JOURNAL OF BACTERIOLOGY, July 1997, p. 4164–4171 Vol. 179, No. 13
0021-9193/97/$04.00+0

Duplication of the *pepF* Gene and Shuffling of DNA Fragments on the Lactose Plasmid of *Lactococcus lactis*

MICHÈLE NARDI,[1] PIERRE RENAULT,[2] AND VÉRONIQUE MONNET[1]*

Unité de Recherches de Biochimie et Structure des Protéines[1] and Laboratoire de Génétique Microbienne,[2] INRA, 78352 Jouy en Josas Cedex, France

Received 22 October 1996/Accepted 24 April 1997

The gene corresponding to the lactococcal oligopeptidase PepF1 (formerly PepF [V. Monnet, M. Nardi, A. Chopin, M.-C. Chopin, and J.-C. Gripon, J. Biol. Chem. 269:32070–32076, 1994]) is located on the lactose-proteinase plasmid of *Lactococcus lactis* subsp. *cremoris* NCDO763. Use of the *pepF1* gene as a probe with different strains showed that *pepF1* is present on the chromosome of *Lactococcus lactis* subsp. *lactis* IL1403, whereas there is a second, homologous gene, *pepF2*, on the chromosome of strain NCDO763. From hybridization, PCR amplification, and sequencing experiments, we deduced that (i) *pepF1* and *pepF2* exhibit 80% identity and encode two proteins which are 84% identical and (ii) *pepF2* is included in an operon composed of three open reading frames and is transcribed from two promoters. The protein, encoded by the gene located downstream of *pepF2*, shows significant homology with methyltransferases. Analysis of the sequences flanking *pepF1* and *pepF2* indicates that only a part of the *pepF2* operon is present on the plasmid of strain NCDO763, while the operon is intact on the chromosome of strain IL1403. Traces of several recombination events are visible on the lactose-proteinase plasmid. This suggests that the duplication of *pepF* occurred by recombination from the chromosome of an *L. lactis* subsp. *lactis* strain followed by gene transfer. We discuss the possible functions of PepF and the role of its amplification.

The composition and the functioning of the proteolytic system of lactic acid bacteria, essential for optimal growth in milk, are now well-known, especially for lactococci (36). In this system, a cell envelope-associated proteinase hydrolyzes casein into oligopeptides which constitute the main source of nitrogen (18). The oligopeptides are then partly taken up by a specific transport system (47) and internally hydrolyzed by peptidases into amino acids, which are used for synthesis of the bacterial proteins. We already know that oligopeptides could be hydrolyzed from their extremities by about 10 different exopeptidases; they are also probably cleaved internally by two recently characterized oligopeptidases, PepO and PepF (29, 33). The two oligopeptidases are metalloenzymes, but they differ in their substrate specificities and are grouped in different peptidase families (32). In a previous paper (33), the sequence of the *pepF* gene was reported and it was shown that PepF belongs to the M3 peptidase family, which includes the mammalian oligopeptidase 24.15, according to the classification of Rawlings and Barrett (38). This gene is located on a large, 55-kb plasmid which also carries the genes coding for lactose and casein utilization, essential for optimal growth in milk (21).

Three observations suggested the existence of another gene (*pepF2*), homologous to that carried by the plasmid (now named *pepF1*) but located on the chromosome of strain NCDO763. First, we copurified with PepF1 another protein displaying the same properties; second, we measured a residual PepF activity in strains cured of the plasmid carrying the *pepF1* gene; and third, by hybridization experiments, we identified a DNA fragment having some homology with *pepF1*, used as a probe. The hybridization experiments also showed that *pepF1* was present on the chromosome of strain IL1403.

In the present study, we characterized the two genes coding for PepF and their flanking sequences in *Lactococcus lactis* subsp. *cremoris* NCDO763 and *L. lactis* subsp. *lactis* IL1403. Our observations suggest that the duplication of *pepF* occurred by recombination events, traces of which are numerous on the conjugative lactose-proteinase plasmid. We discuss the possible roles of PepF1 and PepF2, which may be an example of the first step of gene divergence in the evolution of the lactococcal genome.

* Corresponding author. Mailing address: Unité de Recherches de Biochimie et Structure des Protéines, INRA, 78352 Jouy en Josas Cedex, France. Phone: (33) 0134652149. Fax: (33) 0134652163. E-mail: monnet@jouy.inra.fr.

MATERIALS AND METHODS

Purification and N-terminal sequencing of PepF2. PepF1 and PepF2 were copurified as previously described (33). The purified oligopeptidases were separated by electrophoresis under denaturing and reducing conditions using 12.5% acrylamide gels as described by Laemmli (22). The oligopeptidases were blotted on a polyvinylidene difluoride membrane (Problott; Applied Biosystems, San Jose, Calif.) with a Bio-Rad apparatus (1 h, 100 V). The blotting was done in 10 mM CAPS buffer (pH 11) containing 10% methanol and 0.5 mM dithiothreitol (27). The proteins were stained with Ponceau red. The N-terminal sequence was determined directly on the part of the membrane containing PepF2 with an automatic sequencer (model 477A; Applied Biosystems).

Bacterial strains and plasmids, transformation, and growth conditions. The bacterial strains and plasmids used in this study are listed in Table 1. According to its genotype (15), strain NCDO763 is designated an *L. lactis* subsp. *cremoris* strain throughout the paper. *L. lactis* strains were grown at 30°C in M17 medium (46) supplemented with 0.5% lactose or glucose. TIL80 and MG1363 were grown on a minimal medium containing vitamins, salts, nucleic acids, buffer, and glucose as described for the chemically defined medium by Otto et al. (34) and the following amino acids: serine (0.34 g/liter), histidine (0.11 g/liter), methionine (0.12 g/liter), isoleucine (0.20 g/liter), valine (0.33 g/liter), leucine (0.47 g/liter), and glutamic acid (0.20 g/liter) (5). The *Escherichia coli* strain was grown in Luria-Bertani medium (42) at 37°C. When needed, erythromycin (5 μg/ml for *L. lactis* subsp. *lactis* and 150 μg/ml for *E. coli*) or ampicillin (50 μg/ml for *E. coli*) was added to the culture medium. *L. lactis* electroporation was done as previously described (17). *E. coli* was transformed as described by Sambrook et al. (42).

