Shuzo). The 3′ ends were labelled with dideoxyadenosine 5′-[α-^{32}P]triphosphate (Amersham) and terminal transferase (Amersham; Yousaf et al. 1984).

Results and discussion

Restriction mapping of the enlarged regions

In order to elucidate the structural differences between the ϕFSW and ϕFSV genome DNAs, the 1.1 kbp *Bgl*II fragment of ϕFSW DNA (from the 22.4 kbp to the 23.5 kbp *Bgl*II sites in Fig. 1) and the corresponding enlarged fragments of the ϕFSV-B and ϕFSV-C DNAs (the 2.4 kbp *Bgl*II fragments) were cloned into the *E. coli* plasmid vector pTS101. The resulting plasmids were named pMSK41, pMSK126 and pMSK117, respectively, and subjected to more detailed restriction analysis. When the restriction map of the ϕFSV-B or ϕFSV-C fragment was compared with that of ϕFSW DNA, all of the restriction sites of ϕFSW which had been determined were found to be conserved in ϕFSV-B or ϕFSV-C (Fig. 2). However, the *Taq*I-*Eco*RI region in ϕFSV-B and the *Hha*I-*Hpa*II region in ϕFSV-C were 1.3 kbp larger than the corresponding ϕFSW regions and had new sites for *Apa*I and *Sac*I respectively. The restriction sites within these enlarged regions of ϕFSV-B and ϕFSV-C were identical when mapped with *Apa*I, *Ban*II, *Dra*I, *Hin*cII, *Hin*dIII, *Sac*I and *Taq*I. These results suggest that the ϕFSV-B and ϕFSV-C DNAs have the same 1.3 kbp insert with the same orientation but at different sites. The distance between the insertion sites in the ϕFSV-B and ϕFSV-C DNAs was calculated to be 0.7 kbp.

Heteroduplex analysis

The DNA structures of the mapped regions were investigated by heteroduplex analysis. The recombinant plasmid DNAs carrying the phage fragments were linearized by cutting with *Pst*I at a unique site in the vector-derived region and then annealed to each other. Figure 3 shows the electron micrographs of observed representative heteroduplex molecules. The molecules formed between the recombinant plasmid DNAs of pMSK41 and pMSK117 (from ϕFSW and ϕFSV-C) were linear with a single-stranded loop of 1.20 ± 0.35 kbp. The location of the loop was 4.86 ± 0.08 kbp inwards from one end and 1.84 ± 0.08 kbp inwards from the other end (Fig. 3-1). When heteroduplex molecules were formed between the DNAs of pMSK41 and pMSK126 (from ϕFSW and ϕFSV-B), the linear molecules also had a single-stranded loop of a similar size but its location was shifted inwards by about 0.7 kbp (data not shown). These observations show that ϕFSV-B and ϕFSV-C had inserts 0.7 kbp apart on the cloned fragments but retained most of the ϕFSW genome sequence.

When pMSK126 DNA was hybridized to pMSK117 DNA, heteroduplex molecules observed had two double-stranded loops, 1.28 ± 0.08 kbp and 0.72 ± 0.06 kbp, at the same site. The position was determined to be 1.87 ± 0.07 kbp inwards from one end and 4.14 ± 0.08 kbp inwards from the other end (Fig. 3-2). The length of the larger loop corresponded to the size of the insert and that