DNA techniques. Plasmid and chromosomal DNAs were prepared as previously described (25, 42, 44). The classical pulsed-field gel electrophoresis procedure for *L. lactis* was used (24). Restriction endonucleases, T4 DNA ligase, and T4 polymerase were obtained from Boehringer Mannheim or Eurogentec and used as recommended by the suppliers. After agarose gel electrophoresis, Southern blotting and DNA or RNA hybridization were performed either as described by Amersham International (ECL gene detection system) or by the

TABLE 1. Strains and plasmids

Strain or plasmid	Description	Source or reference
Strains		
L. lactis		
IL1403	Plasmid free	Chopin et al. (4)
NCDO763 *lac*$^+$ *prt*$^+$	Wild-type strain	Davies et al. (6)
TIL75	NCDO763 cured of the *lac-prt* plasmid	V. Juillard (17a)
TIL70	NCDO763 *lac*$^+$ *prt*$^+$ *pepF1*	Monnet et al. (33)
TIL79	NCDO763 *lac*$^+$ *prt*$^+$ *pepF2*	This work
TIL246	NCDO763 containing pTIL125 (high copy number)	This work
TIL248	NCDO763 containing pTIL125 (low copy number)	This work
TIL247	NCDO763 containing pTIL126 (high copy number)	This work
TIL249	NCDO763 containing pTIL126 (low copy number)	This work
TIL250	NCDO763 containing pJIM2366 (high copy number)	This work
TIL251	NCDO763 containing pJIM2366 (low copy number)	This work
NCDO712 *lac*$^+$ *prt*$^+$	Wild-type strain	M. J. Gasson (12)
MG1363	Plasmid-free derivative of NCDO712	M. J. Gasson (12)
TIL80	MG1363 *pepF1 pepF2*	This work
E. coli TG1		T. J. Gibson (14)
Plasmids		
pBluescript SK$^+$	Ap^r, M13 ori, pBr322 ori	Stratagene
pIL253	Em^r; 4.6 kb	Simon and Chopin (45)
pJIM2366	Derivative of pIL252; Em^r *lux A/B*	Renault et al. (39)
pTIL28	1.4-kb fragment *pepF1* in SK$^+$; Ap^r	Monnet et al. (33)
pTIL34	1.1-kb fragment *pepF2* in SK$^+$; Ap^r	This work
pTIL35	2.3-kb *Sau*3A fragment of pIL253 in pTIL34; Em^r Ap^r; 6.9 kb	This work
pTIL36	*pepF2* C-terminal part in pTIL35 9 kb	This work
pTIL37	*pepF2* N-terminal part in pTIL35; 12 kb	This work
pTIL125	pJIM2366 containing the P1 promoter of *pepF2*	This work
pTIL126	pJIM2366 containing the P2 promoter of *pepF2*	This work

procedure described by Sambrook et al. (42). A *pepF1* probe (33) was prepared from the *Bcl*I/*Cla*I DNA fragment (positions 416 to 1934) purified from pTIL28.

PCR amplifications were done on a Perkin-Elmer DNA thermal cycler 480. The internal *pepF2* fragment was amplified by using degenerate oligonucleotides 1 [5′ GAAGC(A/T/C/G)TTAACTTGGGA 3′, corresponding to the N-terminal sequence of PepF2] and 2 [5′ (T/A)(G/A)TG(C/G/T/A)CC(C/G/T/A)GT(C/T)TC(G/A)TG 3′, corresponding to the sequence of the zinc-binding motif of PepF1]. Annealing was performed at 44°C. For comparison of the *pepF* genes of different strains, inverse PCR was done with oligonucleotides deduced from two identical parts of the sequences of *pepF1* and *pepF2*: oligonucleotides 3 (5′ CATAGAAACGGGGAGGAT 3′) and (5′ AGGAGTCCATCTTTAAA 3′) (see Fig. 5). The conditions were as follows. The total DNAs of different strains were completely digested by *Hin*dIII, religated in diluted solution (500 ng per ml), and amplified by PCR (with annealing at 48°C). Oligonucleotides 5 (5′ TTGGCATGGAACCTCAAGCTC 3′) and 6 (5′ GAGGAATTTAGCAGTAC 3′), used to amplify the fragments flanking *pepF1*, were deduced from the sequences of IS*904* (37) and *lacR* (8), respectively. Finally, to complete the sequences, two oligonucleotides deduced from the sequence of the amplified fragment (with oligonucleotides 3 and 5), oligonucleotides 7 (5′ CCCCATAAG CCAAAGCAC 3′) and 8 (5′ GAAAAACCCTAAAGGTC 3′), were used. To test the functionality of the two promoters located upstream of *pepF2*, P1 and P2, we amplified two fragments containing either P1 or P2 with oligonucleotides 9 (5′ ATGGACGACGCGTCCGAA 3′) and 10 (5′ CCTAAATGCTGGTCTGAC 3′) for P1 and with oligonucleotides 11 (5′ AAGGTCCATACATTGCTC 3′) and 12 (5′ GACCGTTTCCCGACATTT 3′) for P2. The oligonucleotides were synthesized by Eurogentec or Genosys.

Construction of plasmids. A 1.1-kb fragment obtained by PCR amplification of the DNA of strain TIL75 with oligonucleotides 1 and 2, deduced from the N-terminal sequence of PepF2 and from the active site of PepF1, respectively, was cloned into plasmid pBluescript SK$^+$ in *E. coli* to give pTIL34. To clone the entire *pepF2* gene and its flanking regions, we used the integrative plasmid pTIL35 (see Fig. 2), which was constructed by cloning the erythromycin resistance gene of pIL253 (*Sau*3A fragment) into the *Bam*HI site of pTIL34 and which conferred to *E. coli* TG1 both Ap and Em resistance. After transformation of *L. lactis* subsp. *cremoris* NCDO763, the Em^r clones containing pTIL35 integrated into the chromosome were checked by hybridization with pTIL35 used as a probe. A clone containing a single copy of pTIL35 integrated into the chromosome was chosen and named TIL79. To clone the fragments adjacent to pTIL35, we mapped on the chromosome of TIL79 the *Spe*I, *Sac*I, *Eco*RI, *Cla*I, *Xho*I, *Xba*I, and *Pst*I sites (see Fig. 2) which are unique in the polylinker of pBluescript SK$^+$. *Eco*RI and *Sac*I sites present at 7 and 4 kb, respectively, from the PCR fragment were chosen to recover the *pepF*-flanking fragments. Chromosomal DNAs digested with *Sac*I or *Eco*RI and self-ligated were used to transform *E. coli* TG1. Two clones containing the pTIL36 (7-kb) and pTIL37 (9-kb) plasmids (see Fig. 2) were selected, and their inserts were sequenced.

To check the functionality and measure the strength of the two potential promoters, P1 and P2, located upstream of *pepF2*, we cloned them upstream of the luciferase gene, whose expression can be easily measured. P1 and P2 were subcloned from PCR fragments (nucleotides 553 to 1140 and 1715 to 3064, respectively; see Fig. 4) into the pJIM2366 plasmid (39) between the *Xba*I and *Not*I sites. Plasmid pJIM2366 contains the luciferase gene and the erythromycin resistance gene; furthermore, it can switch from high copy number to low copy number per cell by a simple *Kpn*I digestion followed by ligation (39). The resulting plasmids, pTIL125 and pTIL126, containing P1 and P2, respectively, were introduced directly or after *Kpn*I digestion into *L. lactis* subsp. *cremoris* NCDO763. The resulting strains (TIL246 to TIL251; Table 1) were grown in milk and in M17 medium. The cells were counted and the expression of the *lux* gene was monitored on a Bertold Lumat LB9501 apparatus as previously described (39). The CFU per milliliter and luciferase activities were calculated as the averages of several measures during exponential growth in more than two independent experiments.

DNA sequencing. The amplified fragments were extracted from 0.7% agarose with spin-X (Costar, Cambridge, United Kingdom) and sequenced. For sequencing, series of overlapping clones were generated by using exonuclease III and mung bean nuclease (Stratagene, La Jolla, Calif.) or DNase I (Boehringer Mannheim). The dideoxy chain termination method of DNA sequencing was carried out on double-stranded DNA plasmids with *Taq* dye primer (cycle sequencing kit) and the *Taq* dye deoxy terminator (Applied Biosystems) on a PCR apparatus (DNA thermal cycler 480; Perkin-Elmer). Sequencing reaction mixtures were primed with fluorescent oligonucleotides (Applied Biosystems), and the products were analyzed on an automatic sequencer (370A DNA sequencer; Applied Biosystems). The reported sequences were determined at least twice for both strands. The DNA and protein sequences were analyzed with the Genetics Computer Group sequence analysis software package from the University of Wisconsin (7), the Genmark program (3), and MAIL FASTA (National Center for Biotechnology Information).

Nucleotide sequence accession numbers. The Genbank, EMBL, and DDBJ nucleotide sequence accession numbers are X99710 for *pepF2* and its flanking sequences and X99798 for *pepF1p* and its flanking sequences.

RESULTS

Presence of two PepF oligopeptidases in *L. lactis* subsp. *cremoris* NCDO763. During the purification of oligopeptidase PepF1, another protein which has the same endopeptidase activity on bradykinin was copurified (33). The N-terminal sequence of this protein, determined up to the 14th residue (Val-Lys-Asn-Arg-Asn-Glu-Ile-Pro-Glu-Ala-Leu-Thr-?-Ile), was identical to 10 of the 13 amino acids of the N-terminal sequence of PepF1 (identical residues underlined) (33), which suggested the existence of two homologous oligopeptidases in strain NCDO763. At the DNA level, hybridization with a *pepF1* probe under stringent conditions revealed a single 4-kb and 3.2-kb *Hin*dIII band in strains NCDO763 and IL1403, respectively. No band was detected in TIL75 (NCDO763 cured of the 55-kb lactose-proteinase plasmid) or in the plasmid-free MG1363 strain. This confirmed that *pepF1* is located on the lactose-proteinase plasmid of strain NCDO763 and showed that it is located on the chromosome of the plasmid-free IL1403 strain. We have named the plasmid copy of *pepF1* *pepF1p* and the chromosomal copy *pepF1c. pepF1p* was further mapped by pulsed-field gel electrophoresis to the *Sma*2 and *Apa*11 DNA restriction fragment of strain IL1403. Under less stringent conditions, a 1.6-kb band of low intensity could be visualized in strains NCDO763 and MG1363, but another band in addition to the 3.2-kb fragment could not be detected in strain IL1403 (Fig. 1). This implied the existence of a second gene, homologous but nonidentical to *pepF1*, located on the chromosome of strain NCDO763 but absent in IL1403. We therefore decided to characterize the second *pepF* gene, named *pepF2.*

***pepF2* is included in an operon.** A 1.1-kb fragment of the *pepF2* gene was amplified by PCR with oligonucleotides deduced from the N-terminal sequence of PepF2 and from the active site of PepF1. The nucleotide sequence of the fragment showed 80% identity with the corresponding nucleotide sequence of *pepF1* and encoded a peptide 84% identical to the corresponding part of PepF1. From these results, we concluded that a second oligopeptidase gene, *pepF2*, homologous to *pepF1*, was indeed present on the chromosome of strain NCDO763. To sequence the entire *pepF2* gene and its flanking sequences, we cloned the 1.1-kb fragment in an integrative plasmid to produce pTIL35, which was used to transform *L. lactis*. After integration into the chromosome and appropriate digestions, two fragments containing either the sequence upstream or downstream of the 1.1-kb fragment were self-ligated to generate pTIL37 and pTIL36, respectively. A total of 4,680 bp were sequenced from pTIL34, pTIL36, and pTIL37, in which four open reading frames (ORF1 through ORF4) were detected by the Genmark program (Fig. 2), 3A, and 4).

The presence of a potential lactococcal promoter (P1) upstream of ORF2 and of a potential terminator downstream of ORF4 suggested that ORF2, *pepF2*, and ORF4 are organized in an operon. To test this hypothesis, we measured the mRNA spanning this region with a *pepF2* probe and detected a small amount of 3.7-kb transcript, whose size fits well with the predicted size (3.6 kb) of the operon. A second potential −10 extended promoter sequence (P2) (41) was found upstream of *pepF2*, which would produce a 2.6-kb transcript (Fig. 4). We could not clearly state if the signal detected at this position corresponded to this transcript because it would comigrate with the abundant 23S RNA, which often produces artifactual bands. Consequently, P1 and P2 were cloned upstream of the luciferase gene in *L. lactis* and their expression was tested during exponential growth in milk and in M17 medium. At low copy number, the expression of the *lux* gene under the control of P1 was not detectable, whereas the same gene under the control of P2 is expressed in M17 medium and in milk (Table 2). At high copy number, with P2, the luciferase activity was saturating and the growth rate of strain TIL247 was reduced. This indicated that P2 is much stronger than P1 under the conditions used. However, the activity of the P1 promoter was evidenced in the high-copy-number plasmid (Table 2).

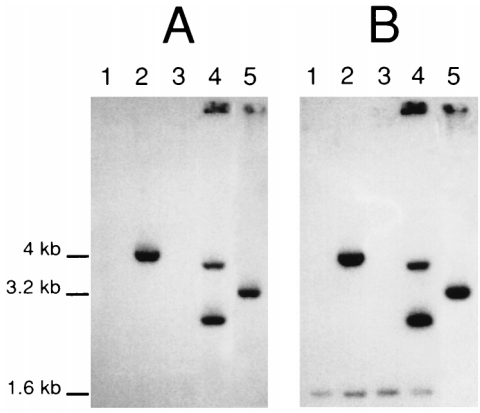

FIG. 1. Southern hybridization of *Hin*dIII-digested total genomic DNAs from several *L. lactis* strains using a 1.4-kb fragment of *pepF1* as a probe under stringent (50% formamide) (A) and nonstringent (20% formamide) (B) conditions. Lanes 1, *L. lactis* subsp. *cremoris* 1363; lanes 2, *L. lactis* subsp. *cremoris* NCDO763; lanes 3, NCDO763 *prt lac*; lanes 4, NCDO763 *pepF1*; lanes 5, *L. lactis* subsp. *lactis* IL1403.

We found no significant homology between the protein encoded by ORF1 and other proteins in the databases. The protein encoded by ORF2 presents a weak overall homology (15% identity, 39% similarity) with the tomato protein vsf-1, which is involved in the regulation of vascular system-specific gene expression (46a).

PepF2 is homologous (80% identity) to PepF1 (33). Both the zinc-binding motif (91% identity between the two PepF proteins for 45 amino acids) and the area showing homology with creatine and arginine kinases found for PepF1 are well-conserved, also for PepF2. No indication of different locations and functions for the two PepF oligopeptidases was found within the sequences.