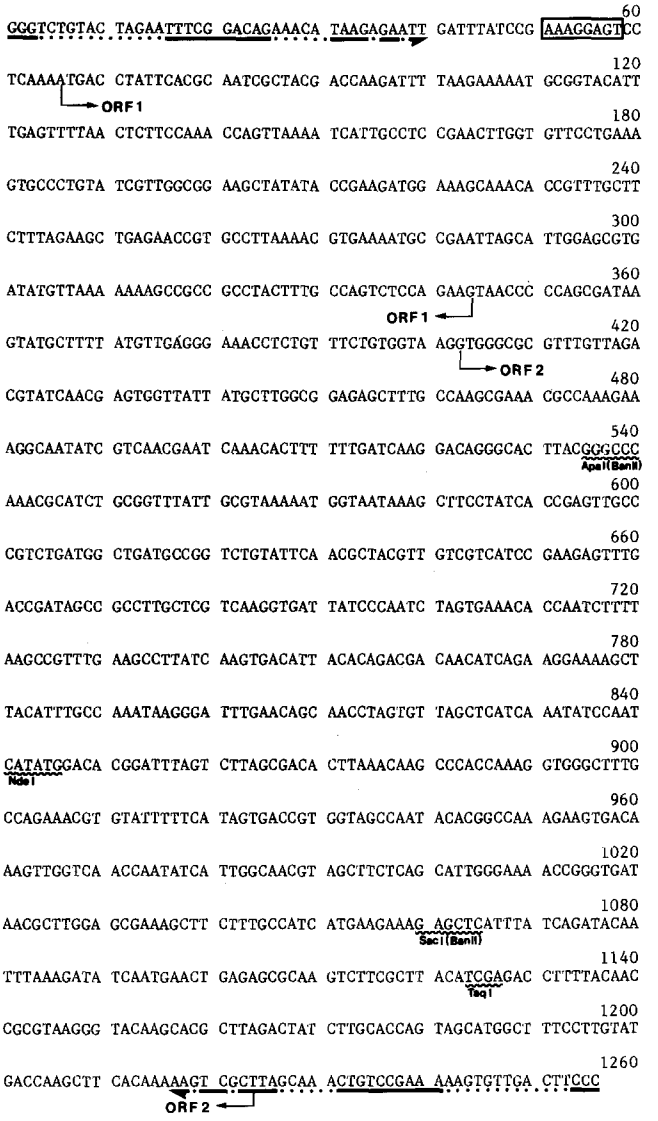

Fig. 5. Entire nucleotide sequence of the ISL*1* DNA in ϕFSV-B. Only the 5′-3′ strand is shown. Termninal inverted repeats are indicated with *arrows* and several restriction sites with *wavy lines*. The sequence similar to that of the ribosome binding site in *Escherichia coli* is *boxed*. The open reading frames (ORF 1 and ORF 2) described in the text are indicated

of the smaller one to the distance between the two insertion sites in ϕFSV-B and ϕFSV-C. These facts clearly indicate that the sequences of the inserted DNA segments in ϕFSV-B and ϕFSV-C and their orientations with respect to their locations on the ϕFSW genomes are the same.

Nucleotide sequences in the junction region between the phages and the inserted segments

In order to analyse further the nature of the insertion, the nucleotide sequences of the junction regions of the ϕFSV DNAs were determined by the method of Maxam and Gilbert and compared with those of the corresponding regions of ϕFSW (Fig. 4). In both the ϕFSV-B and ϕFSV-C genomes flanking the inserts, the ϕFSW sequence had been conserved completely, except for one base at the 19th position from the right junction in ϕFSV-B. At the left and right junctions of the insert, a 3 bp sequence of ϕFSW had been duplicated in the same orientation. However, the duplicated sequence of ϕFSV-B was different from that of ϕFSV-C and no common sequences were found within 50 bp of the ϕFSW sequences flanking the inserts in ϕFSV-B and ϕFSV-C. When both of the corresponding terminal sequences of the inserts in ϕFSV-B and ϕFSV-C were compared, the sequences up to 50 bp were exactly the same. This result suggests that the inserts in ϕFSV-B and ϕFSV-C were the same at the sequence level. The termini of the inserts were able to form the inside inverted repeats as shown by arrows in Fig. 4. Since these structural properties are commonly found in well known transposable elements as a consequence of transposition and ϕFSV-B and ϕFSV-C had been derived from ϕFSW with the insertion of the 1.3 kbp segment, we concluded that the 1.3 kbp segment inserted into the ϕFSW genome was a transposable element. Judging from the size, the segment was likely to be one of the IS elements and was designated as ISL*1* as we could not find any genes unrelated to the insertion function.