ORF4 displays a significant homology with several methyltransferases from various origins (plants, mammals, and bacteria). The highest homologies were found with bacterial *O*-methyltransferases from *Synechocystis* (31% identity, 55% similarity) (21) and *Streptomyces* (27% identity, 56% similarity) (16), caffeoyl-coenzyme A 3-*O*-methyltransferases from plants (26% identity, 52% similarity) (28), and catechol-*O*-methyltransferases from mammals (27% identity, 55% similarity) (2). ORF4 was also related, though to a lesser extent, to protein L-isoaspartate-*O*-methyltransferases (10, 19).

Recombination events around the plasmid *pepF1* gene. Our results revealed the existence of two homologous genes, *pepF1* and *pepF2*, in *L. lactis*. Hybridization experiments showed first that *pepF1* is present both on the lactose-proteinase plasmid of strain NCDO763 (*pepF1p*) and on the chromosome of strain IL1403 (*pepF1c*) and second that *pepF2* is located on the chromosome of strain NCDO763. The high homology observed between these genes suggested that they are evolutionarily related. To evaluate the relationship between them and the

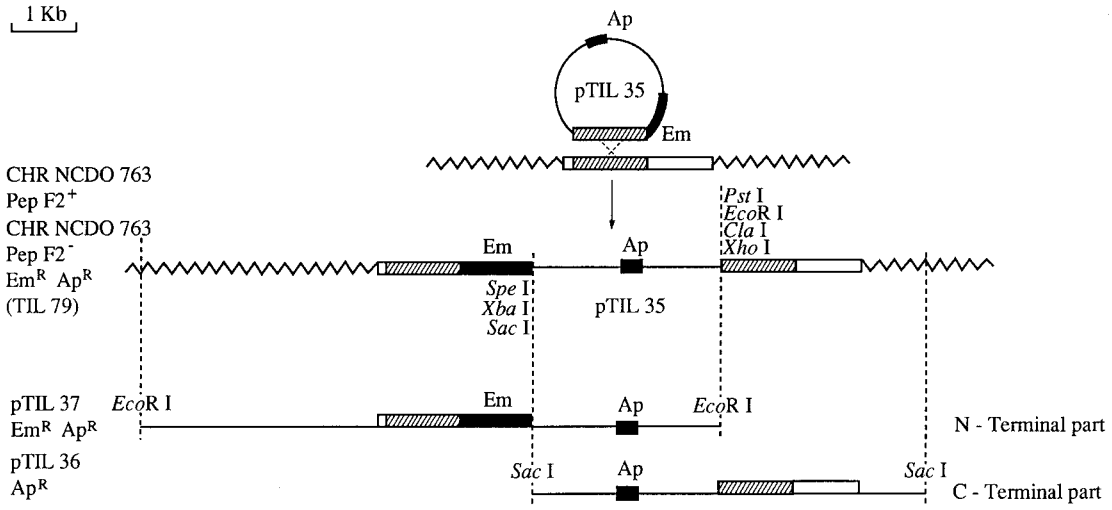

FIG. 2. Schematic representation of Campbell-like integration of pTIL35 (pBluescript SK^+ containing the erythromycin resistance gene of pIL253 and a 1.1-kb fragment of the *pepF2* gene) into the chromosomal *pepF2* gene of *L. lactis* subsp. *cremoris* NCDO763. pTIL36 and pTIL37 were obtained after digestion by *Sac*I and *Eco*RI, respectively, and were used to sequence the fragments upstream and downstream of *pepF2*.

possibility that one was derived from the other, we sequenced larger regions upstream and downstream of the *pepF1p* and *pepF1c* genes in strains NCDO763 and IL1403. Two striking observations were made. First, the sequences flanking *pepF1c* in strain IL1403 are 80% identical to those flanking *pepF2* in strain NCDO763. Second, the sequences flanking *pepF1c* and *pepF1p* are identical from base 339 of ORF2 to base 113 of ORF4 (Fig. 3B and C; Fig. 5). Beyond these limits, the sequences carried on the plasmid of strain NCDO763 were totally different from those located on the chromosomes of strains IL1403 and NCDO763. ORF2 and ORF4 are therefore truncated and the operon is incomplete on the plasmid of strain NCDO763. Moreover, the −10 extended P2 promoter is present on the plasmid.

A sequence identical to IS*904* (37) is present on the plasmid, 1.4 kb upstream of the truncated ORF2. Downstream of the

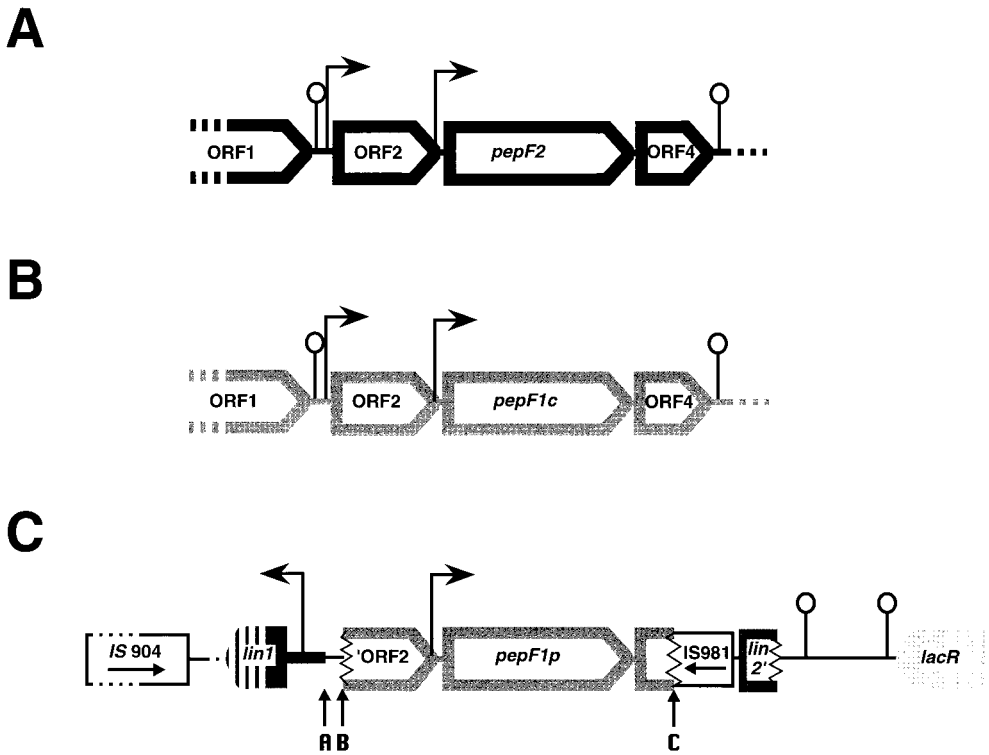

FIG. 3. Organization of the sequences flanking the *pepF* genes on the chromosomes of strains NCDO763 (A) and IL1403 (B) and the lactose-proteinase plasmid of strain NCDO763 (C). ORFs (thick arrows), IS elements (boxes), promoters (broken arrows), terminators (lollipops), and junctions between truncated elements (broken lines) are shown. Junctions A to C in panel C correspond to the sequences presented in Fig. 5. Regions which were not sequenced (hatched regions) are also indicated.

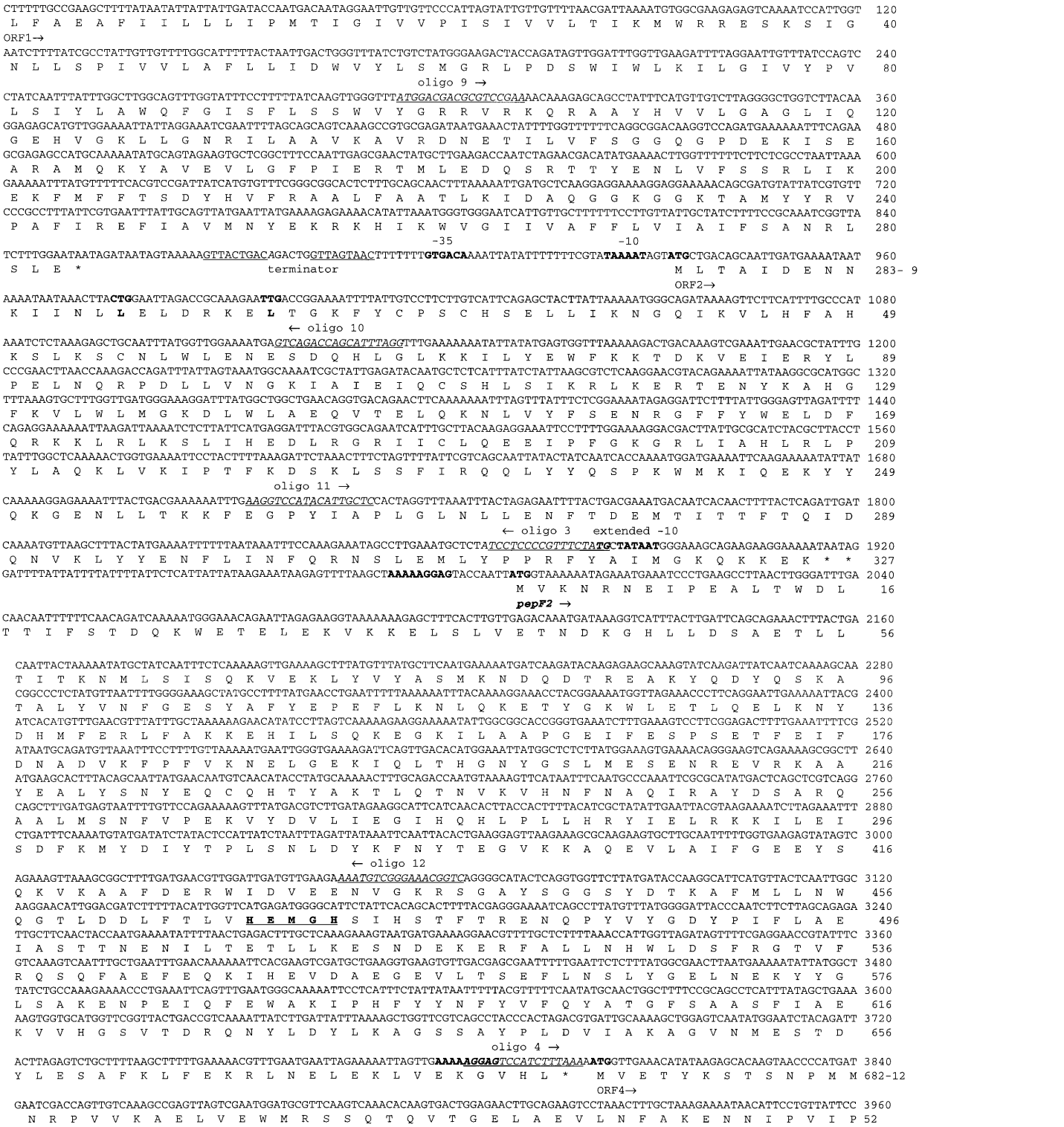

FIG. 4. Nucleotide and deduced amino acid sequences of ORF1, ORF2, ORF3 (*pepF2*), and ORF4 from the chromosome of *L. lactis* subsp. *cremoris* NCDO763. The numbers to the right refer to nucleotides (top lines) and amino acid residues (bottom lines). The putative −10 and −35 regions and ribosome binding sites (boldface), stop codons (asterisks), putative terminators (underlined), and the zinc-binding motif of PepF2 (underlined and boldface) are indicated.

TABLE 2. Luciferase expression under the control of promoter sequences P1 and P2 found upstream of *pepF2* at high and low copy numbers in *L. lactis*[a]

Promoter	Luciferase activity (lx/10^7 CFU) in the indicated medium	
	Milk	M17
Low copy number		
None (control)	NS[b]	NS
P1	NS	NS
P2	1,550	4,750
High copy number		
None (control)	25	38
P1	110	275
P2	(34,860)[c]	(326,920)[c]

[a] Luciferase activities were measured during growth in milk or M17 medium when the cell concentration reached 10^8 CFU/ml.
[b] NS, value obtained not significant compared to the background level.
[c] Luciferase activities in these experiments were saturating and thus probably underestimated.

insertion element (IS), we found a sequence almost identical to that found downstream of the terminator of the *L. lactis llaII* restriction/modification genes and a sequence coding for a DNA invertase from *L. lactis* (49) (Fig. 3C) homologous to *binR* from *Staphylococcus aureus* (40). A small DNA fragment of 68 bp is inserted between the latter sequence and the truncated ORF2; the junctions between these fragments are shown in Fig. 5A and B.

Downstream of *pepF1p*, the truncated sequence of ORF4 directly abuts a sequence identical to a truncated IS*981* (bases 591 to 1, from accession no. M33933) (35) (Fig. 5C). At 34 bp downstream of the IS, there is a sequence, *lin2*, that could code for an invertase, different from that encoded by *lin1* present upstream of *pepF1*. The 133-amino-acid protein encoded by *lin2* seems to be truncated at its C terminus, since the usual length for bacterial invertases is 180 amino acids. The region between *lin2* and *lacR* does not contain a potential ORF and does not show any homology to known sequences. This fragment contains two potential terminators (Fig. 3C).

Northern (RNA) blotting showed that the abundant *pepF1p* transcripts from the plasmid of strain NCDO763 were significantly smaller (3.2 kb) than the transcripts of the operons located on the chromosomes of strains NCDO763 and IL1403 (3.7 kb). The size of the *pepF1p* transcript corresponded to an mRNA initiated at P2 and ending at the potential terminator found downstream of the truncated *lin2′* sequence (Fig. 3C). Since we demonstrated that P2 was functional on a plasmid, its presence upstream of *pepF1p* on the lactose plasmid most probably induced the amplification of expression of *pepF*.