The nucleotide sequence of the ISL1 DNA

The ISL*1* DNA inserted in the ϕFSV-B genome was sequenced as shown in Fig. 5. The ISL*1* from ϕFSV-B was 1,256 nucleotides long and its G+C content was calculated to be 43.7%, a value close to that of the host chromosomal DNA (45%; T. Kuroshima, unpublished). In the previous paper (Shimizu-Kadota and Tsuchida 1984), we showed that the 9.1 kbp fragment of ϕFSW DNA between 18.9 and 28.0 kbp on the ϕFSW genome was not able to hybridize to the host chromosomal DNA but the corresponding fragment of the ϕFSV-B or ϕFSV-C DNA containing the ISL*1* sequence could do so. Moreover, it has been shown that the 1.4 kbp *Hae*II-*Eco*RI fragment of the ϕFSV-B DNA, consisting of the entire ISL*1* and 0.1 kbp ϕFSW-derived sequences, is able to hybridize to the host chromosomal DNA (Shimizu-Kadota, unpublished). Therefore, it is clear that the ISL*1* sequence was derived from the host chromosome.

The determined DNA sequence of ISL*1* was compared by two-dimensional dot-matrix homology analysis (Maizel, Jr and Lenk 1981) with that of the known IS elements IS*1* (Ohtsubo and Ohtsubo 1978), IS*2* (Ghosal et al. 1979), IS*4* (Klaer et al. 1981), IS*5* (Engler and van Bree 1981; Schoner and Kahn 1981), IS*10*R (Halling et al. 1982), IS*15* (Trieu-Cuot and Courvalin 1984), IS*26* (Mollet et al. 1983), IS*50*R (Auerswald et al. 1981), IS*102* (Bernardi and Bernardi 1981) and IS*903* (Grindley and Joyce 1980; Oka et al. 1981) of *E. coli,* and ISH*1* (Simsek et al. 1982), ISH*2* (Das-Sarma et al. 1983) and ISH*50* (Xu and Doolittle 1983) of *Halobacterium.* However, no significant sequence homologies were found in the forward or reverse strands. Therefore, ISL*1* is identified as a new transposable element, which belongs to *Lactobacillus casei.*

The protein coding capacity of ISL*1* is illustrated in Fig. 6. Two long open reading frames, designated as ORF 1 and ORF 2, were found on the upper strand, although several smaller ones of less than 250 bases could be assigned on both strands. ORF 1, from positions 66 to 344, and ORF 2, from positions 403 to 1,224, are large enough to code for proteins of 10.7 and 31.7 kilodaltons, respectively.

Some genes derived from lactobacilli have been expressed in *E. coli* cells (Lee et al. 1982; Davies and Gronenborn 1982; Williams et al. 1984), suggesting that they have