Function of *pepF* operon. The amplification of PepF activity by gene duplication in *L. lactis* raises the question of the function of PepF. Preliminary investigations of the growth of strains MG1363 (*pepF2*$^+$ Δ*pepF1*) and TIL80 (*pepF2* Δ*pepF1*) on different media revealed that the two strains grow at the

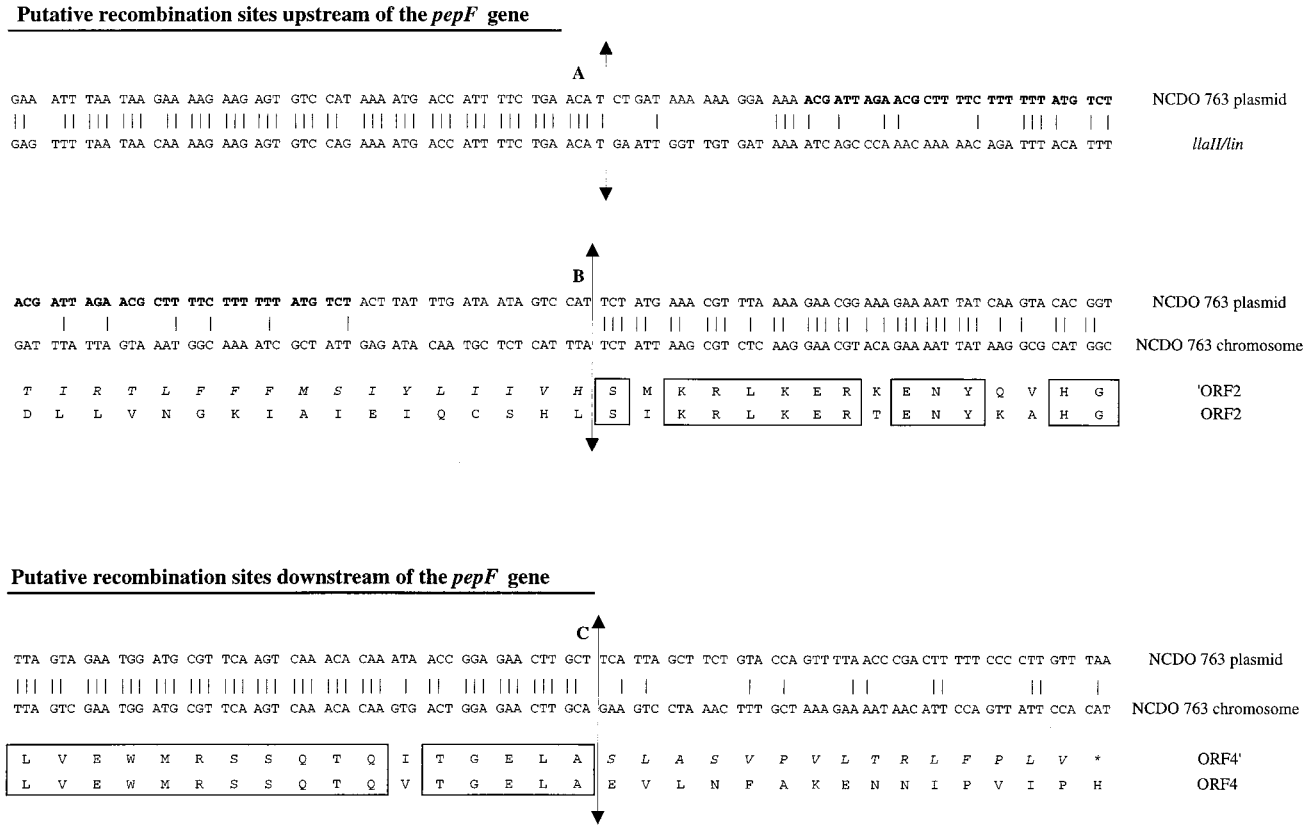

FIG. 5. Junctions corresponding to probable recombination sites on the plasmid of *L. lactis* subsp. *cremoris* NCDO763 upstream (A and B) and downstream (C) of the *pepF1* gene (see Fig. 3). The sequence of the plasmid in the first line is aligned with the *llaII/lin* sequence (accession no. U16027 and Z48180) and the chromosomal *pepF* operon as indicated. The ORFs corresponding to the sequences are indicated below. Identical amino acids (boxed), amino acids corresponding to totally different proteins (italics), and DNA sequences corresponding to the same part of the plasmid (boldface) are shown.

same rate in M17 rich medium. However, different phenotypes were evidenced in minimum medium containing only the seven amino acids essential to the two strains. TIL80, devoid of PepF, grew significantly more slowly (maximum growth rate [μ_{max}] = 0.12 h^{-1}) in this medium than its parental strain (μ_{max} = 0.14 h^{-1}). This suggests that PepF could be involved in protein turnover, since starvation for required nutrients leads to an elevated rate of the degradation of proteins (30). Similar results obtained with TIL70 (*pepF2*$^+$ *pepF1p*) and TIL79 (*pepF2 pepF1p*$^+$) compared to NCDO763 seem to indicate that the presence of both genes is beneficial in minimal medium.

DISCUSSION

We report the presence of two homologous genes coding for oligopeptidases in *L. lactis* subsp. *cremoris* NCDO763. One of the genes (*pepF1p*) is carried on the 55-kb plasmid which also contains the genes necessary for lactose and casein utilization. The other (*pepF2*) is located on the chromosome.

During a search for *pepF* genes in the chromosome of several lactococcal strains, we observed that *pepF2* is present on the chromosome of *L. lactis* subsp. *cremoris* NCDO763, while *pepF1* (identical to the plasmid copy of the gene in strain NCDO763) is present on the chromosome of *L. lactis* subsp. *lactis* IL1403. The two chromosomal genes are 80% identical, which corresponds to the divergence usually observed for *L. lactis* subsp. *lactis* and *cremoris* genes (16). The almost complete identity between the chromosomal gene of strain IL1403 and the plasmid gene of strain NCDO763 indicates that the two genes diverged more recently than the two subspecies and strongly supports the hypothesis of a recent gene transfer. The fact that *pepF* is part of a gene cluster in the chromosome of the two subspecies whereas it is restricted to an intact copy of *pepF1* on the plasmid suggests that *pepF1p* is a derivative of *pepF1c*. The precise mechanism which allowed this event is not known. However, the presence of IS*904* and IS*981* upstream and downstream, respectively, of *pepF1p* suggests the involvement of IS-directed mobilization of *pepF1* from the chromosome to the plasmid. Since the lactose-proteinase plasmid is conjugative (13), it could have then been transferred to different strains, such as *L. lactis* subsp. *cremoris* NCDO763. The sequences upstream and downstream of *pepF1p* possess several fragments identical to previously identified sequences. The direct linkages of these fragments are visible proof of numerous reorganizations that have occurred during the evolution of the lactose-proteinase plasmid.

Two duplicated *pepO* genes coding for homologous PepO oligopeptidases were recently shown to exist in lactococci (16a). Therefore, both the PepO and the PepF oligopeptidase genes are duplicated in *L. lactis*. The occurrence of such independent duplications of oligopeptidase genes suggests that these events are beneficial to the cell.

The first role attributed to gene duplication is amplification of the level of expression of a gene. This event can be fixed in a population by natural selection if the amplification confers a selective advantage. The best-known examples of amplification concern genes conferring a selective advantage by coding for resistance to antibiotics or drugs. The *tufA* and *tufB* genes of *E. coli* are another example of duplicated genes. The two genes are 98% identical, and their products are functionally indistinguishable. However, the TufA protein appears to be produced at a higher level (11, 31), which suggests that the duplication enables the cell to provide larger amounts of the translation elongation factor EF-Tu when demand is high (1). In the case of *pepF*, its presence on the lactose plasmid allows a fourfold increase in enzymatic activity in *L. lactis* subsp. *cremoris* NCDO763, which fits with the above hypothesis. The selective advantage that the strain derives from *pepF* overexpression remains to be determined.

Once retained on a plasmid, a duplicated gene can be assembled into new pathways to respond to environmental changes and participate in an important manner in the evolution of the genome. In the present case, the enzymatic properties of *pepF* genes seem to remain unchanged but the expression of the two genes is under the control of different promoters and in different backgrounds (the chromosome for *pepF2* and the plasmid for *pepF1p*). Studies of the regulation of *pepF* genes are in progress in our laboratory and will provide a better understanding of the function of PepF.

Although PepF is biochemically well characterized, its role in *L. lactis* remains unclear. A mutant devoid of PepF activity (TIL80) is fully viable in M17 rich medium. However, the slower growth of *pepF1* or *pepF2* mutants than of the parental strains in minimum medium suggests that this oligopeptidase could be involved in protein turnover. This hypothesis is reinforced by the fact that *pepF* is cotranscribed with a gene coding for a methyltransferase. Most of the methyltransferases showing homology to ORF4 are involved in defense and survival responses and metabolism of altered proteins (43). The latter role is attributed to L-isoaspartyl methyltransferases, which help the cells to repair or degrade the altered proteins that accumulate in aging cells and limit their viability (10, 24, 26). A similar function may be postulated for ORF4 and the chromosome-encoded PepF. Lastly, a gene coding for a protein showing 30% identity with PepF was detected in the genome of *Mycoplasma genitalium* (9). This genome is one of the smallest known for a self-replicating organism and probably possesses a minimal functional gene set. This suggests that PepF oligopeptidases could play an important role in the cells. Such a fundamental function was already proposed for the homologous oligopeptidase from *Salmonella typhimurium* (OpdA), which is involved in protein turnover during carbon deficiency and in the degradation of signal peptides (48).

The presence of an isolated copy of *pepF1p* (without the methyltransferase) suggests that the plasmid and chromosome *pepF* genes have different functions. The linkage of *pepF1p* with the genes necessary for optimal growth in milk would indicate a role in peptide assimilation. However, experiments testing growth in milk showed that *pepF* amplification does not confer a significant advantage during exponential growth. Alternatively, the two *pepF* genes may still have the same function and a sufficient amount of the methyltransferase from the chromosome could be expressed while PepF activities could be limiting in milk cultures.

ACKNOWLEDGMENTS

We thank Marie-Christine and Alain Chopin for their help in initiating the work, Patricia Anglade for protein sequencing, and Costa Anagnostopoulos and Dusko Ehrlich for critically reading the manuscript.

REFERENCES

1. **An, G., and J. D. Friesen.** 1980. The nucleotide sequence of *tuf*B and four nearby tRNA structural genes of *E. coli*. Gene **12:**33–39.
2. **Bertocci, B., V. Miggiano, M. Da Prada, Z. Dembic, H.-W. Lahm, and P. Malherbe.** 1991. Human catechol-*O*-methyltransferase: cloning and expression of the membrane associated form. Proc. Natl. Acad. Sci. USA **88:**1416–1420.
3. **Borodovsky, M., and J. McIninch.** 1993. Genmark: parallel gene recognition for both DNA strands. Computer Chem. **17:**123–133.
4. **Chopin, A., M.-C. Chopin, A. Moillot-Batt, and P. Langella.** 1984. Two plasmid-determined restriction and modification system in *Streptococcus lactis*. Plasmid **11:**260–263.

5. **Cocaign-Bousquet, M., C. Garrigues, L. Novak, N. D. Lindley, and P. Loubiere.** 1995. Rational development of a simple synthetic medium for the sustained growth of *Lactococcus lactis*. J. Appl. Bacteriol. **79:**108–116.
6. **Davies, F. L., H. M. Underwood, and M. J. Gasson.** 1981. The role of plasmid profiles for strain identification in lactic streptococci and the relationship between *Streptococcus lactis* 712, ML3 and C2. J. Appl. Bacteriol. **51:**325–337.
7. **Devereux, J., P. Haberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
8. **De Vos, W. M., I. J. Boerrigter, R. J. van Rooyen, B. Reiche, and W. Hengstenberg.** 1990. Characterization of the lactose-specific enzymes of the phosphotransferase system in Lactococcus lactis. J. Biol. Chem. **265:**22554–22560.
9. **Fraser, C. M., J. D. Gocayne, O. White, M. D. Adams, R. A. Clayton, R. D. Fleischmann, C. J. Bult, A. R. Kerlavage, G. Sutton, J. M. Kelly, J. L. Fritchman, J. F. Weidman, K. V. Small, M. Sandusky, J. L. Fuhrmann, D. T. Nguyen, T. R. Utterback, D. M. Saudek, C. A. Phillips, J. M. Merrick, J.-F. Tomb, B. A. Dougherty, K. F. Bott, P.-C. Hu, T. S. Lucier, S. N. Peterson, H. O. Smith, C. A. Hutchison III, and J. C. Venter.** 1995. The minimal gene complement of *Mycoplasma genitalium*. Science **270:**397–403.
10. **Fu, J. C., L. Ding, and S. Clarke.** 1991. Purification, gene cloning and sequence analysis of an L-isoaspartyl protein carboxyl methyltransferase from *E. coli*. J. Biol. Chem. **266:**14562–14572.
11. **Furano, A. V.** 1977. The elongation factor Tu coded by the *tufA* gene of *E. coli* K-12 is almost identical to that coded by the *tufB* gene. J. Biol. Chem. **252:**57–62.
12. **Gasson, M. J.** 1983. Plasmid complements of *Streptococcus lactis* NCDO712 and other lactic streptococci after protoplast-induced curing. J. Bacteriol. **154:**1–9.
13. **Gasson, M. J.** 1990. In vivo genetic systems in lactic acid bacteria. FEMS Microbiol. Rev. **87:**43–60.
14. **Gibson, T. J.** 1984. Studies on the Epstein Barr virus genome. Ph.D. Thesis. University of Cambridge, Cambridge, England.
15. **Godon, J.-J., C. Delorme, S. D. Ehrlich, and P. Renault.** 1992. Divergence of genomic sequences between *Lactococcus lactis* subsp. *lactis* and *Lactococcus lactis* subsp. *cremoris*. Appl. Environ. Microbiol. **58:**4045–4047.
16. **Hara, O., and C. R. Hutchinson.** 1992. A macrolide 3-*O*-acyltransferase gene from the midecamycin-producing species *Streptomyces mycarofaciens*. J. Bacteriol. **174:**5141–5144.

16a. **Hellendoorn, M., and I. Mierau.** Personal communication.

17. **Holo, H., and I. F. Nes.** 1989. High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media. Appl. Environ. Microbiol. **55:**3119–3123.