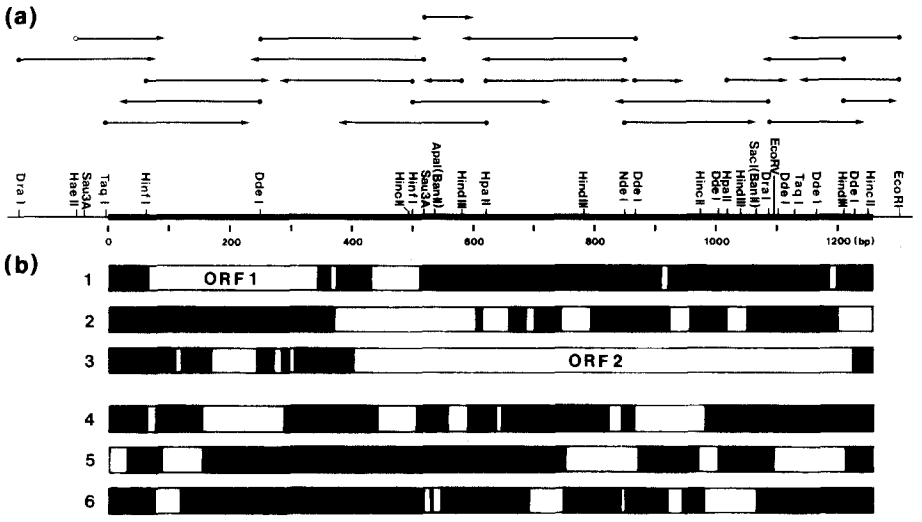

Fig. 6. a Detailed restriction map of ISL*1* and the adjacent region in ϕFSV-B with the sequencing strategy. The region of the inserted ISL*1* is shown as a *thick line*. The sequencing strategy is shown with *arrows* as in Fig. 2.
b Open reading frames in the ISL*1* in ϕFSV-B. Phases 1–3 are from the 5′-3′ strand and phases 4–6 are from the 3′-5′ strand; *open boxes* represent possible reading frames

the necessary initiation and termination signals. However, no sequences homologous to the common promoter or transcriptional terminator of *E. coli* (Rosenberg and Court 1979; Gold et al. 1981) were found for ORFs 1 and 2. Only an 8 bp sequence similar to that of the ribosome binding site was found 8 bp upstream from the putative initiation codon of ORF 1 (boxed in Fig. 6). If ORFs 1 and 2 are actively expressed in lactobacilli, the initiation and termination signals of these ORFs could be different from those of lactobacilli which are expressed in *E. coli*.

Function of ISL1 in the ϕFSV genome

The genomes of two out of three virulent mutants, ϕFSV-B and ϕFSV-C, had the same ISL*1* with the same orientation at positions separated by 0.7 kbp. Therefore, the insertion of ISL*1* into the ϕFSW genome may be responsible for the acquisition of virulence. If this is the case, two possibilities are proposed: (1) the ISL*1* insertion inactivates the phage operator(s) or (2) ISL*1* provides a new promoter for the phage genome which is functional in the presence of the phage repressor. Transposable elements in *E. coli* are known to have promoters which activate genes adjacent to the target by insertion (Kleckner 1981; Iida et al. 1983). In either alternative, the phage genes essential for lytic growth could be expressed constitutively in the ϕFSW lysogen.

Acknowledgements. We are grateful to Prof. E. Ohtsubo for invaluable discussions. We thank Mr. T. Sako for providing us with the pTS101 DNA, Mr. K. Satake and Mrs. Y. Tsuchida for helping with the computer analysis of nucleotide sequence and Mrs. T. Ichinose for typing the manuscript. We also thank Dr. M. Mutai for his encouragement throughout the course of the work.

References

Auerswald E-A, Ludwig G, Schaller H (1981) Structural analysis of Tn*5*. Cold Spring Harbor Symp Quant Biol 45:107–113

Bernardi A, Bernardi F (1981) Complete sequence of an IS element present in pSC101. Nucleic Acids Res 9:2905–2911

DasSarma S, RajBhandary UL, Khorana HG (1983) High-frequency spontaneous mutation in the bacterio-opsin gene in *Halobacterium halobium* is mediated by transposable elements. Proc Natl Acad Sci USA 80:2201–2205

Davies RW, Gronenborn AM (1982) Molecular cloning of the gene for dihydrofolate reductase from *Lactobacillus casei*. Gene 17:229–233

Davis RW, Simon M, Davidson N (1971) Electron microscope heteroduplex methods for mapping regions of base sequence homology in nucleic acids. Methods Enzymol 31:413–428

Engler JA, van Bree MP (1981) The nucleotide sequence and protein-coding capability of the transposable element IS*5*. Gene 14:155–163

Ghosal D, Sommer H, Saedler H (1979) Nucleotide sequence of the transposable DNA-element IS*2*. Nucleic Acids Res 6:1111–1122

Gold L, Pribnow D, Schneider T, Shinedling S, Singer BS, Stormo G (1981) Translational initiation in prokaryotes. Annu Rev Microbiol 35:365–403

Grindley NDF, Joyce CM (1980) Genetic and DNA sequence analysis of the kanamycin resistance transposon Tn*903*. Proc Natl Acad Sci USA 77:7176–7180