17a. **Juillard, V.** Unpublished results.

18. **Juillard, V., D. Le Bars, E. R. S. Kunji, W. N. Konings, J.-C. Gripon, and J. Richard.** 1995. Oligopeptides are the main source of nitrogen for *Lactococcis lactis* during growth in milk. Appl. Environ. Microbiol. **61:**3024–3030.
19. **Kagan, R. M., and S. Clarke.** 1995. Protein L-isoaspartyl methyltransferase from the nematode *Caenorhabditis elegans*: genomic structure and substrate specificity. Biochemistry **34:**10794–10806.
20. **Kaneko, T., A. Tanaka, S. Sato, H. Kotani, T. Sazuka, N. Miyajima, M. Sugiura, and S. Tabata.** 1995. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome. DNA Res. **2:**153–166.
21. **Kiwaki, M., H. Ikemura, M. Shimizu-Kadota, and A. Hirashima.** 1989. Molecular characterization of a cell wall-associated proteinase gene from *Streptococcus lactis* NCDO763. Mol. Microbiol. **3:**359–369.
22. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
23. **Le Bourgeois, P., M. Lautier, M. Mata, and P. Ritzenthaler.** 1992. Physical and genetic map of the chromosome of *Lactococcus lactis* subsp. *lactis* IL1403. J. Bacteriol. **174:**6752–6762.
24. **Li, C., and S. Clarke.** 1992. A protein methyltransferase specific for altered aspartyl residues is important in *E. coli* stationary phase survival and heat-shock resistance. Proc. Natl. Acad. Sci. USA **89:**9885–9889.
25. **Loureiros Dos Santos, A. L., and A. Chopin.** 1987. Shotgun cloning in *Streptococcus lactis*. FEMS Microbiol. Lett. **42:**209–212.
26. **Lowenson, J. D., and S. Clarke.** 1992. Recognition of D-aspartyl residues in polypeptides by the erythrocyte L-isoaspartyl/D-aspartyl protein methyltransferase. J. Biol. Chem. **267:**5985–5995.
27. **Matsudaira, P.** 1987. Sequence from picomoles quantities of protein electroblotted onto polyvinylidene difluoride membrane. J. Biol. Chem. **262:**10035–10038.
28. **Meng, H., and W. H. Campbell.** 1995. Cloning of *Aspen* (*Populus tremuloides*) *xulem* caffeoyl-CoA 3-*O*-methyltransferase. Plant Physiol. **108:**1749.
29. **Mierau, I., P. S. T. Tan, A. J. Haandrikman, B. Mayo, J. Kok, K. J. Leenhouts, W. N. Konings, and G. Venema.** 1993. Cloning and sequencing of the gene for a lactococcal endopeptidase, an enzyme with sequence similarity to mammalian enkephalinase. J. Bacteriol. **175:**2087–2096.
30. **Miller, C. M.** 1987. Protein degradation and proteolytic modification, p. 680–691. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
31. **Miller, D. L.** 1978. A comparison of the activities of the products of the two genes for elongation factor Tu. Mol. Gen. Genet. **159:**57–62.
32. **Monnet, V.** 1995. Oligopeptidases from *Lactococcus lactis*. Methods Enzymol. **248:**579–592.
33. **Monnet, V., M. Nardi, A. Chopin, M.-C. Chopin, and J.-C. Gripon.** 1994. Biochemical and genetic characterization of PepF, an oligopeptidase from *Lactococcus lactis*. J. Biol. Chem. **269:**32070–32076.
34. **Otto, R., B. Ten Brink, H. Veldkamp, and W. N. Konings.** 1983. The relation between growth rate and electrochemical proton gradient of Streptococcus cremoris. FEMS Microbiol. Lett. **16:**69–74.
35. **Polzin, K. M., and L. L. McKay.** 1991. Identification, DNA sequencing, and distribution of IS*981*, a new, high-copy-number insertion sequence in lactococci. Appl. Environ. Microbiol. **57:**734–743.
36. **Poolman, B., E. R. S. Kunji, A. Hagting, V. Juillard, and W. N. Konings.** 1995. The proteolytic pathway of *Lactococcus lactis*. J. Appl. Bacteriol. Symp. Suppl. **79:**65S–75S.
37. **Rauch, P. J. G., M. M. Beerthuyzen, and W. M. de Vos.** 1990. Nucleotide sequence of IS904 from *Lactococcus lactis* subsp. *lactis* strain NIZO R5. Nucleic Acids Res. **18:**4253–4254.
38. **Rawlings, N. D., and A. J. Barrett.** 1995. Evolutionary families of metallopeptidases. Methods Enzymol. **248:**183–228.
39. **Renault, P., G. Corthier, N. Goupil, C. Delorme, and S. D. Ehrlich.** 1996. Plasmid vectors for Gram-positive bacteria switching from high to low copy number. Gene **183:**175–182.
40. **Rowland, S. J., and K. G. H. Dyke.** 1989. Characterization of the staphylococcal beta-lactamase transposon Tn552. EMBO J. **8:**2761–2773.
41. **Sabelnikov, A. G., B. Greenberg, and S. A. Lacks.** 1995. An extended −10 promoter alone directs transcription of the *DpnII* operon of *Streptococcus pneumoniae*. J. Mol. Biol. **250:**144–155.
42. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
43. **Schmitt, D., A.-E. Pakusch, and U. Matern.** 1991. Molecular cloning, induction and taxonomic distribution of caffeoyl-CoA 3-*O*-methyltransferase, an enzyme involved in disease resistance. J. Biol. Chem. **266:**17416–17423.
44. **Simon, D., A. Rouault, and M.-C. Chopin.** 1985. Protoplast formation of group N streptococci with cryptic plasmids. FEMS Microbiol. Lett. **26:**239–241.
45. **Simon, D., and A. Chopin.** 1988. Construction of a vector plasmid family for molecular cloning in *Streptococcus lactis*. Biochimie **70:**559–566.
46. **Terzaghi, B. E., and W. E. Sandine.** 1975. Improved medium for streptococci and their bacteriophages. Appl. Microbiol. **29:**87–813.

46a. **Torres-Schumann, S., C. Ringli, D. Heierli, N. Amrheim, and B. Keller.** 1996. In vitro binding of the tomato bZIP transcriptional activator VSF-1 to a regulatory element that controls xylem-specific gene expression. Plant J. **9:**283–296.

47. **Tynkkynen, S., G. Buist, E. Kunji, J. Kok, B. Poolman, G. Venema, and A. Haandrikman.** 1993. Genetic and biochemical characterization of the oligopeptide transport system of *Lactococcus lactis*. J. Bacteriol. **175:**7523–7532.
48. **Vimr, E. R., L. Green, and C. G. Miller.** 1983. Oligopeptidase-deficient mutants of *Salmonella typhimurium*. J. Bacteriol. **153:**1259–1265.
49. **Waterfield, N. R., R. W. Le Page, P. W. Wilson, and J. M. Wells.** 1995. The isolation of lactococcal promoters and their use in investigating bacterial luciferase synthesis in *Lactococcus lactis*. Gene **165:**9–15.