Halling SM, Simons RW, Way JC, Walsh RB, Kleckner N (1982) DNA sequence organization of IS*10*-right of Tn*10* and comparison with IS*10*-left. Proc Natl Acad Sci USA 79:2608–2612

Iida S, Meyer J, Arber W (1983) Prokaryotic IS Elements. In: Shapiro JA (ed) Mobile genetic elements. Academic Press, New York, pp 159–221

Klaer R, Kühn S, Tillmann E, Fritz H-J, Starlinger P (1981) The sequence of IS*4*. Mol Gen Genet 181:169–175

Kleckner N (1981) Transposable elements in prokaryotes. Annu Rev Genet 15:341–404

Lee L-J, Hansen JB, Jagusztyn-Krynicka EK, Chassy BM (1982) Cloning and expression of the β-D-phosphogalactoside galactohydrolase gene of *Lactobacillus casei* in *Escherichia coli* K-12. J Bacteriol 152:1138–1146

Maizel Jr JV, Lenk RP (1981) Enhanced graphic matrix analysis of nucleic acid and protein sequences. Proc Natl Acad Sci USA 78:7665–7669

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, New York

Maxam AM, Gilbert W (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol 65:499–560

Mollet B, Iida S, Shepherd J, Arber W (1983) Nucleotide sequence of IS*26*, a new prokaryotic mobile genetic element. Nucleic Acids Res 11:6319–6330

Ohtsubo H, Ohtsubo E (1978) Nucleotide sequence of an insertion element, IS*1*. Proc Natl Acad Sci USA 75:615–619

Oka A, Sugisaki H, Takanami M (1981) Nucleotide sequence of the kanamycin resistance transposon Tn*903*. J Mol Biol 147:217–226

Rosenberg M, Court D (1979) Regulatory sequences involved in

the promotion and termination of RNA transcription. Annu Rev Genet 13:319–353
Sako T (1985) Overproduction of staphylokinase in *Escherichia coli* and its characterization. Eur J Biochem, in press
Schoner B, Kahn M (1981) The nucleotide sequence of IS*5* from *Escherichia coli*. Gene 14:165–174
Shapiro JA (ed) (1983) Mobile genetic elements. Academic Press, New York
Shimizu-Kadota M, Sakurai T (1982) Prophage curing in *Lactobacillus casei* by isolation of a thermoinducible mutant. Appl Environ Microbiol 43:1284–1287
Shimizu-Kadota M, Tsuchida N (1984) Physical mapping of the virion and the prophage DNAs of a temperate *Lactobacillus phage* ϕFSW. J Gen Microbiol 130:423–430
Shimizu-Kadota M, Sakurai T, Tsuchida N (1983) Prophage origin of a virulent phage appearing on fermentations of *Lactobacillus casei* S-1. Appl Environ Microbiol 45:669–674
Shimizu-Kadota M, Kiwaki M, Hirokawa H, Tsuchida N (1984) A temperate *Lactobacillus* phage converts into virulent phage possibly by transposition of the host sequence. Dev Ind Microbiol 25:151–159
Simsek M, DasSarma S, RajBhandary UL, Khorona HG (1982) A transposable element from *Halobacterium halobium* which inactivates the bacteriorhodopsin gene. Proc Natl Acad Sci USA 79:7268–7272
Trieu-Cuot P, Courvalin P (1984) Nucleotide sequence of the transposable element IS*15*. Gene 30:113–120
Williams SA, Hodges RA, Strike TL, Snow R, Kunkee RE (1984) Cloning the gene for the malolactic fermentation of wine from *Lactobacillus delbrueckii* in *Escherichia coli* and yeasts. Appl Environ Microbiol 47:288–293
Xu W-L, Doolittle WF (1983) Structure of the archaebacterial transposable element ISH*50*. Nucleic Acids Res 11:4195–4199
Yousaf SI, Carroll AR, Clarke BE (1984) A new and improved method for 3′-end labelling DNA using [α-^{32}P]ddATP. Gene 27:309–313

Communicated by M. Takanami

Received February 4, 1985